2005 Multivariate Prototypes

jw### refers to page numbers in R.A. Johnson & D.W. Wichern *Applied Multivariate Statistical Analysis* 4th Edition Prentice Hall 1998. Prototype sheets are arranged in this file by jw chapters with supplementary materials.

Chapter 01 – Multivariate Data
- **jw010 – Using MathCad to Evaluate Data from Disk Files**
- **jw031 – Euclidean and Statistical Distance**

Chapter 02 – Matrix Algebra
- **jw060 – Linear Transformations**
- **jw066 – Points in Space, the Linear Transformation, and Use of the Variance/Covariance Matrix to Calculate Distances**
- **jw084 – Matrix Algebra Toolbox for Multivariate Methods**
- **jw098 – Linear Transformations and the Calculation of Eigenvectors and Eigenvalues**
- **jw099 – Example Calculation of Eigenvectors & Eigenvalues**

Chapter 03 – Sample Geometry
- **jw139 – Basic Setup of Multivariate Methods**
- **jw142 – Linear Combinations of Variables**

Chapter 04 – Multivariate Normal Distribution
- **jw155 – The Multivariate Normal Distribution and the Variance-Covariance Matrix**
- **jw179 – Constructing Q-Q & χ^2 Plot**
- **jw749 – Statistical Distributions and Finding Critical Values**

Chapter 05 – Hotelling's T^2
- **jw214 – Hotelling's T^2**
- **jw214A - Hotelling's T^2 Verifying Examples**
- **stattest – Test Statistics for Univariate Data & Multivariate Data Compared**
- **SYSTAT Test Output – Example Run in SYSTAT**
- **t-tests – Univariate Means Tests**

Chapter 06 – MANOVA & More Hotelling
- **ANOVA 1 – Analysis of Variance – Univariate Case**
- **ANOVA 2 – Univariate Case Example**
- **Two-Way ANOVA – Two Way ANOVA Fixed Effects – Univariate Case**
- **jw272 – Hotelling's T^2 – Paired Design**
- **jw278 – Hotelling's T^2 Profile Analysis for One Population (= Repeated Measures Design with Contrasts)**
- **jw283 – Hotelling's T^2 – Two Populations Small Sample Sizes, Equal Σs**
- **jw283A – Hotelling's T^2 Verifying Example**
- **jw290 – Hotelling's T^2 – Two Populations, Unequal Σs for Large Samples Only**
- **jw290 – Hotelling's T^2 Verifying Example**
- **jw298 – MANOVA – One Way**
- **jw298A – MANOVA One Way – Verifying Example**
- **jw298B – MANOVA One Way – Calculating Standard Discriminant (Canonical) Coefficients in R.C. Mencher Methods of Multivariate Analysis 4th ed. p. 306-307**
- **jw313 – Two Way MANOVA – Fixed Effects**
- **jw318 – Hotelling's T^2 Profile Analysis for Two Groups**
- **jw318A – Hotelling's T^2 Profile Analysis Verifying Example**

jw318B – Hotelling's T^2 Profile Analysis Verifying Calculations in Rencher p. 141-165
Hotellings Conversion – Conversion to Hotelling's T^2 from Typical Computer Output of MANOVA
Chapter 07 – Regression & General Linear Model (GLM)
jw294 – Comparing ANOVA and Regression - One Way Case
jw368 – Classical Linear Regression
jw372 – ANOVA as a Regression – Example Calculations
jw372A – Comparing ANOVA and Regression – Two Way Case
jw376 – Classical Linear Regression, Confidence Invervals and Prediction
jw383 – Multivariate and Univariate Multiple Linear Regression
Chapter 08 – Principal Components Analysis (PCA) & PCO
jw439 – Principal Components Analysis (PCA)
PCA & SVD – PCA and Singular Value Decomposition
PCO – Principal Coordinates Analysis
PCO Map Example – Use of PCA in mapping the cities of New York State
Chapter 11 – Discriminant Functions
jw606 – Classification and Discriminant Functions Using Linear and Quadratic ECM Rules
jw611 – Classification and Discriminant Functions Using Linear Fisher's Discriminant Functions for Two Groups
jw631 – Fisher's Sample Linear Discriminants for Several Populations
jw631A – Fisher's Discriminants – Verifying Example 11.13
jw631B – Fisher's Discriminants – Verifying Example in Table 11.7 Salmon Data
jw631C – Fisher's Discriminants – Verifying Table 11.7 Crude Oil Data
Chapter 12 – Clustering & Correspondence Analysis
jw670 – Common Similarity/Distance Measures
jw683 – Clustering Data Charts for Hand Analysis Using Single Link, Complete Link, UPGMA, & WPGMA
jw715 – Correspondence Analysis
jw715A – Correspondence Analysis – Verifying Example 12.18

ORIGIN ≡ 1

USING MATHCAD TO EVALUATE DATA FROM DISK FILES
jw10.mcd

Prepared by:
Wm Stein

X := READPRN("\DATA\T1-2.DAT")

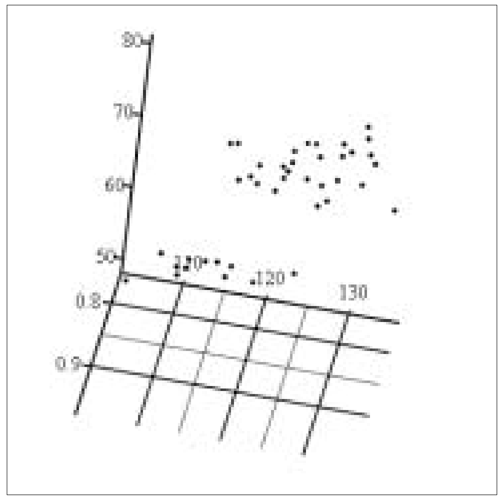
X =

	1	2	3
1	0.801	121.41	70.42
2	0.824	127.7	72.47
3	0.841	129.2	78.2
4	0.816	131.8	74.89
5	0.84	135.1	71.21
6	0.842	131.5	78.39
7	0.82	126.7	69.02
8	0.802	115.1	73.1
9	0.828	130.8	79.28
10	0.819	124.6	76.48
11	0.826	118.31	70.25
12	0.802	114.2	72.88
13	0.81	120.3	68.23
14	0.802	115.7	68.12
15	0.832	117.51	71.62
16	0.796	109.81	53.1

n := rows(X) n = 41

p := cols(X) p = 3

$X_{1,1} = 0.801$



$\left(X^{\langle 1\rangle}, X^{\langle 2\rangle}, X^{\langle 3\rangle}\right)$

Index variables:

i := 1 .. n j := 1 .. n k := 1 .. p

Means:

$xbar_1 := \frac{1}{n}\cdot\sum_j X_{j,1}$ $xbar_1 = 0.812$

$xbar_2 := \frac{1}{n}\cdot\sum_j X_{j,2}$ $xbar_2 = 120.953$

$xbar_3 := \frac{1}{n}\cdot\sum_j X_{j,3}$ $xbar_3 = 67.723$

Variances:

$\sigma_{1,1} := \frac{1}{n}\cdot\sum_j \left(X_{j,1} - xbar_1\right)^2$ $\sigma_{1,1} = 1.234 \times 10^{-3}$

$\sigma_{2,2} := \frac{1}{n}\cdot\sum_j \left(X_{j,2} - xbar_2\right)^2$ $\sigma_{2,2} = 57.874$

$\sigma_{3,3} := \frac{1}{n}\cdot\sum_j \left(X_{j,3} - xbar_3\right)^2$ $\sigma_{3,3} = 93.519$

Covariances:

$\sigma_{1,2} := \frac{1}{n}\cdot\sum_j \left(X_{j,1} - xbar_1\right)\cdot\left(X_{j,2} - xbar_2\right)$ $\sigma_{1,2} = 0.164$

$\sigma_{2,3} := \frac{1}{n}\cdot\sum_j \left(X_{j,2} - xbar_2\right)\cdot\left(X_{j,3} - xbar_3\right)$ $\sigma_{2,3} = 59.505$

$\sigma_{1,3} := \frac{1}{n}\cdot\sum_j \left(X_{j,1} - xbar_1\right)\cdot\left(X_{j,3} - xbar_3\right)$ $\sigma_{1,3} = 0.22$

2006 jw031

ORIGIN ≡ 1

EUCLIDEAN AND STATISTICAL DISTANCE
jw031.mcd

Prepared by:
Wm Stein

Using data in jw Table 1.2 but deleting outlier specimen 25:

X := READPRN("\DATA\T1-2-25.DAT")

X =

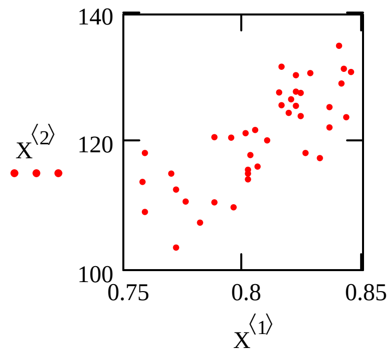
	1	2	3
1	0.801	121.41	70.42
2	0.824	127.7	72.47
3	0.841	129.2	78.2
4	0.816	131.8	74.89
5	0.84	135.1	71.21
6	0.842	131.5	78.39
7	0.82	126.7	69.02
8	0.802	115.1	73.1
9	0.828	130.8	79.28
10	0.819	124.6	76.48
11	0.826	118.31	70.25
12	0.802	114.2	72.88
13	0.81	120.3	68.23
14	0.802	115.7	68.12
15	0.832	117.51	71.62
16	0.796	109.81	53.1

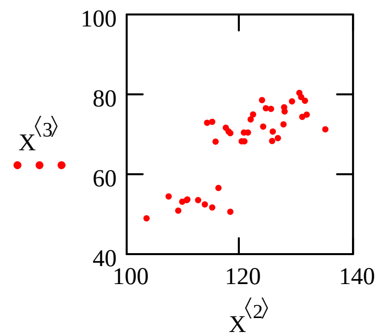
$n := \text{rows}(X)$ $\quad n = 40$

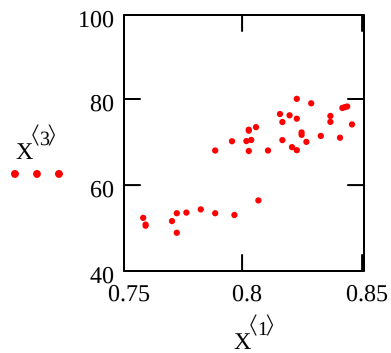
$p := \text{cols}(X)$ $\quad p = 3$

$X_{1,1} = 0.801$

Scatter Plots:







Index variables:

$i := 1..n \qquad j := 1..n \qquad k := 1..p$

Means:

$$x_{bar_1} := \frac{1}{n}\cdot\sum_j X_{j,1} \qquad x_{bar_1} = 0.808$$

$$x_{bar_2} := \frac{1}{n}\cdot\sum_j X_{j,2} \qquad x_{bar_2} = 120.825$$

$$x_{bar_3} := \frac{1}{n}\cdot\sum_j X_{j,3} \qquad x_{bar_3} = 67.614$$

Variances:

$$\sigma_{1,1} := \frac{1}{n}\cdot\sum_j \left(X_{j,1} - x_{bar_1}\right)^2 \qquad \sigma_{1,1} = 0.001$$

$$\sigma_{2,2} := \frac{1}{n}\cdot\sum_j \left(X_{j,2} - x_{bar_2}\right)^2 \qquad \sigma_{2,2} = 58.642$$

$$\sigma_{3,3} := \frac{1}{n}\cdot\sum_j \left(X_{j,3} - x_{bar_3}\right)^2 \qquad \sigma_{3,3} = 95.366$$

Covariances:

$$\sigma_{1,2} := \frac{1}{n}\cdot\sum_j \left(X_{j,1} - x_{bar_1}\right)\cdot\left(X_{j,2} - x_{bar_2}\right) \qquad \sigma_{1,2} = 0.147$$

$$\sigma_{2,3} := \frac{1}{n}\cdot\sum_j \left(X_{j,2} - x_{bar_2}\right)\cdot\left(X_{j,3} - x_{bar_3}\right) \qquad \sigma_{2,3} = 60.415$$

$$\sigma_{1,3} := \frac{1}{n}\cdot\sum_j \left(X_{j,1} - x_{bar_1}\right)\cdot\left(X_{j,3} - x_{bar_3}\right) \qquad \sigma_{1,3} = 0.207$$

2006 jw031

Pearson's product-moment correlation coefficients:

$$r_{1,2} := \frac{1}{n}\cdot\sum_j \frac{\sigma_{1,2}}{\sqrt{\sigma_{1,1}}\cdot\sqrt{\sigma_{2,2}}} \qquad r_{1,2} = 0.776 \qquad r_{1,1} := \frac{1}{n}\cdot\sum_j \frac{\sigma_{1,1}}{\sqrt{\sigma_{1,1}}\cdot\sqrt{\sigma_{1,1}}} \qquad r_{1,1} = 1$$

$$r_{1,3} := \frac{1}{n}\cdot\sum_j \frac{\sigma_{1,3}}{\sqrt{\sigma_{1,1}}\cdot\sqrt{\sigma_{3,3}}} \qquad r_{1,3} = 0.856 \qquad r_{2,2} := \frac{1}{n}\cdot\sum_j \frac{\sigma_{2,2}}{\sqrt{\sigma_{2,2}}\cdot\sqrt{\sigma_{2,2}}} \qquad r_{2,2} = 1$$

$$r_{2,3} := \frac{1}{n}\cdot\sum_j \frac{\sigma_{2,3}}{\sqrt{\sigma_{2,2}}\cdot\sqrt{\sigma_{3,3}}} \qquad r_{1,2} = 0.776 \qquad r_{3,3} := \frac{1}{n}\cdot\sum_j \frac{\sigma_{3,3}}{\sqrt{\sigma_{3,3}}\cdot\sqrt{\sigma_{3,3}}} \qquad r_{3,3} = 1$$

Basic Descriptive Statistics as Arrays:

$$X_{bar} := \begin{pmatrix} x_{bar_1} \\ x_{bar_2} \\ x_{bar_3} \end{pmatrix} \qquad X_{bar} = \begin{pmatrix} 0.808 \\ 120.825 \\ 67.614 \end{pmatrix} \qquad S_n := \begin{pmatrix} \sigma_{1,1} & \sigma_{1,2} & \sigma_{1,3} \\ \sigma_{1,2} & \sigma_{2,2} & \sigma_{2,3} \\ \sigma_{1,3} & \sigma_{2,3} & \sigma_{3,3} \end{pmatrix} \qquad S_n = \begin{pmatrix} 0.001 & 0.147 & 0.207 \\ 0.147 & 58.642 & 60.415 \\ 0.207 & 60.415 & 95.366 \end{pmatrix}$$

$$R_n := \begin{pmatrix} r_{1,1} & r_{1,2} & r_{1,3} \\ r_{1,2} & r_{2,2} & r_{2,3} \\ r_{1,3} & r_{2,3} & r_{3,3} \end{pmatrix} \qquad R_n = \begin{pmatrix} 1 & 0.776 & 0.856 \\ 0.776 & 1 & 0.808 \\ 0.856 & 0.808 & 1 \end{pmatrix}$$

Standardizing Matrix X:

$$X_{std_{j,k}} := \frac{\left(X_{j,k} - x_{bar_k}\right)}{\sqrt{\sigma_{k,k}}}$$

Note: Each variable is separately centered on its mean, and variance is adjusted to unity

$X_{std} =$

	1	2	3
1	-0.2771	0.07643	0.28736
2	0.64993	0.89781	0.49728
3	1.33512	1.09368	1.08404
4	0.32748	1.43321	0.74509
5	1.29482	1.86414	0.36826
6	1.37543	1.39403	1.1035
7	0.4887	0.76722	0.144
8	-0.2368	-0.74757	0.5618
9	0.81115	1.30262	1.19463
10	0.4484	0.49299	0.90791
11	0.73054	-0.32839	0.26995
12	-0.2368	-0.86509	0.53927
13	0.08565	-0.06852	0.0631
14	-0.2368	-0.66922	0.05184
15	0.97237	-0.43286	0.41024
16	-0.47863	-1.43836	-1.48622

Standardized Means:

$$xstd_{bar_1} := \frac{1}{n}\cdot\sum_j X_{std_{j,1}} \qquad xstd_{bar_1} = 0$$

$$xstd_{bar_2} := \frac{1}{n}\cdot\sum_j X_{std_{j,2}} \qquad xstd_{bar_2} = -6.689 \times 10^{-15}$$

$$xstd_{bar_3} := \frac{1}{n}\cdot\sum_j X_{std_{j,3}} \qquad xstd_{bar_3} = 3.403 \times 10^{-15}$$

Standardized Variances:

$$\sigma_{std_{1,1}} := \frac{1}{n}\cdot\sum_j \left(X_{std_{j,1}} - xstd_{bar_1}\right)^2 \qquad \sigma_{std_{1,1}} = 1$$

$$\sigma_{std_{2,2}} := \frac{1}{n}\cdot\sum_j \left(X_{std_{j,2}} - xstd_{bar_2}\right)^2 \qquad \sigma_{std_{2,2}} = 1$$

$$\sigma_{std_{3,3}} := \frac{1}{n}\cdot\sum_j \left(X_{std_{j,3}} - xstd_{bar_3}\right)^2 \qquad \sigma_{std_{3,3}} = 1$$

Standardized Covariances:

$$\sigma_{std_{1,2}} := \frac{1}{n}\cdot\sum_j \left(X_{std_{j,1}} - xstd_{bar_1}\right)\cdot\left(X_{std_{j,2}} - xstd_{bar_2}\right) \qquad \sigma_{std_{1,2}} = 0.776$$

$$\sigma_{std_{2,3}} := \frac{1}{n}\cdot\sum_j \left(X_{std_{j,2}} - xstd_{bar_2}\right)\cdot\left(X_{std_{j,3}} - xstd_{bar_3}\right) \qquad \sigma_{std_{2,3}} = 0.808$$

$$\sigma_{std_{1,3}} := \frac{1}{n}\cdot\sum_j \left(X_{std_{j,1}} - xstd_{bar_1}\right)\cdot\left(X_{std_{j,3}} - xstd_{bar_3}\right) \qquad \sigma_{std_{1,3}} = 0.856$$

Standardized Covariances in Matrix form:

Note: variance/covariances of stadardized data are the same as correlation coefficients from the raw data.

$$S_{std.n} := \begin{pmatrix} \sigma_{std_{1,1}} & \sigma_{std_{1,2}} & \sigma_{std_{1,3}} \\ \sigma_{std_{1,2}} & \sigma_{std_{2,2}} & \sigma_{std_{2,3}} \\ \sigma_{std_{1,3}} & \sigma_{std_{2,3}} & \sigma_{std_{3,3}} \end{pmatrix} \qquad S_{std.n} = \begin{pmatrix} 1 & 0.776 & 0.856 \\ 0.776 & 1 & 0.808 \\ 0.856 & 0.808 & 1 \end{pmatrix}$$

Standardized Pearson's product-moment correlation coefficients:

$$r_{std_{1,2}} := \frac{1}{n}\cdot\sum_j \frac{\sigma_{std_{1,2}}}{\sqrt{\sigma_{std_{1,1}}}\cdot\sqrt{\sigma_{std_{2,2}}}} \qquad r_{std_{1,2}} = 0.776$$

$$r_{std_{1,1}} := \frac{1}{n}\cdot\sum_j \frac{\sigma_{std_{1,1}}}{\sqrt{\sigma_{std_{1,1}}}\cdot\sqrt{\sigma_{std_{1,1}}}} \qquad r_{std_{1,1}} = 1$$

$$r_{std_{1,3}} := \frac{1}{n}\cdot\sum_j \frac{\sigma_{std_{1,3}}}{\sqrt{\sigma_{std_{1,1}}}\cdot\sqrt{\sigma_{std_{3,3}}}} \qquad r_{std_{1,3}} = 0.856$$

$$r_{std_{2,2}} := \frac{1}{n}\cdot\sum_j \frac{\sigma_{std_{2,2}}}{\sqrt{\sigma_{std_{2,2}}}\cdot\sqrt{\sigma_{std_{2,2}}}} \qquad r_{std_{2,2}} = 1$$

$$r_{std_{2,3}} := \frac{1}{n}\cdot\sum_j \frac{\sigma_{std_{2,3}}}{\sqrt{\sigma_{std_{2,2}}}\cdot\sqrt{\sigma_{std_{3,3}}}} \qquad r_{std_{1,2}} = 0.776$$

$$r_{std_{3,3}} := \frac{1}{n}\cdot\sum_j \frac{\sigma_{std_{3,3}}}{\sqrt{\sigma_{std_{3,3}}}\cdot\sqrt{\sigma_{std_{3,3}}}} \qquad r_{std_{3,3}} = 1$$

Standardized Correlation coefficients in Matrix form:

Note: Once the data are standardized, correlations are the same as variance/covariances!

$$R_{std.n} := \begin{pmatrix} r_{std_{1,1}} & r_{std_{1,2}} & r_{std_{1,3}} \\ r_{std_{1,2}} & r_{std_{2,2}} & r_{std_{2,3}} \\ r_{std_{1,3}} & r_{std_{2,3}} & r_{std_{3,3}} \end{pmatrix} \qquad R_{std.n} = \begin{pmatrix} 1 & 0.776 & 0.856 \\ 0.776 & 1 & 0.808 \\ 0.856 & 0.808 & 1 \end{pmatrix}$$

DISTANCES OF DATA POINTS IN X FROM A FIXED REFERENCE Q:

Squared Euclidean Distances:

$$Q := \begin{pmatrix} xbar_1 \\ xbar_2 \\ xbar_3 \end{pmatrix}$$

Pick a point of reference Q from which to measure distances to each point in the data:

Note: I choose to use the sample mean in all three variables in this example.

$$Q = \begin{pmatrix} 0.807875 \\ 120.82475 \\ 67.61375 \end{pmatrix}$$

Squared Euclidean Components:

$$Comp_{sq_{j,k}} := \left(X_{j,k} - Q_k\right)^2$$

Squared Euclidean Distances:

$$D_j := \sum_k Comp_{sq_{j,k}}$$

Note: In multivariate statistics often convenient to use square distances rather than distance order to avoid taking square r lot of the time.

$Comp_{sq} =$

	1	2	3
1	0.00005	0.34252	7.87504
2	0.00026	47.26906	23.58316
3	0.0011	70.14481	112.06869
4	0.00007	120.45611	52.94381
5	0.00103	203.78276	12.93301
6	0.00116	113.96096	116.12756
7	0.00015	34.51856	1.97754
8	0.00003	32.77276	30.09894
9	0.00041	99.50561	136.10139
10	0.00012	14.25251	78.61039
11	0.00033	6.32397	6.94981
12	0.00003	43.88731	27.73339
13	$4.51563\cdot10^{-6}$	0.27536	0.37976
14	0.00003	26.26306	0.25629
15	0.00058	10.98757	16.05004
16	0.00014	121.32472	210.64894

$D =$

	1
1	8.2176
2	70.85249
3	182.2146
4	173.39999
5	216.71681
6	230.08969
7	36.49625
8	62.87174
9	235.60741
10	92.86303
11	13.27411
12	71.62074
13	0.65513
14	26.51939
15	27.03819
16	331.9738

Squared Statistical Distances:

$$Q := \begin{pmatrix} xbar_1 \\ xbar_2 \\ xbar_3 \end{pmatrix}$$

Pick a point from which to measure distances to each point in the data:

$$Q = \begin{pmatrix} 0.807875 \\ 120.82475 \\ 67.61375 \end{pmatrix}$$

Squared Statisticcal Components:

$$SComp_{sq_{j,k}} := \frac{\left(X_{j,k} - Q_k\right)^2}{S_{n_{k,k}}}$$

Squared Statistical Distances:

$$SD_j := \sum_k SComp_{sq_{j,k}}$$

Now let's compare what happens when we use the standardized data

$SComp_{sq} =$

	1	2	3
1	0.07678	0.00584	0.08258
2	0.42241	0.80606	0.24729
3	1.78255	1.19614	1.17515
4	0.10724	2.05408	0.55517
5	1.67655	3.47501	0.13561
6	1.8918	1.94332	1.21771
7	0.23883	0.58863	0.02074
8	0.05607	0.55886	0.31562
9	0.65796	1.69682	1.42715
10	0.20106	0.24304	0.8243
11	0.53369	0.10784	0.07288
12	0.05607	0.74839	0.29081
13	0.00734	0.0047	0.00398
14	0.05607	0.44785	0.00269
15	0.94551	0.18737	0.1683
16	0.22909	2.06889	2.20885

$SD =$

	1
1	0.1652
2	1.47575
3	4.15384
4	2.71649
5	5.28717
6	5.05283
7	0.8482
8	0.93055
9	3.78193
10	1.26841
11	0.7144
12	1.09527
13	0.01601
14	0.50661
15	1.30117
16	4.50683

2006 jw031

DISTANCES OF DATA POINTS IN Xstd FROM A FIXED REFERENCE Q:

Squared Euclidean Distances:

$$Qstd := \begin{pmatrix} xstd_{bar_1} \\ xstd_{bar_2} \\ xstd_{bar_3} \end{pmatrix}$$

Pick a point Q from which to measure distances to each point in the data:

Note: Means here are zero.

$$Qstd = \begin{pmatrix} 0 \\ -6.689094 \times 10^{-15} \\ 3.402834 \times 10^{-15} \end{pmatrix}$$

Squared Euclidean Components:

$$Compstd_{sq_{j,k}} := \left(Xstd_{j,k} - Q_k\right)^2$$

$Compstd_{sq}$ =

	1	2	3
1	1.17717	14580.15799	4532.84246
2	0.02495	14382.47185	4504.61997
3	0.27799	14335.52816	4426.20213
4	0.23078	14254.34086	4471.41709
5	0.23711	14151.62745	4521.95603
6	0.32212	14263.69686	4423.61368
7	0.10187	14413.81028	4552.167
8	1.09134	14779.82847	4495.96442
9	0.00001	14285.53946	4411.49892
10	0.12922	14479.732	4449.66884
11	0.00598	14678.08315	4535.18681
12	1.09134	14808.41827	4498.98605
13	0.52161	14615.18384	4563.0897
14	1.09134	14760.78395	4564.61162
15	0.02706	14703.40736	4516.31128
16	1.65509	14948.26899	4774.80598

Squared Euclidean Distances:

$$Dstd_j := \sum_k Compstd_{sq}$$

$Dstd$ =

	1
1	19114.17761
2	18887.11677
3	18762.00827
4	18725.98873
5	18673.8206
6	18687.63265
7	18966.07915
8	19276.88422
9	18697.03839
10	18929.53006
11	19213.27594
12	19308.49565
13	19178.79516
14	19326.48691
15	19219.7457
16	19724.73006

Squared Statistical Distances:

$$Qstd := \begin{pmatrix} xstd_{bar_1} \\ xstd_{bar_2} \\ xstd_{bar_3} \end{pmatrix}$$

Pick a point from which to measure distances to each point in the data:

$$Qstd = \begin{pmatrix} 0 \\ -6.689094 \times 10^{-15} \\ 3.402834 \times 10^{-15} \end{pmatrix}$$

Squared Statisticcal Components:

$$SCompstd_{sq_{j,k}} := \frac{\left(Xstd_{j,k} - Q_k\right)^2}{Sn_{k,k}}$$

$SCompstd_{sq}$ =

	1	2	3
1	1912.36209	248.62822	47.53112
2	40.52816	245.25718	47.23518
3	451.60431	244.45667	46.4129
4	374.90573	243.07222	46.88702
5	385.19735	241.3207	47.41697
6	523.28953	243.23176	46.38576
7	165.49093	245.79157	47.73376
8	1772.91724	252.0331	47.14442
9	0.01742	243.60424	46.25872
10	209.92726	246.91571	46.65897
11	9.71629	250.29809	47.5557
12	1772.91724	252.52063	47.17611
13	847.37548	249.2255	47.84829
14	1772.91724	251.70834	47.86425
15	43.9587	250.72993	47.35778
16	2688.76013	254.90543	50.06834

Squared Statistical Distances:

$$SDstd_j := \sum_k SComp_{sq_{j,k}}$$

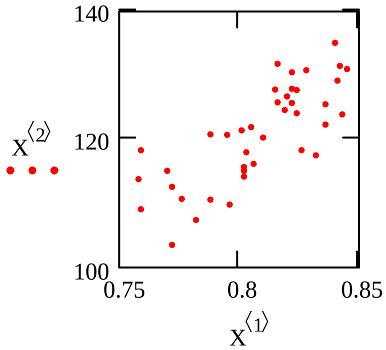
$SDstd$ =

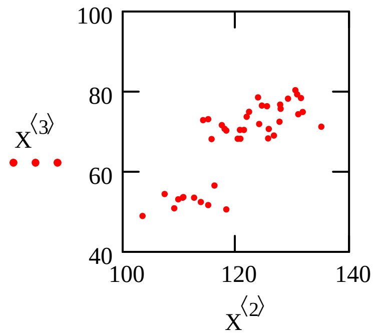
	1
1	0.1652
2	1.47575
3	4.15384
4	2.71649
5	5.28717
6	5.05283
7	0.8482
8	0.93055
9	3.78193
10	1.26841
11	0.7144
12	1.09527
13	0.01601
14	0.50661
15	1.30117
16	4.50683

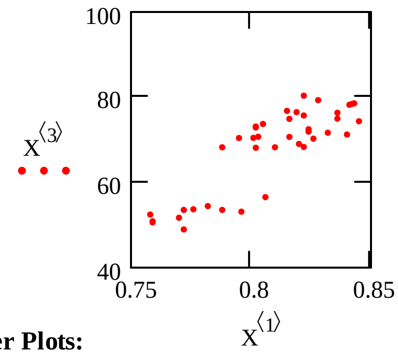
Note: If the data are standardized, then variances of each variable wil be the same BUT IF covariance between the variables is greater than zero, Statistical distance is no the same as Euclidean distance.

2006 jw031

Original Scatter Plots: **THE PHENOMENON OF COVARIANCE:**



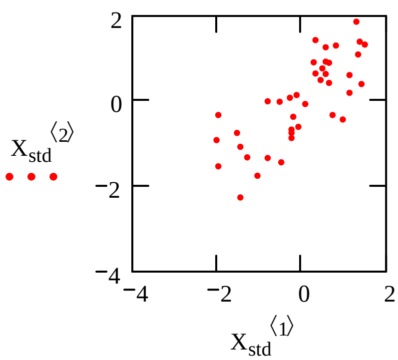


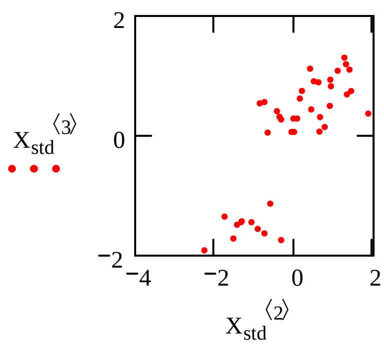


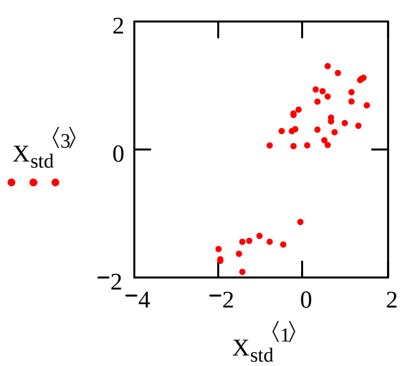
$X_{std} =$

	1	2	3
1	-0.2771	0.07643	0.28736
2	0.64993	0.89781	0.49728
3	1.33512	1.09368	1.08404
4	0.32748	1.43321	0.74509
5	1.29482	1.86414	0.36826
6	1.37543	1.39403	1.1035
7	0.4887	0.76722	0.144
8	-0.2368	-0.74757	0.5618
9	0.81115	1.30262	1.19463
10	0.4484	0.49299	0.90791
11	0.73054	-0.32839	0.26995
12	-0.2368	-0.86509	0.53927
13	0.08565	-0.06852	0.0631
14	-0.2368	-0.66922	0.05184
15	0.97237	-0.43286	0.41024
16	-0.47863	-1.43836	-1.48622

Standardized Scatter Plots:







Squarring each value in the standardized data matrix X_{std}:

$$Y_{j,k} := \left(X_{std_{j,k}}\right)^2$$

$Y =$

	1	2	3
1	0.07678	0.00584	0.08258
2	0.42241	0.80606	0.24729
3	1.78255	1.19614	1.17515
4	0.10724	2.05408	0.55517
5	1.67655	3.47501	0.13561
6	1.8918	1.94332	1.21771
7	0.23883	0.58863	0.02074
8	0.05607	0.55886	0.31562
9	0.65796	1.69682	1.42715
10	0.20106	0.24304	0.8243
11	0.53369	0.10784	0.07288
12	0.05607	0.74839	0.29081
13	0.00734	0.0047	0.00398
14	0.05607	0.44785	0.00269
15	0.94551	0.18737	0.1683
16	0.22909	2.06889	2.20885

Note: The effect of standardizing data is to center each axis on zero, and to give distribution for each variable equal size. However, because covariance still exists between the variables, it still matters what direction distance is measured. Euclidean distance along the direction of maximum scatter is statistically less surprising than the same distance in the direction of mimimum scatter.

IN A NUTSHELL, THIS IS WHAT MULTIVARIATE METHODS ARE ALL ABOUT!

2005 jw060

ORIGIN ≡ 1

LINEAR TRANSFORMATIONS
jw060.mcd

Prepared by:
Wm Stein

Reading in a grid of data points:

X := READPRN("\DATA\grid.txt")   n := 11   **< Predefined width of grid.**

X =

	1	2
1	-5	-5
2	-4	-5
3	-3	-5
4	-2	-5
5	-1	-5
6	0	-5
7	1	-5
8	2	-5
9	3	-5
10	4	-5
11	5	-5
12	-5	-4
13	-4	-4
14	-3	-4
15	-2	-4
16	-1	-4

Matrix M (2 X 2) specifying a linear transformation:

$$M := \begin{pmatrix} 1.0 & 0.3 \\ 0.3 & 1.0 \end{pmatrix}$$

<Change these values to see different patterns!

The transformation:

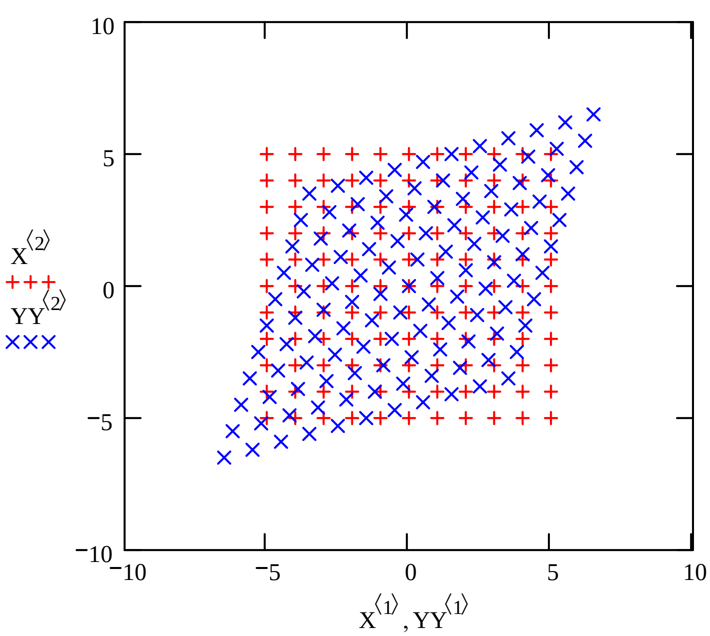
$$Y := M \cdot X^T$$

Matrix Y of transformed points:

Y =

	1	2	3	4	5	6	7	8	9	10
1	-6.5	-5.5	-4.5	-3.5	-2.5	-1.5	-0.5	0.5	1.5	2.5
2	-6.5	-6.2	-5.9	-5.6	-5.3	-5	-4.7	-4.4	-4.1	-3.8

Plotting points: $XX := X^T$ $YY := Y^T$



Note how the points have been moved by "warping" the underlying two-dimensional space.

Every Matrix M specifies a different warping of the space.

Displacement of points:

$$Z := X - Y^T$$ **< The displacement matrix.**

X =

	1	2
1	-5	-5
2	-4	-5
3	-3	-5
4	-2	-5
5	-1	-5
6	0	-5
7	1	-5
8	2	-5
9	3	-5
10	4	-5
11	5	-5
12	-5	-4
13	-4	-4
14	-3	-4
15	-2	-4
16	-1	-4

Y^T =

	1	2
1	-6.5	-6.5
2	-5.5	-6.2
3	-4.5	-5.9
4	-3.5	-5.6
5	-2.5	-5.3
6	-1.5	-5
7	-0.5	-4.7
8	0.5	-4.4
9	1.5	-4.1
10	2.5	-3.8
11	3.5	-3.5
12	-6.2	-5.5
13	-5.2	-5.2
14	-4.2	-4.9
15	-3.2	-4.6
16	-2.2	-4.3

Vector Plot:

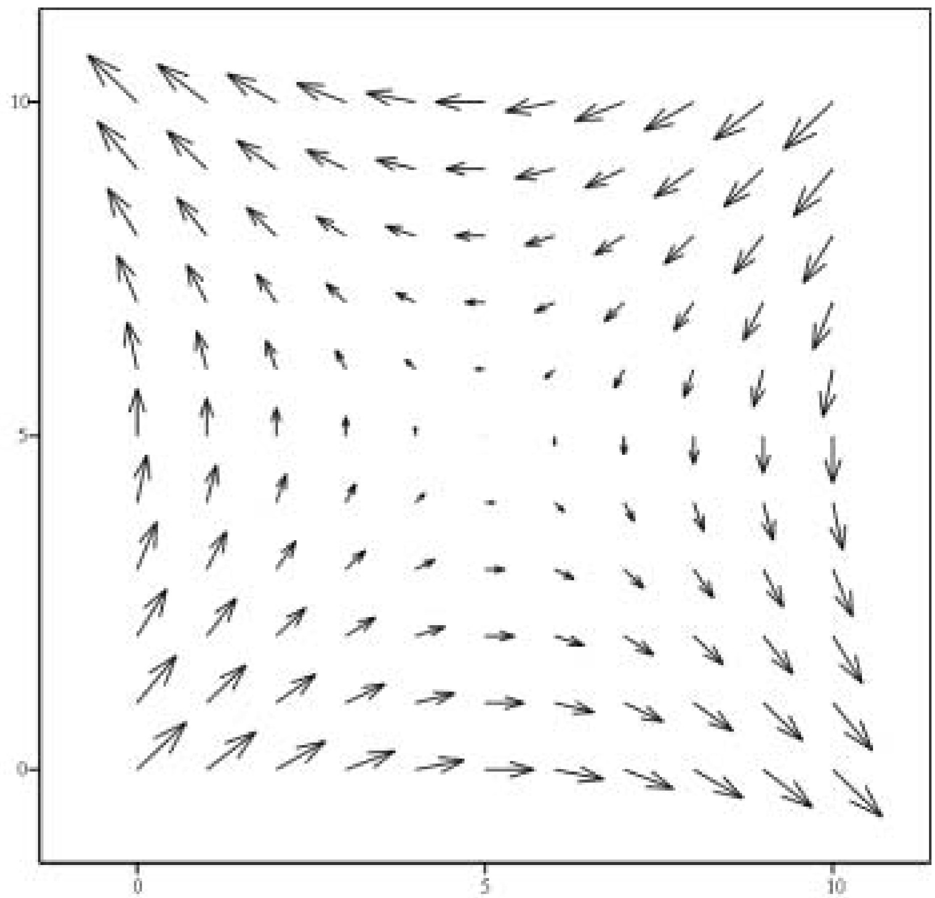
This graphing function for vector plots in Mathcad is not well implemented, so vectors are more-or-less correct (depending on values chosen for M). But you can get the idea here...

In the example here, vector heads represent points in X, and vector tails represent points in Y.

$ZZ := Z^T \qquad i := 1..2 \qquad k := 1..n$

$j := 1..n^2 \qquad kk := 1..n$

$Zx_{kk,k} := ZZ_{1,(kk-1)\cdot n+k} \quad Zy_{kk,k} := ZZ_{2,(kk-1)\cdot n+k}$



(Zy, Zx)

From this, it is easy to see that Linear Transformations in two dimension have two directions in the fabric of the warp that involve lengthening or shortening vector lengths only - but not changing their directions. In general there are n - many such directions in an n-dimensional space.

These are the Eigenvector directions!
The amount of streatch or shrink are recorded by the Eigenvalues.

ORIGIN ≡ 1

POINTS IN SPACE, THE LINEAR TRANSFORMATION, AND USE OF THE VARIANCE/COVARIANCE MATRIX TO CALCULATE DISTANCES
jw066.mcd

Prepared by:
Wm Stein

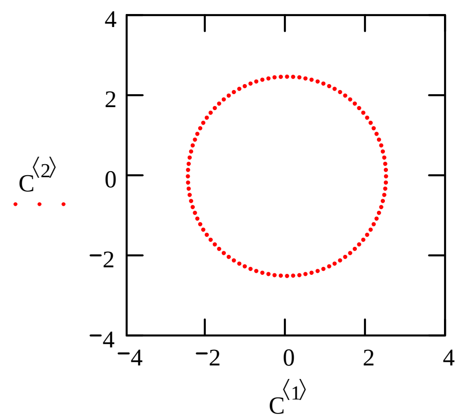
Points of Equal distance from the origin of a coordinate system:

$j := 1 .. 100$ $\theta_j := \frac{j}{100} \cdot 2 \cdot \pi$ $c := 2.488$

$Cx1_j := c \cdot \cos(\theta_j)$ $Cx2_j := c \cdot \sin(\theta_j)$

$C := \text{augment}(Cx1, Cx2)$ **< points on a circle of radius c**

All of the points on this circle are some unforgetable number = c units of Euclidean Distance from the center.



C =

	1	2
1	2.48309	0.15622
2	2.46838	0.31183
3	2.44393	0.4662
4	2.40983	0.61874
5	2.36623	0.76883
6	2.31328	0.91589
7	2.25121	1.05934
8	2.18025	1.1986
9	2.10069	1.33314
10	2.01283	1.46241
11	1.91704	1.58591
12	1.81367	1.70315
13	1.70315	1.81367
14	1.58591	1.91704
15	1.46241	2.01283
16	1.33314	2.10069

Calculating squared distances for all vectors in C:
using squared Euclidean distance in matrix form:

$n := \text{rows}(C)$

$i := 1 .. n$

$D_i := \left(C \cdot C^T\right)_{i,i}$

D =

	1
1	6.19014
2	6.19014
3	6.19014
4	6.19014
5	6.19014
6	6.19014
7	6.19014
8	6.19014
9	6.19014
10	6.19014
11	6.19014
12	6.19014
13	6.19014
14	6.19014
15	6.19014
16	6.19014

$c^2 = 6.19014$

The main diagonal of data coordinates matrix times its transpose.

Note: Like others in the field of Multivariate statistics, I will sometimes be sloppy and fail to make the distinction between DISTANCES per se and DISTANCES squared. Whereas the former sometimes has physical meaning, the latter is often used for calculations with matrix algebra.

2005 jw066

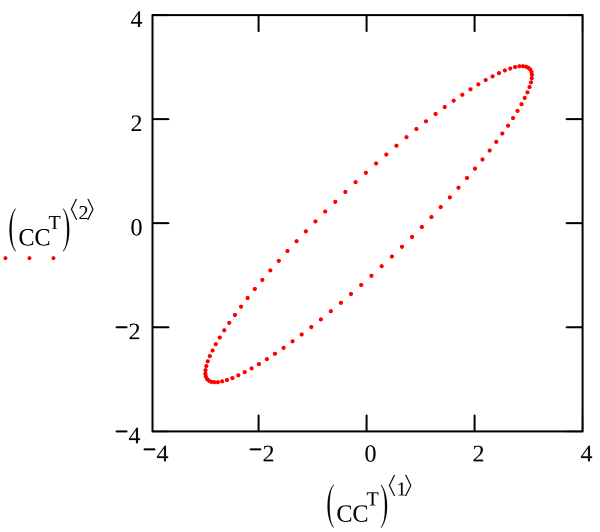
Now let's subject the points of the circle to a Linear Transformation:

$$M := \begin{pmatrix} 1 & 0.7 \\ 0.7 & 1 \end{pmatrix}$$

Note: I have deliberately chosen a symmetric matrix, but you can choose whatever you want.

$$CC := M \cdot C^T$$

The circle is now an ellipse!



$CC^T =$

	1	2
1	2.59245	1.89439
2	2.68666	2.0397
3	2.77027	2.17696
4	2.84295	2.30562
5	2.90441	2.42519
6	2.95441	2.53519
7	2.99275	2.63519
8	3.01927	2.72478
9	3.03388	2.80362
10	3.03652	2.87139
11	3.02717	2.92784
12	3.00588	2.97272
13	2.97272	3.00588
14	2.92784	3.02717
15	2.87139	3.03652
16	2.80362	3.03388

Now play with other values of M to see what happens!

In general, Linear transformations serve as an excellent means to model covariance in real data. Because all points in the plane are "warped" by the Linear transformation, what is now an ellipse can be viewed as retaining an equal STATISTICAL distance from the center.

Calculating squared statistical distances for all vectors in C:
using squared Statistical distance in matrix form:

Mean Vector and Variance-Covariance Matrix of Transformed Data Matrix CC:

$n := \text{rows}(CC^T)$ $\quad i := 1 .. n$ $\quad I_{vec_i} := 1$ $\quad I := \text{identity}(n)$

$$X_{bar} := \frac{1}{n} \cdot CC \cdot I_{vec} \qquad X_{bar} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$S_{CC} := \frac{1}{n-1} \cdot CC \cdot \left(I - \frac{1}{n} \cdot I_{vec} \cdot I_{vec}^T \right) \cdot CC^T \qquad S_{CC} = \begin{pmatrix} 4.65824 & 4.37687 \\ 4.37687 & 4.65824 \end{pmatrix}$$

$$ED_i := \left(CC^T \cdot CC \right)_{i,i}$$ **<- Eucledian distances (squared) for points on the ellipse**

$$SD_i := \left(CC^T \cdot S_{CC}^{-1} \cdot CC \right)_{i,i}$$ **<- Statistical distances (squared) for points on the ellipse**

$ED =$

	1
1	10.30948
2	11.37851
3	12.41356
4	13.39829
5	14.31718
6	15.15574
7	15.90074
8	16.54043
9	17.06473
10	17.46536
11	17.73601
12	17.87242
13	17.87242
14	17.73601
15	17.46536
16	17.06473

$SD =$

	1
1	1.98
2	1.98
3	1.98
4	1.98
5	1.98
6	1.98
7	1.98
8	1.98
9	1.98
10	1.98
11	1.98
12	1.98
13	1.98
14	1.98
15	1.98
16	1.98

Note: Although scaled differently, the important thing to note is that Statistical Distances are identical for all of the points on the ellipse whereas Euclidean Distances are not. (oops - forgot to say squared...)

Can we can undo the Linear Transformation specified by M??

Finding the inverse of matrix M:

$$M = \begin{pmatrix} 1 & 0.7 \\ 0.7 & 1 \end{pmatrix} \qquad M^{-1} = \begin{pmatrix} 1.96078 & -1.37255 \\ -1.37255 & 1.96078 \end{pmatrix} \qquad M \cdot M^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Transforming the data using the inverse matrix:

$$CB := M^{-1} \cdot CC$$

$CC^T =$

	1	2
1	2.59245	1.89439
2	2.68666	2.0397
3	2.77027	2.17696
4	2.84295	2.30562
5	2.90441	2.42519
6	2.95441	2.53519
7	2.99275	2.63519
8	3.01927	2.72478
9	3.03388	2.80362
10	3.03652	2.87139
11	3.02717	2.92784
12	3.00588	2.97272
13	2.97272	3.00588
14	2.92784	3.02717
15	2.87139	3.03652
16	2.80362	3.03388

$CB^T =$

	1	2
1	2.48309	0.15622
2	2.46838	0.31183
3	2.44393	0.4662
4	2.40983	0.61874
5	2.36623	0.76883
6	2.31328	0.91589
7	2.25121	1.05934
8	2.18025	1.1986
9	2.10069	1.33314
10	2.01283	1.46241
11	1.91704	1.58591
12	1.81367	1.70315
13	1.70315	1.81367
14	1.58591	1.91704
15	1.46241	2.01283
16	1.33314	2.10069

$C =$

	1	2
1	2.48309	0.15622
2	2.46838	0.31183
3	2.44393	0.4662
4	2.40983	0.61874
5	2.36623	0.76883
6	2.31328	0.91589
7	2.25121	1.05934
8	2.18025	1.1986
9	2.10069	1.33314
10	2.01283	1.46241
11	1.91704	1.58591
12	1.81367	1.70315
13	1.70315	1.81367
14	1.58591	1.91704
15	1.46241	2.01283
16	1.33314	2.10069

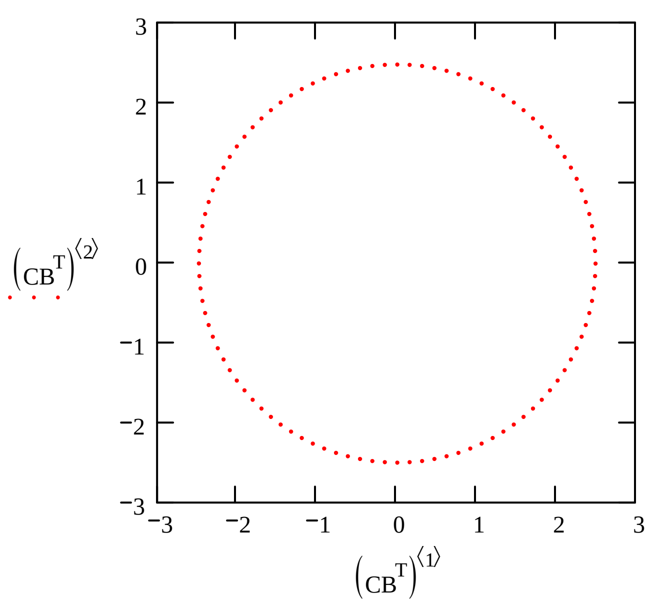
Good Grief, that's all there ist to it!

Linear transforming the data through the inverse of a matrix undoes the original transformation.

However, the problem is that in real situations involving elliptically scattered data, we lack prior knowledge of the original Linear Transformation represented by M.

Until some statistician tells me otherwise, it would seem that recovering the "original scale" of the data is not possible from the variance/covariance matrix alone.

Such a concept of "original scale" in prior real data probabally doesn't mean much anyway, and statistically this limitation is unimportant.



PLAYING WITH REAL-LOOKING DATA:

Creating A Multivariate Normal Distribution:

$n := 1000$ **<Number of points**

$\mu_1 := 0$ $\mu_2 := 0$ $\mu_3 := 0$

$\sigma_1 := 1$ $\sigma_2 := 1$ $\sigma_3 := 1$

$X_1 := \mathrm{rnorm}(n, \mu_1, \sigma_1)$ $X_2 := \mathrm{rnorm}(n, \mu_2, \sigma_2)$ $X_3 := \mathrm{rnorm}(n, \mu_2, \sigma_2)$

$X := \mathrm{augment}(X_1, X_2, X_3)$ **<Data array**

$X =$

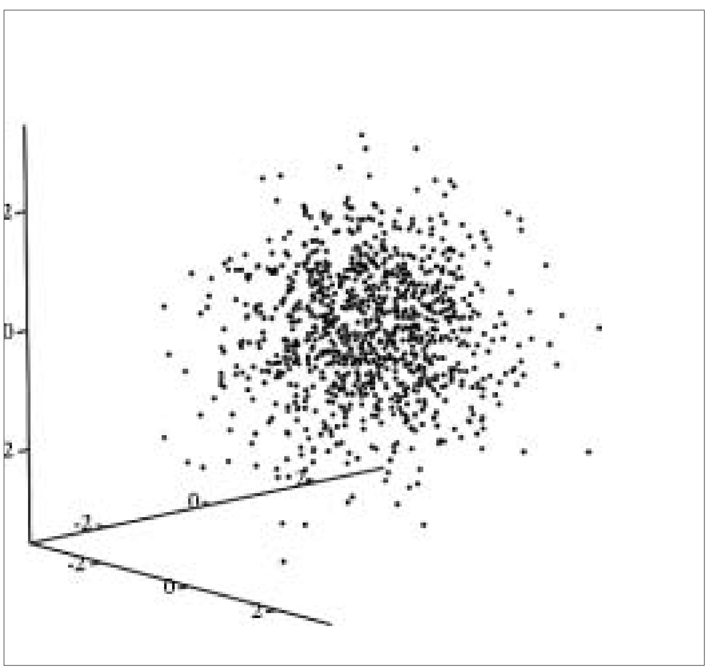
	1	2	3
1	-0.43897	0.63434	0.17454
2	-0.67941	0.29209	0.55938
3	-0.47329	0.14146	-0.97268
4	-0.95147	0.34771	0.7135
5	-1.68568	-0.38343	0.7757
6	0.04353	-0.06994	0.25824
7	-0.12063	-0.9638	0.55301
8	0.55643	-0.55921	-0.63348
9	2.19179	-2.05691	1.43786
10	0.80873	-0.16582	-0.2259
11	0.98514	0.10421	0.99399
12	0.86223	0.71155	0.4259
13	0.91557	-1.22359	-0.34793
14	0.673	1.53579	0.17871
15	-1.04431	-0.09462	0.68205
16	0.06908	-0.85186	0.1789

Mean Vector and Variance-Covariance Matrix of Data Matrix X:

$i := 1..n$ $\quad I_{vec_i} := 1$ $\quad I := \mathrm{identity}(n)$

$$X_{bar} := \frac{1}{n} \cdot X^T \cdot I_{vec} \qquad X_{bar} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$S_X := \frac{1}{n-1} \cdot X^T \cdot \left(I - \frac{1}{n} \cdot I_{vec} \cdot I_{vec}{}^T \right) \cdot X \qquad S_X = \begin{pmatrix} 0.9701 & -0.00983 & 0.02039 \\ -0.00983 & 0.95573 & 0.01914 \\ 0.02039 & 0.01914 & 1.0472 \end{pmatrix}$$



$\left(X^{\langle 1 \rangle}, X^{\langle 2 \rangle}, X^{\langle 3 \rangle}\right)$

What we have here is a nice spherical ball of points with fuzzy edges.

Calculating Euclidean Distances (squared) for the first 50 points:

$j := 1..50$

$$DD_j := \left(X \cdot X^T\right)_{j,j}$$

$DD =$

	1
1	0.62555
2	0.85981
3	1.19012
4	1.53527
5	3.59027
6	0.07347
7	1.24929
8	1.02363
9	11.10224
10	0.73258
11	1.96937
12	1.43113
13	2.45649
14	2.84353

Simulating Correlation between variables in Matrix X:

$$M_{sim} := \begin{pmatrix} 1 & .2 & .4 \\ .2 & 1 & .3 \\ .4 & .3 & 1 \end{pmatrix} \qquad Y := \left(M_{sim} \cdot X^T\right)^T$$

< New Data Matrix Y

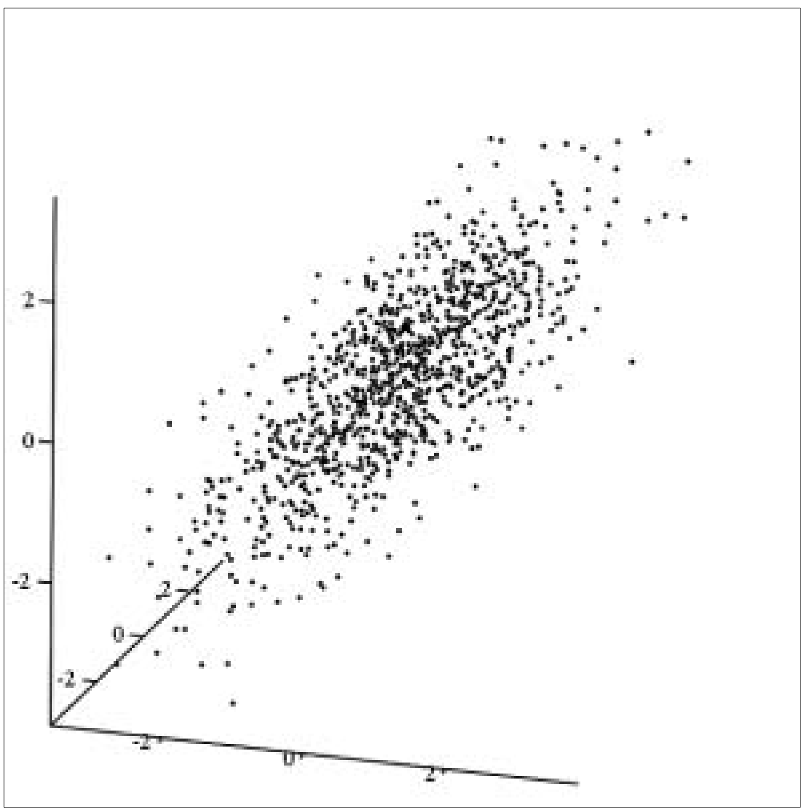
Y =

	1	2	3
1	-0.24229	0.59891	0.18926
2	-0.39724	0.32402	0.37524
3	-0.83406	-0.245	-1.11955
4	-0.59652	0.37146	0.43723
5	-1.45209	-0.48786	-0.0136
6	0.13284	0.01623	0.25467
7	-0.09219	-0.82203	0.21562
8	0.19119	-0.63797	-0.57868
9	2.35555	-1.18719	1.6975
10	0.68521	-0.07185	0.04784
11	1.40358	0.59943	1.41931
12	1.1749	1.01176	0.98425
13	0.53168	-1.14486	-0.34878
14	1.05164	1.72401	0.90865
15	-0.79042	-0.09887	0.23594
16	-0.02973	-0.78437	-0.04902

Mean Vector and Variance-Covariance Matrix of Data Matrix Y:

$$Y_{bar} := \frac{1}{n} \cdot Y^T \cdot I_{vec} \qquad Y_{bar} = \begin{pmatrix} -0.03461 \\ 0.04403 \\ 0.03023 \end{pmatrix}$$

$$S_Y := \frac{1}{n-1} \cdot Y^T \cdot \left(I - \frac{1}{n} \cdot I_{vec} \cdot I_{vec}^T\right) \cdot Y \qquad S_Y = \begin{pmatrix} 1.19132 & 0.51716 & 0.89031 \\ 0.51716 & 1.09878 & 0.70135 \\ 0.89031 & 0.70135 & 1.31387 \end{pmatrix}$$



$$\left(Y^{\langle 1 \rangle}, Y^{\langle 2 \rangle}, Y^{\langle 3 \rangle}\right)$$

Calculating Statistical Distances (squared) for the first 50 points:

$$j := 1 .. 50$$

$$SDD_j := \left(Y \cdot S_Y^{-1} \cdot Y^T\right)_{j,j}$$

SDD =

	1
1	0.64177
2	0.86914
3	1.14132
4	1.55742
5	3.73841
6	0.07098
7	1.30558
8	1.02393
9	11.2466
10	0.75513
11	1.91505
12	1.45701
13	2.51928
14	2.97404

What we have here is an ellipsoidal "squashed football" with fuzzy edges.

Using M_{sim}^{-1} to undo Correlation between variables in Matrix Y:

$$Z := M_{sim}^{-1} \cdot Y^{T}$$

< Z = Recovering Matrix X

$Z^T =$

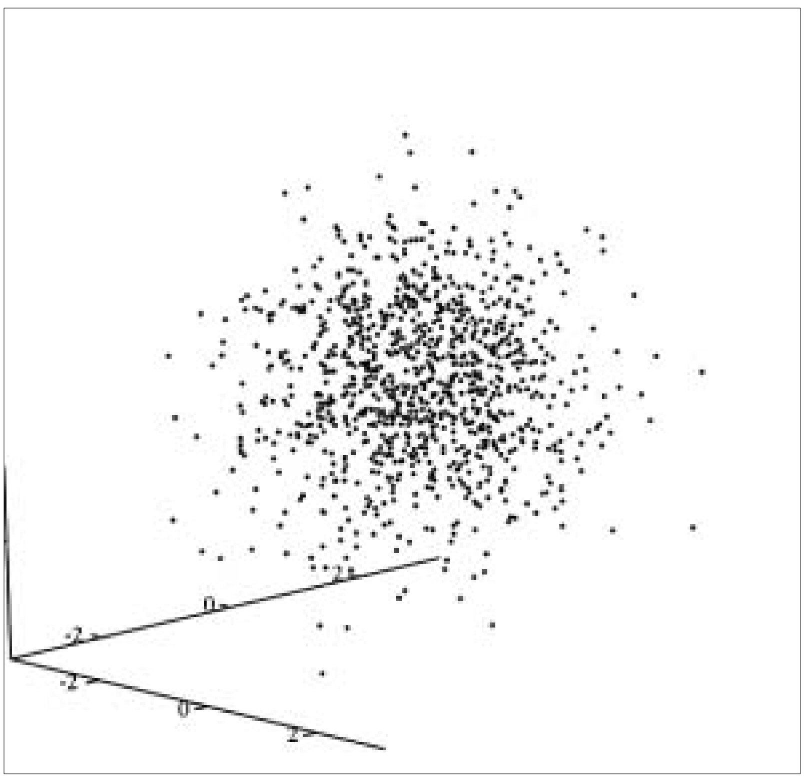
	1	2	3
1	-0.43897	0.63434	0.17454
2	-0.67941	0.29209	0.55938
3	-0.47329	0.14146	-0.97268
4	-0.95147	0.34771	0.7135
5	-1.68568	-0.38343	0.7757
6	0.04353	-0.06994	0.25824
7	-0.12063	-0.9638	0.55301
8	0.55643	-0.55921	-0.63348
9	2.19179	-2.05691	1.43786
10	0.80873	-0.16582	-0.2259
11	0.98514	0.10421	0.99399
12	0.86223	0.71155	0.4259
13	0.91557	-1.22359	-0.34793
14	0.673	1.53579	0.17871
15	-1.04431	-0.09462	0.68205
16	0.06908	-0.85186	0.1789

Mean Vector and Variance-Covariance Matrix of Recovered Data Matrix Z:

$$Z_{bar} := \frac{1}{n} \cdot Z \cdot I_{vec} \qquad Z_{bar} = \begin{pmatrix} -0.05975 \\ 0.04368 \\ 0.04103 \end{pmatrix}$$

$$S_Z := \frac{1}{n} \cdot Z \cdot \left(I - \frac{1}{n} \cdot I_{vec} \cdot I_{vec}^{T} \right) \cdot Z^T$$

$$S_Z = \begin{pmatrix} 0.96913 & -0.00982 & 0.02037 \\ -0.00982 & 0.95477 & 0.01912 \\ 0.02037 & 0.01912 & 1.04615 \end{pmatrix}$$



$$\left[\left(Z^T\right)^{\langle 1 \rangle}, \left(Z^T\right)^{\langle 2 \rangle}, \left(Z^T\right)^{\langle 3 \rangle} \right]$$

Compare Z with X.

Lucky we had M_{sim}!

$X =$

	1	2	3
1	-0.43897	0.63434	0.17454
2	-0.67941	0.29209	0.55938
3	-0.47329	0.14146	-0.97268
4	-0.95147	0.34771	0.7135
5	-1.68568	-0.38343	0.7757
6	0.04353	-0.06994	0.25824
7	-0.12063	-0.9638	0.55301
8	0.55643	-0.55921	-0.63348
9	2.19179	-2.05691	1.43786
10	0.80873	-0.16582	-0.2259
11	0.98514	0.10421	0.99399
12	0.86223	0.71155	0.4259
13	0.91557	-1.22359	-0.34793
14	0.673	1.53579	0.17871
15	-1.04431	-0.09462	0.68205
16	0.06908	-0.85186	0.1789

2005 jw084

ORIGIN ≡ 1

MATRIX ALGEGRA TOOLBOX FOR MULTIVARIATE METHODS

jw084.mcd

Prepared by: Wm Stein

VECTORS:

We'll use vectors x_n and scalar c here:

$$x_1 := \begin{pmatrix} 3 \\ -1 \\ 4 \end{pmatrix} \qquad x_2 := \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix} \qquad c := 5$$

Vector Addition & Scalar Multiplication:

$$x_1 + x_2 = \begin{pmatrix} 4 \\ 1 \\ 2 \end{pmatrix} \qquad c \cdot x_2 = \begin{pmatrix} 5 \\ 10 \\ -10 \end{pmatrix}$$

Vector Multiplication:

$$x_1 \cdot x_2 = -7 \qquad {x_1}^T \cdot x_2 = (-7)$$

$$x_1 \cdot {x_2}^T = \begin{pmatrix} 3 & 6 & -6 \\ -1 & -2 & 2 \\ 4 & 8 & -8 \end{pmatrix}$$

Several vector products have been defined. We will mostly use the "inner product" (= dot product) which is the same as matrix product.

In general, the dot product is not commutative. Note how Mathcad treats vector multiplication here. (-7) is a (1X1) vector whereas -7 is a scalar number, and $x_1^T x_2$ does not equal $x_1 x_2^T$.

In general, it is best to specify the product you want by using the Transpose function on the Matrix pad.

Vector Length & Angle between two vectors:

$$L_{x1} := \left|\sqrt{{x_1}^T \cdot x_1}\right| \quad L_{x1} = 5.09902 \quad L_{x2} := \left|\sqrt{{x_2}^T \cdot x_2}\right| \quad L_{x2} = 3$$

Note use of absolute value (Determinant) function on Matrix pad to convert datatype for L_x & L_y from matrix to scalar numbers.

$$|x_1| = 5.09902 \qquad |x_2| = 3$$

Length can also be found by using the Determinant function alone on the Matrix pad, but it is important not to forget the original definition.

$$\cos_\theta := \left(\frac{{x_1}^T \cdot x_2}{L_{x1} \cdot L_{x2}}\right) \qquad \cos_\theta = (-0.457604)$$

θ is the angle between vectors x & y
$\cos_\theta$ is the cosine of θ

$$\theta := \mathrm{acos}(\cos_\theta) \qquad \theta = (2.046095)\ \mathrm{rad} \quad \theta = (117.232626)\ \mathrm{deg}$$

Note: units for θ are obtained in Mathcad by typing "rad" or "deg" in the trailing box following an evaluation.

When vectors are: PARALLEL, θ = 0 or 180 deg, & $\cos_\theta$ = 1.

PERPENDICULAR, θ = 90 or 270 deg, & $\cos_\theta$ = 0; also $x^T y = 0$

2005 jw084

Projection of one vector onto another:

Projection of x_1 onto x_2:

$$proj_{x1onx2} := \frac{\left|x_1^T \cdot x_2\right|}{\left|x_2^T \cdot x_2\right|} \cdot x_2$$

$$proj_{x1onx2} = \begin{pmatrix} 0.777778 \\ 1.555556 \\ -1.555556 \end{pmatrix}$$

Projection of x_1 onto x_2:

$$proj_{x2onx1} := \frac{\left|x_2^T \cdot x_1\right|}{\left|x_1^T \cdot x_1\right|} \cdot x_1$$

$$proj_{x2onx1} = \begin{pmatrix} 0.807692 \\ -0.269231 \\ 1.076923 \end{pmatrix}$$

Note: this is a way to convert from one coordinate system to another.

MATRICES:

We'll use Matrix A & B and scalar c here:

$$A := \begin{pmatrix} 0 & 3 & 1 \\ 1 & -1 & 1 \end{pmatrix} \qquad B := \begin{pmatrix} 1 & -2 & -3 \\ 2 & 5 & 1 \end{pmatrix} \qquad c := 7$$

Summary numbers about Matrices:

$rows(A) = 2$

$cols(B) = 3$

$rank(A) = 2$

Matrix addition and scalar multiplication:

$$A + B = \begin{pmatrix} 1 & 1 & -2 \\ 3 & 4 & 2 \end{pmatrix} \qquad c \cdot A = \begin{pmatrix} 0 & 21 & 7 \\ 7 & -7 & 7 \end{pmatrix}$$

Note: Matrix rank gives the number of linearly independent columns (= rows) of a matrix.

Matrix Multiplication:

$A \cdot B = ▮$ $\qquad A^T \cdot B^T = ▮$ $\qquad A \cdot B^T = \begin{pmatrix} -9 & 16 \\ 0 & -2 \end{pmatrix}$

$$A^T \cdot B = \begin{pmatrix} 2 & 5 & 1 \\ 1 & -11 & -10 \\ 3 & 3 & -2 \end{pmatrix} \qquad B^T \cdot A = \begin{pmatrix} 2 & 1 & 3 \\ 5 & -11 & 3 \\ 1 & -10 & -2 \end{pmatrix}$$

Note: Multiplication is not commutative and here in Mathcad, you MUST use the Transpose function to multiply properly!

$x_1^T \cdot A = ▮$

$$A \cdot x_1 = \begin{pmatrix} 1 \\ 8 \end{pmatrix} \qquad x_2^T \cdot B^T = \begin{pmatrix} 3 & 10 \end{pmatrix}$$

Special kinds of Matrices:

$$S := \begin{pmatrix} 4 & 7 & 6 \\ 5 & 1 & 1 \\ 3 & 2 & 1 \end{pmatrix} \qquad D := \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -9 \end{pmatrix}$$

S = SQUARE MATRIX: a matrix, where rows(S) = cols(S)

M = SYMMETRIC MATRIX: a kind of Square matrix, where M = M^T

D = DIAGONAL MATRIX: a kind of Symmetric matrix, where D_{ij} = 0 for indices i & j not equal.

I = IDENTITY MATRIX: a kind of Diagonal matrix where I_{ii} = 1.

$$M := \begin{pmatrix} 1 & -5 & 7 \\ -5 & 2 & 4 \\ 7 & 4 & 3 \end{pmatrix} \qquad I := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$M^T = \begin{pmatrix} 1 & -5 & 7 \\ -5 & 2 & 4 \\ 7 & 4 & 3 \end{pmatrix} \qquad D^T = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -9 \end{pmatrix} \qquad I^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Matrix Inverse:

$B^{-1} = \blacksquare$ $\quad (B^T)^{-1} = \blacksquare$ $\quad \begin{pmatrix} rows(B) \\ cols(B) \\ rank(B) \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 2 \end{pmatrix}$

A matrix inverse only exists for Square matrices in which the number of rows or columns is the same as matrix rank. Such matrices are called NONSINGULAR. A matrix that fails this condition is termed SINGULAR

$$S^{-1} = \begin{pmatrix} -0.041667 & 0.208333 & 0.041667 \\ -0.083333 & -0.583333 & 1.083333 \\ 0.291667 & 0.541667 & -1.291667 \end{pmatrix} \quad \begin{pmatrix} rows(S) \\ cols(S) \\ rank(S) \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix}$$

The inverse of a square matrix is the matrix such that when multiplied with the original matrix yields the Identity matrix.

$$D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -9 \end{pmatrix} \quad D^{-1} = \begin{pmatrix} 0.5 & 0 & 0 \\ 0 & 0.25 & 0 \\ 0 & 0 & -0.111111 \end{pmatrix} \quad \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{-9} \end{pmatrix} = \begin{pmatrix} 0.5 & 0 & 0 \\ 0 & 0.25 & 0 \\ 0 & 0 & -0.111111 \end{pmatrix} \quad D \cdot D^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Whereas some matrix inverses (such as for diagonal matrices shown above) are easy to determine most require extensive calculations that are best left to a computer algorithm.

Note however, that due to rounding errors in calculating matrix inverses, it is a wise precaution to calculate the Identity products as a check for bad behavior.

$$S^{-1} \cdot S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad S \cdot S^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Orthogonal Matrix:

$$Q := \begin{pmatrix} \frac{-1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{-1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{-1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{-1}{2} \end{pmatrix} \quad \begin{matrix} i := 1 .. rows(Q) \\ j := 1 .. rows(Q) \end{matrix}$$

$$\left| Q^{\langle j \rangle} \right| = \blacksquare \qquad \left| Q^{\langle i \rangle T} \cdot Q^{\langle j \rangle} \right| = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

An Orthogonal matrix has row (and column) vectors of unit length and perpendicular to each other also meaning that their dot product is zero.

$$Q^{-1} = \begin{pmatrix} -0.5 & 0.5 & 0.5 & 0.5 \\ 0.5 & -0.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & -0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 & -0.5 \end{pmatrix} \quad Q^T = \begin{pmatrix} -0.5 & 0.5 & 0.5 & 0.5 \\ 0.5 & -0.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & -0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 & -0.5 \end{pmatrix}$$

Also, the inverse of an Orthogonal matrix is the same as its Transpose.

$$I^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad I^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Here's another somewhat dippy example of an Orthogonal matrix! But, see also eigenvector Matrix E below...

2005 jw084

Eigenvalues and Eigenvectors:

$$M = \begin{pmatrix} 1 & -5 & 7 \\ -5 & 2 & 4 \\ 7 & 4 & 3 \end{pmatrix} \quad \text{eigenvals}(M) = \begin{pmatrix} 5.759946 \\ 9.08647 \\ -8.846416 \end{pmatrix} \quad \text{eigenvecs}(M) = \begin{pmatrix} 0.340179 & -0.674399 & -0.655335 \\ -0.859124 & 0.06046 & -0.508184 \\ -0.38234 & -0.735887 & 0.558825 \end{pmatrix}$$

Eigenvalues and Eigenvectors may be calculated for any Square matrix, although you may not like the numbers you see - such as negative or complex eigenvalues.

X := READPRN("\DATA\T1-2-25.DAT")

$i := 1 .. \text{rows}(X) \quad n := \text{rows}(X) \quad I_{vec_i} := 1 \quad I := \text{identity}(\text{rows}(X))$

$$S := \frac{1}{n-1} \cdot X^T \cdot \left(I - \frac{1}{n} \cdot I_{vec} \cdot I_{vec}^T \right) \cdot X \qquad S = \begin{pmatrix} 0.000631 & 0.151239 & 0.212717 \\ 0.151239 & 60.146061 & 61.96442 \\ 0.212717 & 61.96442 & 97.811055 \end{pmatrix}$$

$\lambda := \text{eigenvals}(S)$

$E := \text{eigenvecs}(S)$

$$\lambda = \begin{pmatrix} 0.000156 \\ 143.742079 \\ 14.215513 \end{pmatrix} \quad E = \begin{pmatrix} 0.999998 & 0.001815 & -0.000364 \\ -0.000789 & 0.595486 & 0.803365 \\ -0.001675 & 0.803364 & -0.595486 \end{pmatrix}$$

Eigenvectors are normalized to have unit length.

IF the matrix from which they are drawn is Symmetric, THEN Eigenvectors are perpendicular to each other, and the matrix E of all Eigenvectors as columns (rows) is an Orthogonal Matrix.

$i := 1 .. \text{rows}(E) \quad j := 1 .. \text{rows}(E)$

$$\left| E^{\langle i \rangle} \right| = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \qquad \left| E^{\langle i \rangle T} \cdot E^{\langle j \rangle} \right| = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

$$\lambda_1 = 0.000156 \quad \text{eigenvec}(S, \lambda_1) = \begin{pmatrix} 0.999998 \\ -0.000789 \\ -0.001675 \end{pmatrix}$$

$$\lambda_2 = 143.74207 \quad \text{eigenvec}(S, \lambda_2) = \begin{pmatrix} 0.001815 \\ 0.595486 \\ 0.803364 \end{pmatrix}$$

If Eigenvalues are differ from each other, Eigenvalues & Eigenvectors must always be associated in pairs.

It is important not to mix them up!

$$\lambda_3 = 14.215513 \quad \text{eigenvec}(S, \lambda_3) = \begin{pmatrix} 0.000364 \\ -0.803365 \\ 0.595486 \end{pmatrix}$$

The Eigenvec function in Mathcad allows you to form the appropriate pairs.

Here's a safe way to ALWAYS associate Eigenvalues λ with Eigenvectors as columns of E:

$i := 1 .. \text{rows}(S)$

$\lambda := \text{reverse}(\text{sort}(\text{eigenvals}(S))) \qquad E^{\langle i \rangle} := \text{eigenvec}(S, \lambda_i)$

$$\lambda = \begin{pmatrix} 143.742079 \\ 14.215513 \\ 0.000156 \end{pmatrix} \quad E = \begin{pmatrix} 0.001815 & 0.000364 & 0.999998 \\ 0.595486 & -0.803365 & -0.000789 \\ 0.803364 & 0.595486 & -0.001675 \end{pmatrix}$$

Spectral Decomposition of a Matrix:

This technique is often used when treating variance-covariance matrix S of a dataset.

$i := 1 .. \text{rows}(S)$

$P_i := \left(\lambda_i \cdot E^{\langle i \rangle} \cdot E^{\langle i \rangle T}\right)$

$$P = \begin{bmatrix} \begin{pmatrix} 0.0005 & 0.1554 & 0.2096 \\ 0.1554 & 50.9714 & 68.765 \\ 0.2096 & 68.765 & 92.7702 \end{pmatrix} \\ \begin{pmatrix} 1.88 \times 10^{-6} & -0.0042 & 0.0031 \\ -0.0042 & 9.1746 & -6.8006 \\ 0.0031 & -6.8006 & 5.0409 \end{pmatrix} \\ \begin{pmatrix} 0.0002 & -1.2286 \times 10^{-7} & -2.6085 \times 10^{-7} \\ -1.2286 \times 10^{-7} & 9.6921 \times 10^{-11} & 2.0578 \times 10^{-10} \\ -2.6085 \times 10^{-7} & 2.0578 \times 10^{-10} & 4.3692 \times 10^{-10} \end{pmatrix} \end{bmatrix}$$

The original matrix S can be restated as a sum of Partial Matrices P derived from the Eigenvalues and Eigenvectors of S.

Summing the Partials, reconstitutes S.

$$\sum_i P_i = \begin{pmatrix} 0.000631 & 0.151239 & 0.212717 \\ 0.151239 & 60.146061 & 61.96442 \\ 0.212717 & 61.96442 & 97.811055 \end{pmatrix} \qquad S = \begin{pmatrix} 0.000631 & 0.151239 & 0.212717 \\ 0.151239 & 60.146061 & 61.96442 \\ 0.212717 & 61.96442 & 97.811055 \end{pmatrix}$$

Singular Value Decomposition of a Matrix:

Using the Matrix A defined in JW Example p. 102-103:

$$A := \begin{pmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{pmatrix}$$

$m := \text{rows}(A) \qquad k := \text{cols}(A)$

$\gamma_U := \text{reverse}\left(\text{sort}\left(\text{eigenvals}\left(A \cdot A^T\right)\right)\right) \qquad \gamma_U = \begin{pmatrix} 12 \\ 10 \end{pmatrix}$

$\gamma_V := \text{reverse}\left(\text{sort}\left(\text{eigenvals}\left(A^T \cdot A\right)\right)\right) \qquad \gamma_V = \begin{pmatrix} 12 \\ 10 \\ 0 \end{pmatrix}$

Note: as in the text, Eigenvalues for the different products of A are designated γ (gamma) γ_U & γ_V

$i := 1 .. \text{rows}\left(A \cdot A^T\right) \qquad j := 1 .. \text{rows}\left(A^T \cdot A\right)$

Associated Eigenvectors are extracted:

$U^{\langle i \rangle} := \text{eigenvec}\left(A \cdot A^T, \gamma_{U_i}\right) \qquad U = \begin{pmatrix} 0.707107 & -0.707107 \\ 0.707107 & 0.707107 \end{pmatrix}$

$V^{\langle j \rangle} := \text{eigenvec}\left(A^T \cdot A, \gamma_{V_j}\right) \qquad V = \begin{pmatrix} 0.408248 & -0.894427 & -0.182574 \\ 0.816497 & 0.447214 & -0.365148 \\ 0.408248 & 0 & 0.912871 \end{pmatrix}$

$\lambda := \sqrt{\gamma_U} \qquad \lambda = \begin{pmatrix} 3.464102 \\ 3.162278 \end{pmatrix}$ **Singular values λ (lambda) are calculated.**

$i := 1 .. k - m \qquad k - m = 1 \qquad Z_{m,i} := 0 \qquad Z = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

(k-m) Extra columns of zeros are needed for singular value matrix Λ

$\Lambda := \text{augment}(\text{diag}(\lambda), Z)$

$$\Lambda = \begin{pmatrix} 3.464102 & 0 & 0 \\ 0 & 3.162278 & 0 \end{pmatrix}$$

Λ is the matrix of singular values and zeros

$$U \cdot \Lambda \cdot V^T = \begin{pmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{pmatrix} \qquad A = \begin{pmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{pmatrix}$$

Singular Value Decomposition confirmed!

2005 jw084

Singular Value Decomposition may be used for a Data matrix such as X:

$Y := X^T$ **< Transposing data matrix X from above.**

$m := \text{rows}(Y) \quad k := \text{cols}(Y)$

$\gamma_U := \text{reverse}\left(\text{sort}\left(\text{eigenvals}\left(Y \cdot Y^T\right)\right)\right)$

$\gamma_V := \text{reverse}\left(\text{sort}\left(\text{eigenvals}\left(Y^T \cdot Y\right)\right)\right)$

$i := 1 .. \text{rows}\left(Y \cdot Y^T\right) \quad j := 1 .. \text{rows}\left(Y^T \cdot Y\right)$

$U^{\langle i \rangle} := \text{eigenvec}\left(Y \cdot Y^T, \gamma_{U_i}\right)$

$V^{\langle j \rangle} := \text{eigenvec}\left(Y^T \cdot Y, \gamma_{V_j}\right)$

$\lambda := \sqrt{\gamma_U}$

$i := 1 .. k - m \quad k - m = 37 \qquad Z_{m,i} := 0$

$\Lambda := \text{augment}(\text{diag}(\lambda), Z)$

X =

	1	2	3
1	0.801	121.41	70.42
2	0.824	127.7	72.47
3	0.841	129.2	78.2
4	0.816	131.8	74.89
5	0.84	135.1	71.21
6	0.842	131.5	78.39
7	0.82	126.7	69.02
8	0.802	115.1	73.1
9	0.828	130.8	79.28
10	0.819	124.6	76.48
11	0.826	118.31	70.25
12	0.802	114.2	72.88
13	0.81	120.3	68.23
14	0.802	115.7	68.12
15	0.832	117.51	71.62
16	0.796	109.81	53.1
17	0.759	109.1	50.85

Note rounding errors here...

$\left(U \cdot \Lambda \cdot V^T\right)^T =$

	1	2	3
1	0.812739	121.409913	70.420015
2	0.87745	127.699605	72.470069
3	0.867896	129.199801	78.200035
4	0.939834	131.799085	74.890159
5	0.974075	135.099009	71.210172
6	0.900406	131.499568	78.390075
7	0.875535	126.69959	69.020071
8	0.711577	115.100668	73.099884
9	0.901769	130.799455	79.280095
10	0.826318	124.599946	76.480009
11	0.742355	118.310618	70.249892
12	0.698839	114.200762	72.879867
13	0.792961	120.300126	68.229978
14	0.733247	115.700508	68.119912
15	0.721009	117.51082	71.619857
16	0.690789	109.810778	53.099865

Quadratic Forms:

$I := \text{identity}(3)$

$$x_1 = \begin{pmatrix} 3 \\ -1 \\ 4 \end{pmatrix} \qquad M = \begin{pmatrix} 1 & -5 & 7 \\ -5 & 2 & 4 \\ 7 & 4 & 3 \end{pmatrix} \qquad I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$x_1^T \cdot M \cdot x_1 = (225)$$

Quadratic forms $x^T M x$ for some vector x and matrix M are numbers that result from adding squared or cross-product terms of x.

If $x^T M x > 0$ then the quadratic form & Matrix M are called POSITIVE DEFINITE.

A Matrix M is Positive Definite if and only if M's Eigenvalues are all positive.

$$D_{sq} := x_1^T \cdot M \cdot x_1$$

If $0 < D_{sq}$ $x^T M x$, then the quadratic form $x^T M x$ may be viewed as squared distance of the data point in vector x from the center of the coordinate system after performing a Linear Transformation.

See jw060 & jw066 for more details

If Matrix M = I the Identity Matrix, then D_{sq} involves EUCLIDEAN DISTANCES (squared). Otherwise STATISTICAL DISTANCES (sqaured) involving the inverse of M or variance-covariance matrix S are used.

$$\mu := \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}$$

$$D_{sq} := \left(x_1 - \mu\right)^T \cdot M^{-1} \cdot \left(x_1 - \mu\right) \qquad D_{sq} = (2.613391)$$

squared Distances are given by the formula above correcting for the effect of Linear Transformation specifed by M for a fixed reference point μ.

2005 jw098

ORIGIN ≡ 1

LINEAR TRANSFORMATIONS AND THE CALCULATION OF EIGENVECTORS AND EIGENVALUES
jw098.mcd

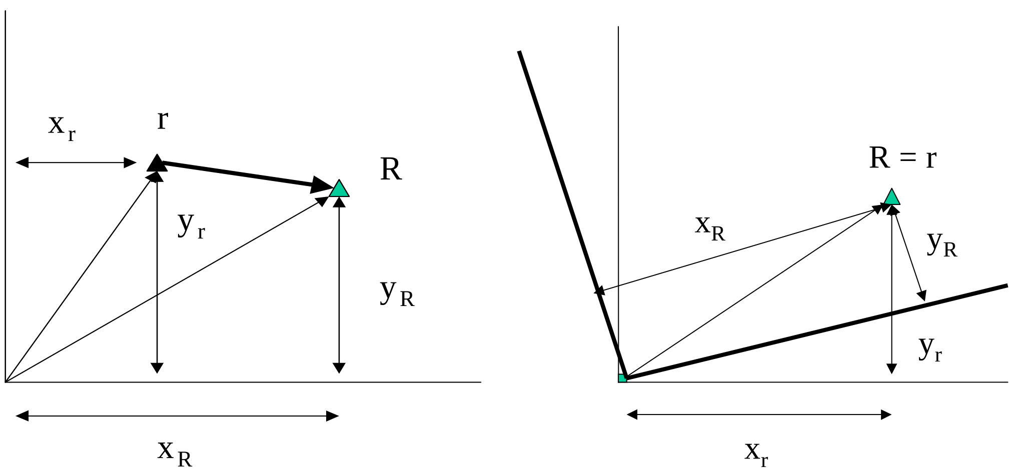
Prepared by:
Wm Stein

Consider a linear transformation from r --> R:

$x_R = a \cdot x_r + b \cdot y_r$ **or** $\begin{pmatrix} x_R \\ y_R \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} x_r \\ y_r \end{pmatrix}$ **<** $R = M \cdot r$

$y_R = c \cdot x_r + d \cdot y_r$

Geometric Interpretations:



Using the left interpretation, suppose the (x,y) plane is covered by an elastic membrane which may be shrunk, streached, or rotated so that:

- **r (x_r,y_r) becomes R (x_R,y_R) for any point on the plane**
- **Let M describe the transformation such that R = Mr**

Now let us ask if there are any vectors r that remain unchanged in direction during transformation? Note, however, that we will allow them to be shrunk or streached. i.e., R = λr where λ is some scalar constant.

$$R = \lambda \cdot r$$

$$\begin{pmatrix} x_R \\ y_R \end{pmatrix} = \lambda \cdot \begin{pmatrix} x_r \\ y_r \end{pmatrix} = \begin{pmatrix} \lambda \cdot x_r \\ \lambda \cdot y_r \end{pmatrix}$$

Such vectors are called <u>eigenvectors</u>, and the λ's are termed <u>eigenvalues</u> (characteristic values) of the transformation matrix M.

To calculate eigenvalues and eigenvectors, remember that:

$$R = M \cdot r = \lambda \cdot r$$

$$R = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} x_r \\ y_r \end{pmatrix} = \lambda \cdot \begin{pmatrix} x_r \\ y_r \end{pmatrix}$$

$$R = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} x_r \\ y_r \end{pmatrix} = \begin{pmatrix} \lambda \cdot x_r \\ \lambda \cdot y_r \end{pmatrix}$$

Converting matrix algebra to simultaneous equations:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} x_r \\ y_r \end{pmatrix} = \begin{pmatrix} \lambda \cdot x_r \\ \lambda \cdot y_r \end{pmatrix}$$

$a \cdot x_r + b \cdot y_r = \lambda \cdot x_r$ **or** $(a - \lambda) \cdot x_r + b \cdot y_r = 0$

$c \cdot x_r + d \cdot y_r = \lambda \cdot y_r$ **or** $c \cdot x_r + (d - \lambda) \cdot y_r = 0$

Eigenvector equation > $$\begin{pmatrix} a - \lambda & b \\ c & d - \lambda \end{pmatrix} \cdot \begin{pmatrix} x_r \\ x_r \end{pmatrix} = 0$$

For x_r & y_r both not equal to zero, the eigenvector equation is solved using the determinant:

$$\left| \begin{pmatrix} a - \lambda & b \\ c & d - \lambda \end{pmatrix} \right| = 0$$ **< Characteristic equation of M**

$$(a - \lambda) \cdot (d - \lambda) - bc = 0$$

$$a \cdot d - a \cdot \lambda - d \cdot \lambda + \lambda^2 - b \cdot c = 0$$

$$(1) \cdot \lambda^2 - (a + d) \cdot \lambda + (a \cdot d - b \cdot c) = 0$$ **< Quadratic equation of the general form:**

$$A \cdot \lambda^2 + B \cdot \lambda + C = 0$$

General solution to the quadratic > $$\lambda = \frac{-B + \sqrt{B^2 - 4 \cdot A \cdot C}}{2 \cdot A} \quad \textbf{or} \quad \lambda = \frac{-B - \sqrt{B^2 - 4 \cdot A \cdot C}}{2 \cdot A}$$

Thus: $(1) \cdot \lambda^2 - (a + d) \cdot \lambda + (a \cdot d - b \cdot c) = 0$

has two general solutions:

$$\lambda = \frac{(a + d) + \sqrt{(a + d)^2 - 4 \cdot (1) \cdot (a \cdot d - b \cdot c)}}{2 \cdot (1)} \quad \textbf{or} \quad \lambda = \frac{(a + d) - \sqrt{(a + d)^2 - 4 \cdot (1) \cdot (a \cdot d - b \cdot c)}}{2 \cdot (1)}$$

To calculate eigenvectors, replace each solved eigenvalue back into the Eigenvector equation.
In general, the solution involves eigenvector/eigenvalue pairs.

Note: Equations will be redundant and will therefore specify the equation of a line.
As a result, it is customary give eigenvectors unit length.

Things to remember about eigenvectors:

- **Eigenvectors found from M are not necessarily perpendicular.**

- **If M is symmetric [i.e., $M = M^T$] then eigenvectors are perpendicular.**
  Only in this case is the right-hand picture above is a viable geometric interpretation.

2005 jw099

ORIGIN ≡ 1

EXAMPLE CALCULATION OF EIGENVECTORS & EIGENVALUES
jw099.mcd

Prepared by:
Wm Stein

Given a linear transformation specified by matrix M:

$$M := \begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix}$$

Mathcad's built-in functions:

$evals := eigenvals(M)$ $\quad evals = \begin{pmatrix} 6 \\ 1 \end{pmatrix}$ **< Vector of eigenvalues**

$evecs := eigenvecs(M)$ $\quad evecs = \begin{pmatrix} 0.894 & 0.447 \\ -0.447 & 0.894 \end{pmatrix}$ **< Matrix of eigenvectors, each column is an eigenvector associated with one of the eigenvalues above. Generally, each column is in the same order as the elements in the eigenvalue vector.**

$eigenvec(M, evals_1) = \begin{pmatrix} -0.894 \\ 0.447 \end{pmatrix}$ $\quad eigenvec(M, evals_2) = \begin{pmatrix} 0.447 \\ 0.894 \end{pmatrix}$ **< A way to check which eigenvector belongs with which eigenvalue!**

Using the symbolic processor to find eigenvalues:

The characteristic equation:

$$\left|\begin{pmatrix} 5-\lambda & -2 \\ -2 & 2-\lambda \end{pmatrix}\right| = 0$$ **[CONTROL=]**

[Symbolic/Load symbolic processor]

[put equation in block]

$6 - 7\cdot\lambda + \lambda^2 = 0$ **[Symbolic/simplify]**

[put variable in block]

$\begin{pmatrix} 6 \\ 1 \end{pmatrix}$ **< eigenvalues** **[Symbolic/solve for variable]**

Solving for eigenvectors:

For the first eigenvalue using first row of M:

$\lambda := 6$ **[first eigenvalue]**

$x := 1 \quad y := 1$ **[initialize x & y with any value]**

Given **[keyword for a set of simultaneous equations]**

$5\cdot x - 2\cdot y = 6\cdot x$

$x^2 + y^2 = 1$ **[standardize eigenvector length to unity]**

$e := Find(x, y)$

$e = \begin{pmatrix} 0.894 \\ -0.447 \end{pmatrix}$ $\quad |e| = 1$ **[eigenvector of length = 1]** $\quad E_1 := e$

$(e_1)^2 + (e_2)^2 = 1$ $\quad \lambda_1 := \lambda$

Same result found using second row of M with same eigenvalue:

$\lambda := 6$

Given

$-2\cdot x + 2\cdot y = 6\cdot y$

$x^2 + y^2 = 1$

$e := Find(x, y)$

$e = \begin{pmatrix} 0.894 \\ -0.447 \end{pmatrix}$ $\quad |e| = 1$ **[same result as above]**

For the second eigenvalue using the first row of M:

$\lambda := 1$

Given

$5 \cdot x - 2 \cdot y = 1 \cdot x$

$x^2 + y^2 = 1$

$e := \text{Find}(x, y)$

$E_2 := e$

$e = \begin{pmatrix} 0.447 \\ 0.894 \end{pmatrix}$ $\qquad |e| = 1$

$\lambda_2 := \lambda$

Equivalent calculation using second row of M:

$\lambda := 1$

Given

$-2 \cdot x + 2 \cdot y = 1 \cdot y$

$x^2 + y^2 = 1$

$e := \text{Find}(x, y)$

$e = \begin{pmatrix} 0.447 \\ 0.894 \end{pmatrix}$

Demonstrating identities in diagonalizing matrix M:

Linear transformation M:	Matrix of eigenvectors:	Diagonal matrix of eigenvalues:
$M = \begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix}$	$\varepsilon := \text{augment}(E_1, E_2)$	$D := \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$
	$\varepsilon = \begin{pmatrix} 0.894 & 0.447 \\ -0.447 & 0.894 \end{pmatrix}$	$D = \begin{pmatrix} 6 & 0 \\ 0 & 1 \end{pmatrix}$

$M \cdot \varepsilon = \begin{pmatrix} 5.367 & 0.447 \\ -2.683 & 0.894 \end{pmatrix}$ **<- equivalent ->** $\varepsilon \cdot D = \begin{pmatrix} 5.367 & 0.447 \\ -2.683 & 0.894 \end{pmatrix}$

$\varepsilon^{-1} \cdot M \cdot \varepsilon = \begin{pmatrix} 6 & 0 \\ 0 & 1 \end{pmatrix}$ **= D [diagonalizing matrix M]**

$\varepsilon \cdot D \cdot \varepsilon^{-1} = \begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix}$ **= M [inverse operation recovering M]**

ORIGIN $\equiv 1$

BASIC SETUP OF MULTIVARIATE METHODS
jw139.mcd

Organization of Data Arrays:

$i := 1 .. 10$ **< Index for number of cases = Rows**

$j := 1 .. 3$ **< Index for number of variables = Columns**

Data matrix uses random numbers from Normal Distribution

$X_{i,j} := \text{rnd}(i)$

$$X = \begin{pmatrix} 0.001 & 0.193 & 0.585 \\ 0.701 & 1.646 & 0.348 \\ 2.131 & 0.912 & 0.274 \\ 0.589 & 3.954 & 0.476 \\ 0.045 & 2.658 & 3.009 \\ 0.997 & 2.705 & 0.342 \\ 5.483 & 3.639 & 6.132 \\ 7.647 & 4.315 & 3.697 \\ 7.76 & 7.017 & 8.971 \\ 6.115 & 2.662 & 8.401 \end{pmatrix}$$

$X_{4,2} = 3.954$ **< A single element of data accessed by the indices**

$n := \text{rows}(X)$ $n = 10$ **< number of rows**

$p := \text{cols}(X)$ $p = 3$ **< number of columns**

SAMPLE STATISTICS THE OLD-FASHIONED WAY:

Mean:

$$X_{bar_j} := \frac{1}{n} \cdot \sum_i X_{i,j}$$

$X_{bar_1} = 3.147$

$X_{bar_2} = 2.9701$

$X_{bar_3} = 3.2236$

$$X_{bar} = \begin{pmatrix} 3.147 \\ 2.9701 \\ 3.2236 \end{pmatrix}$$

Standard Deviation:

$$s_j := \sqrt{\frac{1}{n-1} \cdot \sum_i \left(X_{i,j} - X_{bar_j}\right)^2}$$

$s_1 = 3.2231$

$s_2 = 1.9311$

$s_3 = 3.4697$

Variance:

$s_{11} := (s_1)^2$ $s_{11} = 10.3885$

$s_{22} := (s_2)^2$ $s_{22} = 3.7293$

$s_{33} := (s_3)^2$ $s_{33} = 12.0386$

Covariance:

$$s_{12} := \frac{1}{n-1} \cdot \sum_i \left(X_{i,1} - X_{bar_1}\right) \cdot \left(X_{i,2} - X_{bar_2}\right)$$ $s_{12} = 4.2727$

$$s_{13} := \frac{1}{n-1} \cdot \sum_i \left(X_{i,1} - X_{bar_1}\right) \cdot \left(X_{i,3} - X_{bar_3}\right)$$ $s_{13} = 9.2243$

$$s_{23} := \frac{1}{n-1} \cdot \sum_i \left(X_{i,2} - X_{bar_2}\right) \cdot \left(X_{i,3} - X_{bar_3}\right)$$ $s_{23} = 4.3978$

Note: These calculations are cumbersome because each variance or covariance requires its own explicit equation...

Wouldn't it be nice if the computer could figure out by itself how many equations to use based on numbers of rows and columns?

Sample Variance-Covariance Matrix:

$$S := \begin{pmatrix} s_{11} & s_{12} & s_{13} \\ s_{12} & s_{22} & s_{23} \\ s_{13} & s_{23} & s_{33} \end{pmatrix}$$

$$S = \begin{pmatrix} 10.3885 & 4.2727 & 9.2243 \\ 4.2727 & 3.7293 & 4.3978 \\ 9.2243 & 4.3978 & 12.0386 \end{pmatrix}$$

Sample Correlation Matrix:

$$R := \begin{pmatrix} \frac{s_{11}}{\sqrt{s_{11}}\cdot\sqrt{s_{11}}} & \frac{s_{12}}{\sqrt{s_{11}}\cdot\sqrt{s_{22}}} & \frac{s_{13}}{\sqrt{s_{11}}\cdot\sqrt{s_{33}}} \\ \frac{s_{12}}{\sqrt{s_{11}}\cdot\sqrt{s_{22}}} & \frac{s_{22}}{\sqrt{s_{22}}\cdot\sqrt{s_{22}}} & \frac{s_{23}}{\sqrt{s_{22}}\cdot\sqrt{s_{33}}} \\ \frac{s_{13}}{\sqrt{s_{11}}\cdot\sqrt{s_{33}}} & \frac{s_{23}}{\sqrt{s_{22}}\cdot\sqrt{s_{33}}} & \frac{s_{33}}{\sqrt{s_{33}}\cdot\sqrt{s_{33}}} \end{pmatrix}$$

$$R = \begin{pmatrix} 1 & 0.6865 & 0.8248 \\ 0.6865 & 1 & 0.6563 \\ 0.8248 & 0.6563 & 1 \end{pmatrix}$$

SAMPLE STATISTICS AS MATRIX OPERATIONS

Mean:

$i := 1 .. n$

$I_{vec_i} := 1$ **< I_{vec} is the identity vector**

$$X_{mbar} := \frac{1}{n}\cdot X^T \cdot I_{vec}$$

$$I_{vec} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$$X_{mbar} = \begin{pmatrix} 3.147 \\ 2.9701 \\ 3.2236 \end{pmatrix}$$

Variance-Covariance Matrix:

$I := \text{identity}(10)$ **< I is the identity matrix**

$$S_m := \frac{1}{n-1}\cdot X^T \cdot \left(I - \frac{1}{n}\cdot I_{vec}\cdot I_{vec}^T\right)\cdot X$$

$$S_m = \begin{pmatrix} 10.3885 & 4.2727 & 9.2243 \\ 4.2727 & 3.7293 & 4.3978 \\ 9.2243 & 4.3978 & 12.0386 \end{pmatrix}$$

Note: Compare with the calculation of S above.

Feel the magic??

$$I = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Correlation Matrix:

$$D_s := \begin{pmatrix} \sqrt{S_{1,1}} & 0 & 0 \\ 0 & \sqrt{S_{2,2}} & 0 \\ 0 & 0 & \sqrt{S_{3,3}} \end{pmatrix}$$

< D_s is the sample standard deviation matrix

$$D_s = \begin{pmatrix} 3.2231 & 0 & 0 \\ 0 & 1.9311 & 0 \\ 0 & 0 & 3.4697 \end{pmatrix}$$

$$R_m := D_s^{-1} \cdot S \cdot D_s^{-1}$$

Note: Compare with matrix R above.

POOF

$$R_m = \begin{pmatrix} 1 & 0.6865 & 0.8248 \\ 0.6865 & 1 & 0.6563 \\ 0.8248 & 0.6563 & 1 \end{pmatrix}$$

Correlation Matrix with Mathcad trick:

$$i := 1 .. p$$

$$D_{i,i} := \sqrt{S_{i,i}}$$

$$D = \begin{pmatrix} 3.2231 & 0 & 0 \\ 0 & 1.9311 & 0 \\ 0 & 0 & 3.4697 \end{pmatrix}$$

Note: Compare with R and Rm above.

$$RR := D^{-1} \cdot S \cdot D^{-1}$$

$$RR = \begin{pmatrix} 1 & 0.6865 & 0.8248 \\ 0.6865 & 1 & 0.6563 \\ 0.8248 & 0.6563 & 1 \end{pmatrix}$$

Even easier.

LINEAR COMBINATIONS OF VARIABLES
jw142.mcd

Prepared by:
Wm Stein

Using Data Matrix:

$\mathbf{ORIGIN} := 1$

$\mathbf{i} := 1..4$ **<- variables**

$\mathbf{j} := 1..10$ **<- observations**

$\mathbf{X}_{\mathbf{i},\mathbf{j}} := \mathbf{rnd}(\mathbf{i})$

$$\mathbf{X} = \begin{pmatrix} 1.2684\times10^{-3} & 0.1933 & 0.585 & 0.3503 & 0.8228 & 0.1741 & 0.7105 & 0.304 & 0.0914 & 0.1473 \\ 1.977 & 0.2382 & 0.0178 & 1.0633 & 1.2035 & 0.3325 & 0.9016 & 0.1141 & 1.5666 & 1.0398 \\ 2.6279 & 2.8677 & 1.618 & 1.3862 & 2.5867 & 2.339 & 2.9904 & 1.8345 & 0.7986 & 2.5204 \\ 1.5034 & 2.7087 & 0.0353 & 1.1035 & 2.3516 & 3.3504 & 1.9397 & 2.9749 & 1.8319 & 2.9777 \end{pmatrix}$$

Population Statistics:

$\mathbf{n} := \mathbf{cols}(\mathbf{X})$ $\quad \mathbf{n} = 10$

$\mathbf{p} := \mathbf{rows}(\mathbf{X})$ $\quad \mathbf{p} = 4$

$\mathbf{l}_{\mathbf{j}} := 1$ $\quad \mathbf{l}^{\mathbf{T}} = (1 \ \ 1 \ \ 1 \ \ 1 \ \ 1 \ \ 1 \ \ 1 \ \ 1 \ \ 1 \ \ 1)$

$\mathbf{X}_{\mathbf{bar}} := \frac{1}{\mathbf{n}}\cdot\mathbf{X}\cdot\mathbf{l}$ $\quad {\mathbf{X}_{\mathbf{bar}}}^{\mathbf{T}} = (0.338 \ \ 0.8454 \ \ 2.1569 \ \ 2.0777)$

$\mathbf{I} := \mathbf{identity}(\mathbf{n})$

$\mathbf{S} := \frac{1}{\mathbf{n}-1}\cdot\mathbf{X}\cdot\left(\mathbf{I}-\frac{1}{\mathbf{n}}\cdot\mathbf{l}\cdot\mathbf{l}^{\mathbf{T}}\right)\cdot\mathbf{X}^{\mathbf{T}}$

$$\mathbf{S} = \begin{pmatrix} 0.0772 & -0.0452 & 0.0402 & -0.0759 \\ -0.0452 & 0.4299 & -0.0207 & -0.0862 \\ 0.0402 & -0.0207 & 0.5117 & 0.2928 \\ -0.0759 & -0.0862 & 0.2928 & 1.0264 \end{pmatrix}$$

Linear Combinations of Variables:

$\mathbf{a} := \begin{pmatrix} 2 \\ 3 \\ -1 \\ 5 \end{pmatrix}$ $\quad \mathbf{b} := \begin{pmatrix} 0 \\ 2 \\ 1 \\ 4 \end{pmatrix}$ **<- Vectors containing some numbers**

$\mathbf{a}^{\mathbf{T}} = (2 \ \ 3 \ \ -1 \ \ 5)$ **<- vector transpose**

$\mathbf{b}^{\mathbf{T}} = (0 \ \ 2 \ \ 1 \ \ 4)$

$\mathbf{lc}_{\mathbf{a}} := \mathbf{a}^{\mathbf{T}}\cdot\mathbf{X}$

$\mathbf{lc}_{\mathbf{a}} = (10.8228 \ \ 11.7772 \ \ -0.2181 \ \ 8.0221 \ \ 14.4278 \ \ 15.7589 \ \ 10.8339 \ \ 13.9904 \ \ 13.2436 \ \ 15.7819)$

$\mathbf{lc}_{\mathbf{b}} := \mathbf{b}^{\mathbf{T}}\cdot\mathbf{X}$

$\mathbf{lc}_{\mathbf{b}} = (12.5957 \ \ 14.179 \ \ 1.7948 \ \ 7.9271 \ \ 14.4003 \ \ 16.4057 \ \ 12.5524 \ \ 13.9623 \ \ 11.2595 \ \ 16.5106)$

These are linear combinations of 4 original variables, now expressed as single composite variables based on combinations with vector a or vector b.

Statistics of Linear Combinations:

Mean:

$\mathbf{lca}_{\mathbf{bar}} := \mathbf{mean}\left({\mathbf{lc}_{\mathbf{a}}}^{\mathbf{T}}\right)$ $\quad \mathbf{lcb}_{\mathbf{bar}} := \mathbf{mean}\left({\mathbf{lc}_{\mathbf{b}}}^{\mathbf{T}}\right)$ $\quad \mathbf{lca}_{\mathbf{bar}} = 11.4441$

$\mathbf{lcb}_{\mathbf{bar}} = 12.1587$

Variance:

$\mathbf{lca}_{\mathbf{var}} := \frac{\mathbf{n}}{\mathbf{n}-1}\cdot\mathbf{var}\left(\left(\mathbf{lc}_{\mathbf{a}}\right)^{\mathbf{T}}\right)$ $\quad \mathbf{lca}_{\mathbf{var}} = 22.7384$ **<- population estimate using (n-1)**

Covariance:

$$lc_{cov} := \frac{1}{n-1}\cdot\sum_j \left[\left(lc_a^T\right)_j - lca_{bar}\right]\cdot\left[\left(lc_b^T\right)_j - lcb_{bar}\right]$$

$lc_{cov} = 20.2637$

Identities Demonstrated:

Mean of Linear Combinations:

$lca_{bar} = 11.4441$ $\qquad a^T\cdot X_{bar} = (11.4441)$

Variance of Linear Combinations:

$lca_{var} = 22.7384$ $\qquad a^T\cdot S\cdot a = (22.7384)$

Covariance of Linear Combinations:

$lc_{cov} = 20.2637$ $\qquad a^T\cdot S\cdot b = (20.2637)$

Multiple Linear Combinations:

$q := 1..7$

$p := 1..4$

$A_{q,p} := rnd(p)$

$$A = \begin{pmatrix} 0.599 & 1.47 & 1.7172 & 0.6062 \\ 0.4252 & 1.0342 & 2.2546 & 0.676 \\ 0.4919 & 1.3995 & 0.4426 & 0.5664 \\ 0.6929 & 0.8531 & 2.8998 & 0.613 \\ 0.8217 & 0.3827 & 2.4516 & 0.6223 \\ 0.732 & 0.5592 & 2.0467 & 2.8876 \\ 0.123 & 1.6693 & 1.5511 & 1.7048 \end{pmatrix}$$

Let A be a matrix of linear combination coefficients where each row of A represents coefficients for a single linear transformation.

$LC := A\cdot X$

$$LC = \begin{pmatrix} 8.331 & 7.0324 & 3.1765 & 4.8224 & 8.1295 & 6.6407 & 8.0619 & 5.3035 & 4.8397 & 7.7498 \\ 8.9864 & 8.6251 & 3.939 & 5.12 & 9.0162 & 7.9562 & 9.2879 & 6.3943 & 4.6981 & 8.8333 \\ 4.782 & 3.2318 & 1.0488 & 2.899 & 4.5658 & 3.4838 & 4.0333 & 2.8061 & 3.6285 & 4.3296 \\ 10.2295 & 10.3134 & 5.1342 & 5.8461 & 10.5393 & 9.2408 & 11.1221 & 7.4513 & 4.8387 & 10.123 \\ 8.1356 & 8.9658 & 4.4761 & 4.7799 & 8.9413 & 8.0893 & 9.4669 & 6.6419 & 3.7725 & 8.5506 \\ 10.8264 & 13.966 & 3.8517 & 6.8749 & 13.3602 & 14.7754 & 12.7459 & 12.6315 & 7.8674 & 14.4462 \\ 9.9396 & 9.4873 & 2.6715 & 5.8496 & 10.1315 & 9.9163 & 9.5376 & 8.145 & 6.9883 & 10.7395 \end{pmatrix}$$

Matrix LC represents the set of linear combinations (7 in this case).

Statistics of Linear Combinations:

Mean:

$LC_{bar} := \frac{1}{n}\cdot LC\cdot l$ $\qquad LC_{bar}^T = (6.4088 \;\; 7.2857 \;\; 3.4809 \;\; 8.4838 \;\; 7.182 \;\; 11.1346 \;\; 8.3406)$

2005 jw142

Variance-Covariance Matrix:

$$\mathbf{LC_{varcov}} := \frac{1}{\mathbf{n}-1}\cdot\mathbf{LC}\cdot\left(\mathbf{I}-\frac{1}{\mathbf{n}}\cdot\mathbf{l}\cdot\mathbf{l}^{\mathbf{T}}\right)\cdot\mathbf{LC}^{\mathbf{T}}$$

Identities Demonstrated for Multiple Linear Combinations:

Mean:

$$\mathbf{LC_{bar}} = \begin{pmatrix} 6.4088 \\ 7.2857 \\ 3.4809 \\ 8.4838 \\ 7.182 \\ 11.1346 \\ 8.3406 \end{pmatrix} \qquad \mathbf{A}\cdot\mathbf{X_{bar}} = \begin{pmatrix} 6.4088 \\ 7.2857 \\ 3.4809 \\ 8.4838 \\ 7.182 \\ 11.1346 \\ 8.3406 \end{pmatrix}$$

Covariance:

$$\mathbf{LC_{varcov}} = \begin{pmatrix} 3.142 & 3.5468 & 1.7292 & 4.0712 & 3.3666 & 5.1913 & 4.1162 \\ 3.5468 & 4.2128 & 1.7552 & 4.9531 & 4.2132 & 6.5145 & 4.729 \\ 1.7292 & 1.7552 & 1.1876 & 1.8596 & 1.4413 & 2.6354 & 2.401 \\ 4.0712 & 4.9531 & 1.8596 & 5.9301 & 5.0987 & 7.5163 & 5.2827 \\ 3.3666 & 4.2132 & 1.4413 & 5.0987 & 4.4569 & 6.7177 & 4.4676 \\ 5.1913 & 6.5145 & 2.6354 & 7.5163 & 6.7177 & 13.776 & 8.7081 \\ 4.1162 & 4.729 & 2.401 & 5.2827 & 4.4676 & 8.7081 & 6.3289 \end{pmatrix}$$

$$\mathbf{A}\cdot\mathbf{S}\cdot\mathbf{A}^{\mathbf{T}} = \begin{pmatrix} 3.142 & 3.5468 & 1.7292 & 4.0712 & 3.3666 & 5.1913 & 4.1162 \\ 3.5468 & 4.2128 & 1.7552 & 4.9531 & 4.2132 & 6.5145 & 4.729 \\ 1.7292 & 1.7552 & 1.1876 & 1.8596 & 1.4413 & 2.6354 & 2.401 \\ 4.0712 & 4.9531 & 1.8596 & 5.9301 & 5.0987 & 7.5163 & 5.2827 \\ 3.3666 & 4.2132 & 1.4413 & 5.0987 & 4.4569 & 6.7177 & 4.4676 \\ 5.1913 & 6.5145 & 2.6354 & 7.5163 & 6.7177 & 13.776 & 8.7081 \\ 4.1162 & 4.729 & 2.401 & 5.2827 & 4.4676 & 8.7081 & 6.3289 \end{pmatrix}$$

2005 jw155

ORIGIN ≡ 1 **THE MULTIVARIATE NORMAL DISTRIBUTION AND THE VARIANCE-COVARIANCE MATRIX**
jw155.mcd

Prepared by:
Wm Stein

Creating Multivariate Normal Distribution:

$n := 1000$ **<Number of points**

$\mu_1 := 0$ $\mu_2 := 0$

$\sigma_1 := 1$ $\sigma_2 := 1$

$X_1 := \mathrm{rnorm}(n, \mu_1, \sigma_1)$ $X_2 := \mathrm{rnorm}(n, \mu_2, \sigma_2)$

$X := \mathrm{augment}(X_1, X_2)$ **<Data array**

$X =$

	1	2
1	-0.439	0.634
2	-0.679	0.292
3	-0.473	0.141
4	-0.951	0.348
5	-1.686	-0.383
6	0.044	-0.07
7	-0.121	-0.964
8	0.556	-0.559
9	2.192	-2.057
10	0.809	-0.166
11	0.985	0.104
12	0.862	0.712
13	0.916	-1.224
14	0.673	1.536
15	-1.044	-0.095
16	0.069	-0.852

Mean Vector and Variance-Covariance Matrix of Data Matrix X:

$i := 1 .. n$ $l_{n_i} := 1$ $I := \mathrm{identity}(n)$

$X_{bar} := \frac{1}{n} \cdot X^T \cdot l_n$

$X_{bar} = \begin{pmatrix} -0.05975 \\ 0.04368 \end{pmatrix}$

$S_X := \frac{1}{n-1} \cdot X^T \cdot \left(I - \frac{1}{n} \cdot l_n \cdot l_n^T \right) \cdot X$

$S_X = \begin{pmatrix} 0.9701 & -0.00983 \\ -0.00983 & 0.95573 \end{pmatrix}$

Confidence ellipses based on (eq. 4-8):

Making a set of ellipse points:

$\alpha := 0.05$ $df := 2$ **<Set α and df as appropriate**

$c := \sqrt{\mathrm{qchisq}(1 - \alpha, df)}$ $c = 2.448$ **< radius of circle**

Constructing the points:

$j := 1 .. 100$ $\theta_j := \frac{j}{100} \cdot 2 \cdot \pi$

$x_{1_j} := c \cdot \cos(\theta_j)$ $x_{2_j} := c \cdot \sin(\theta_j)$

$x := \mathrm{augment}(x_1, x_2)$ **< points on a circle of radius c**

$xx := \left(S_X \cdot (x)^T \right)^T$ **< points on an ellipse based on S_X**

$\lambda_X := \mathrm{eigenvals}(S_X)$ $\varepsilon_X := \mathrm{eigenvecs}(S_X)$

Major axis of ellipse M:

$p_{X_1} := c \cdot \sqrt{\lambda_{X_1}} \cdot \varepsilon_X^{\langle 1 \rangle}$ $p_{X_1} = \begin{pmatrix} 2.155 \\ -1.094 \end{pmatrix}$

$M_X := \mathrm{augment}(p_{X_1}, -p_{X_1})^T$

Minor axis of ellipse m:

$p_{X_2} := c \cdot \sqrt{\lambda_{X_2}} \cdot \varepsilon_X^{\langle 2 \rangle}$ $p_{X_2} = \begin{pmatrix} 1.081 \\ 2.128 \end{pmatrix}$

$m_X := \mathrm{augment}(p_{X_2}, -p_{X_2})^T$

Simulating Correlation between variables in Matrix X:

$$S_{new} := \begin{pmatrix} 3 & 2 \\ 2 & 5 \end{pmatrix} \qquad Y := \left(S_{new} \cdot X^T\right)^T$$

< New Data Matrix Y

Y =

	1	2
1	-0.048	2.294
2	-1.454	0.102
3	-1.137	-0.239
4	-2.159	-0.164
5	-5.824	-5.289
6	$-9.295 \cdot 10^{-3}$	-0.263
7	-2.29	-5.06
8	0.551	-1.683
9	2.462	-5.901
10	2.095	0.788
11	3.164	2.491
12	4.01	5.282
13	0.3	-4.287
14	5.091	9.025
15	-3.322	-2.562
16	-1.496	-4.121

Mean Vector and Variance-Covariance Matrix of Data Matrix Y:

$$Y_{bar} := \frac{1}{n} \cdot Y^T \cdot l_n \qquad Y_{bar} = \begin{pmatrix} -0.09191 \\ 0.09887 \end{pmatrix}$$

$$S_Y := \frac{1}{n-1} \cdot Y^T \cdot \left(I - \frac{1}{n} \cdot l_n \cdot l_n^T\right) \cdot Y \qquad S_Y = \begin{pmatrix} 12.43577 & 15.19101 \\ 15.19101 & 27.57689 \end{pmatrix}$$

Confidence ellipses based on (eq. 4-8):

Making a set of ellipse points:

$\alpha := 0.05 \qquad df := 2$ **<Set α and df as appropriate**

$c := \sqrt{qchisq(1 - \alpha, df)} \qquad c = 2.448$ **< radius of circle**

Constructing the points:

$$j := 1 .. 100 \qquad \theta_j := \frac{j}{100} \cdot 2 \cdot \pi$$

$$y1_j := c \cdot \cos(\theta_j) \qquad y2_j := c \cdot \sin(\theta_j)$$

$y := \text{augment}(y1, y2)$ **< points on a circle of radius c**

$yy := \left(S_{new} \cdot (y)^T\right)^T$ **< points on an ellipse based on S_{new}**

$$\lambda_y := \text{eigenvals}(S_Y) \qquad \varepsilon_y := \text{eigenvecs}(S_Y)$$

Major axis of ellipse M:

$$p_{y_1} := c \cdot \sqrt{\lambda_{y_1}} \cdot \varepsilon_y^{\langle 1 \rangle} \qquad p_{y_1} = \begin{pmatrix} 3.625 \\ -2.244 \end{pmatrix}$$

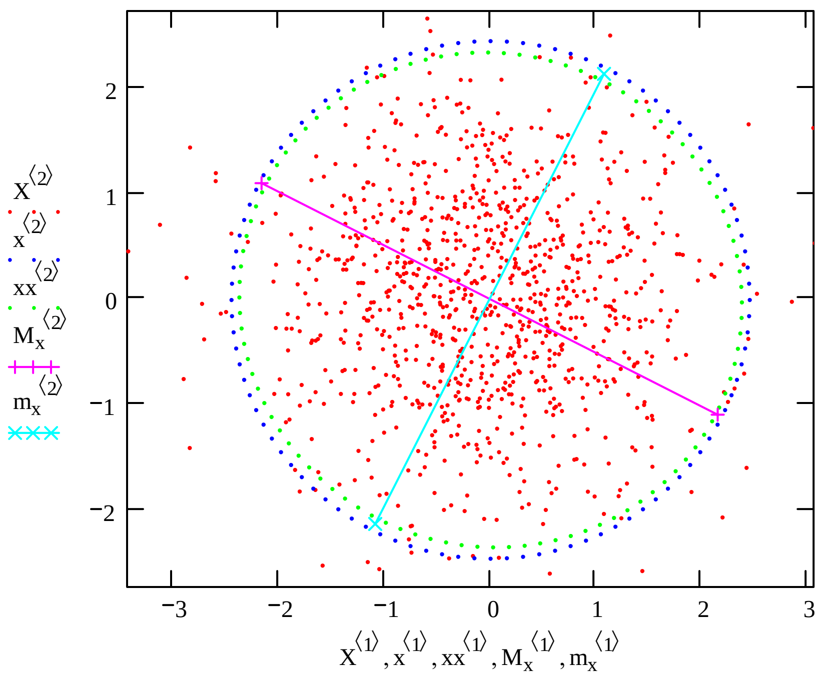
$$M_y := \text{augment}\left(p_{y_1}, -p_{y_1}\right)^T$$

Minor axis of ellipse m:

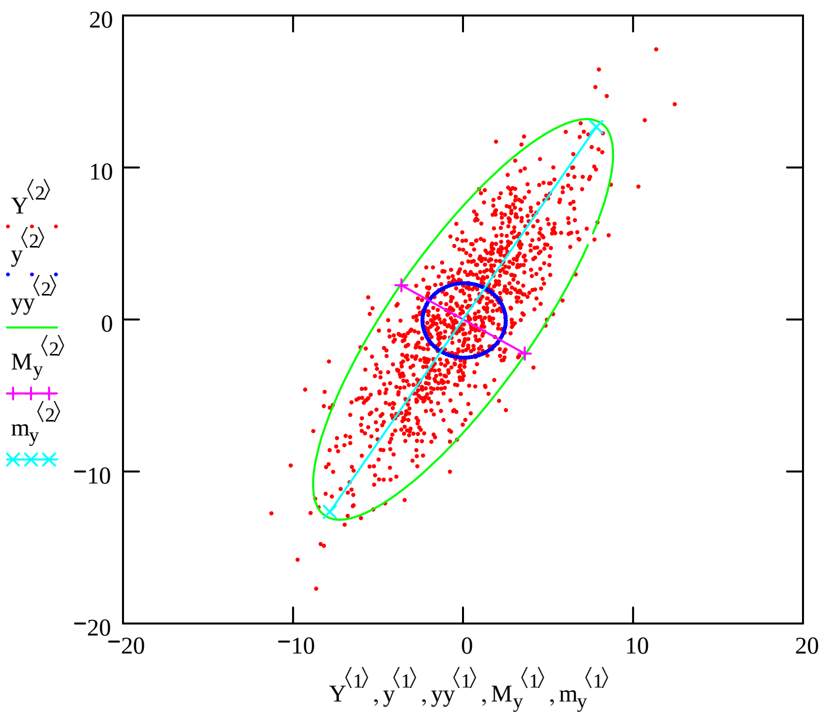
$$p_{y_2} := c \cdot \sqrt{\lambda_{y_2}} \cdot \varepsilon_y^{\langle 2 \rangle} \qquad p_{y_2} = \begin{pmatrix} 7.834 \\ 12.657 \end{pmatrix}$$

$$m_y := \text{augment}\left(p_{y_2}, -p_{y_2}\right)^T$$

2005 jw155



2
1
0
−1
−2
−3
−2
−1
0
1
2
3
$X^{\langle 2 \rangle}$
$x^{\langle 2 \rangle}$
$xx^{\langle 2 \rangle}$
$M_x^{\langle 2 \rangle}$
$m_x^{\langle 2 \rangle}$
$X^{\langle 1 \rangle}, x^{\langle 1 \rangle}, xx^{\langle 1 \rangle}, M_x^{\langle 1 \rangle}, m_x^{\langle 1 \rangle}$



20
10
0
−10
−20
−20
−10
0
10
20
$Y^{\langle 2 \rangle}$
$y^{\langle 2 \rangle}$
$yy^{\langle 2 \rangle}$
$M_y^{\langle 2 \rangle}$
$m_y^{\langle 2 \rangle}$
$Y^{\langle 1 \rangle}, y^{\langle 1 \rangle}, yy^{\langle 1 \rangle}, M_y^{\langle 1 \rangle}, m_y^{\langle 1 \rangle}$

ORIGIN ≡ 1

CONSTRUCTING Q-Q PLOTS
jw179.mcd

Prepared by:
Wm Stein

Using jw Example 4.9 p. 170 in order to compare calculations:

Read in Data:

M := READPRN("DATA\EX4-9.DTA")

$X := M^T$

M =

	1
1	0.8
2	-0.1
3	1.54
4	0.41
5	0.62
6	0.16
7	1.26
8	1.71
9	2.3
10	-1

Calculating summary statistics:

n := cols(X)   n = 10   p := rows(X)   p = 1

i := 1 .. n   j := 1 .. p

Mean vector (X_{bar}):

$l_i := 1$

$X_{bar} := \frac{1}{n} \cdot X \cdot l$ $X_{bar} = (0.77)$

Variance/covariance (S):

I := identity(n)

$S := \frac{1}{n-1} \cdot X \cdot \left(I - \frac{1}{n} \cdot l \cdot l^T \right) \cdot X^T$ $S = (0.941)$

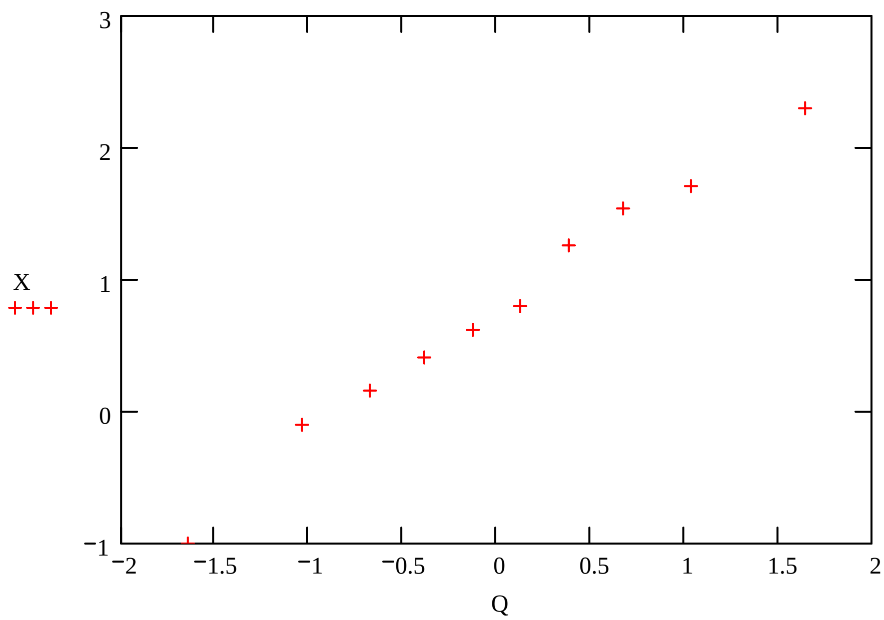
Calculating Probability levels (P) & Quantiles (Q):

X := sort(M)

$P_i := \frac{\left(i - \frac{1}{2} \right)}{n}$

$Q_i := \text{qnorm}\left(P_i, 0, 1 \right)$

$$X = \begin{pmatrix} -1 \\ -0.1 \\ 0.16 \\ 0.41 \\ 0.62 \\ 0.8 \\ 1.26 \\ 1.54 \\ 1.71 \\ 2.3 \end{pmatrix} \quad P = \begin{pmatrix} 0.05 \\ 0.15 \\ 0.25 \\ 0.35 \\ 0.45 \\ 0.55 \\ 0.65 \\ 0.75 \\ 0.85 \\ 0.95 \end{pmatrix} \quad Q = \begin{pmatrix} -1.645 \\ -1.036 \\ -0.674 \\ -0.385 \\ -0.126 \\ 0.126 \\ 0.385 \\ 0.674 \\ 1.036 \\ 1.645 \end{pmatrix}$$



CONSTRUCTING X^2 PLOTS
jw179.mcd

Using jw Problem 1.4 p. 39 & p. 185 in order to compare calculations:

Read in Data:

M := READPRN("DATA\P1-4.DAT")

$X := M^T$

M =

	1	2	3
1	126974	4224	173297
2	96933	3835	160893
3	86656	3510	83219
4	63438	3758	77734
5	55264	3939	128344
6	50976	1809	39080
7	39069	2946	38528
8	36156	359	51038
9	35209	2480	34715
10	32416	2413	25636

Calculating summary statistics:

n := cols(X) n = 10 p := rows(X) p = 3

i := 1 .. n j := 1 .. p

Mean vector (X_{bar}):

$l_i := 1$

$X_{bar} := \frac{1}{n} \cdot X \cdot l$ $X_{bar} = \begin{pmatrix} 62309.1 \\ 2927.3 \\ 81248.4 \end{pmatrix}$

Variance/covariance (S):

I := identity(n)

$S := \frac{1}{n-1} \cdot X \cdot \left(I - \frac{1}{n} \cdot l \cdot l^T\right) \cdot X^T$ $S = \begin{pmatrix} 1.001 \times 10^9 & 2.558 \times 10^7 & 1.512 \times 10^9 \\ 2.558 \times 10^7 & 1.43 \times 10^6 & 4.565 \times 10^7 \\ 1.512 \times 10^9 & 4.565 \times 10^7 & 2.98 \times 10^9 \end{pmatrix}$

Calculating Probability levels (P) & Quantiles (Q):

$i := 1 .. \text{cols}(X^T)$ $j := 1 .. \text{cols}(X^T)$

$d_{i,j} := \left| \left[(X^T)^{\langle i \rangle} - (X^T)^{\langle j \rangle} \right]^T \cdot \left[(X^T)^{\langle i \rangle} - (X^T)^{\langle j \rangle} \right] \right|$

$d = \begin{pmatrix} 0 & 4.382 \times 10^{10} & 1.22 \times 10^{10} \\ 4.382 \times 10^{10} & 0 & 8.736 \times 10^{10} \\ 1.22 \times 10^{10} & 8.736 \times 10^{10} & 0 \end{pmatrix}$

$X^{\langle j \rangle} := \text{sort}(M^{\langle j \rangle})$

$P_i := \frac{\left(i - \frac{1}{2}\right)}{n}$

$Q_i := \text{qnorm}(P_i, 0, 1)$

2005 jw749

STATISTICAL DISTRIBUTIONS AND FINDING CRITICAL VALUES
jw749.mcd

Prepared by:
Wm Stein

Set Index For Plots:

$i := 0 .. 48$

Standard & non-Standard Normal distributions:

$\mu := 0$

$\sigma := 1$

<- set to desired values
Standard Normal: N[μ,σ]) = N[0,1]

$z_i := \frac{(i - 24)\cdot 2}{100}$

$N_i := \mathrm{dnorm}(z_i, \mu, \sigma)$

To find critical value z given:

- cumulative probability P

$P := .95$ **<- set to desired value**

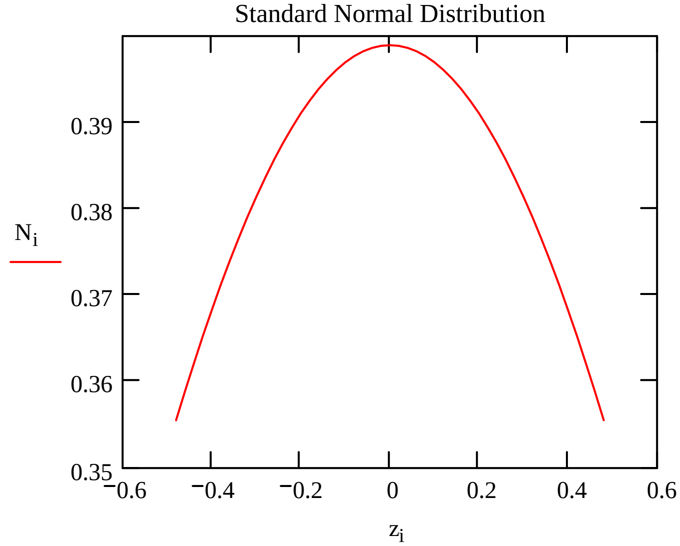
$z := \mathrm{qnorm}(P, \mu, \sigma)$

$z = 1.645$

$z := 2$ **<- set to desired value to go the other direction**

$\mathrm{pnorm}(z, \mu, \sigma) = 0.9772$ **= (P)**

Standard Normal Distribution



Student's t-distribution:

$df_t := 5$ **<- set to desired value**

$x_i := \frac{(i - 24)\cdot 2}{100}$

$t_i := \mathrm{dt}(x_i, df_t)$

To find critical value C given:

- degrees of freedom (df)
- probability of type I error (α)

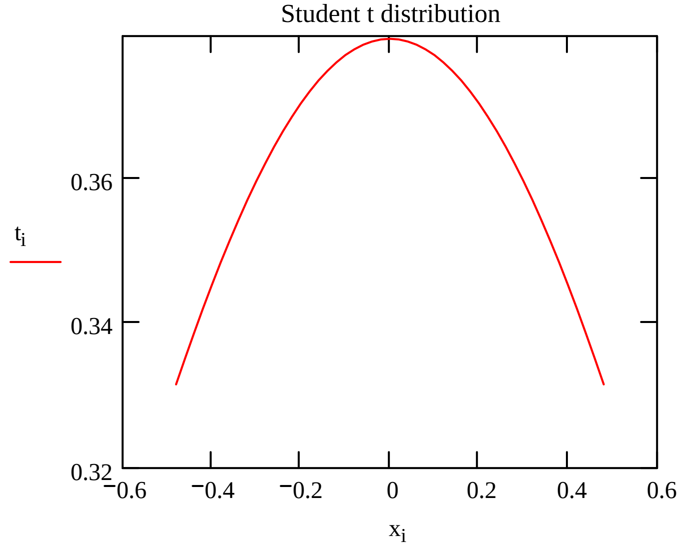
$\alpha_t := .05$ **<- set to desired value**

$C := \mathrm{qt}[(1 - \alpha_t), df_t]$

$C = 2.015$

$\mathrm{pt}(C, df_t) = 0.95$ **= (1 - α)**

Student t distribution



Chi-Square distribution:

$df_{\chi} := 5$ **<- set to desired value**

$y_i := \frac{i}{5}$

$\chi_i := dchisq(y_i, df_{\chi})$

To find critical value C given:

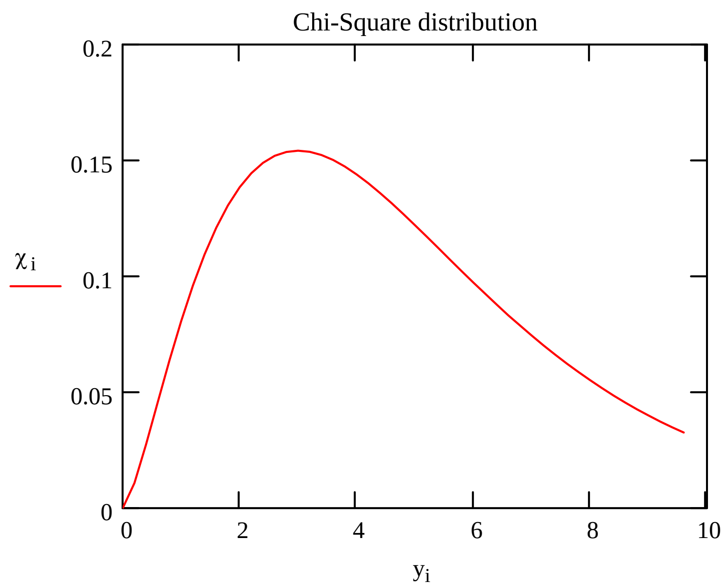
- degrees of freedom (df)
- probability of type I error (α)

$\alpha_{\chi} := 0.05$ **<- set to desired value**

$C := qchisq[(1 - \alpha_{\chi}), df_{\chi}]$

$C = 11.07$

$pchisq(C, df_{\chi}) = 0.95$ **= (1-α)**



F- Distribution:

$df_{F1} := 7$ **<- set to desired value**

$df_{F2} := 12$ **<- set to desired value**

$z_i := \frac{i}{10}$

$F_i := dF(z_i, df_{F1}, df_{F2})$

To find critical value C given:

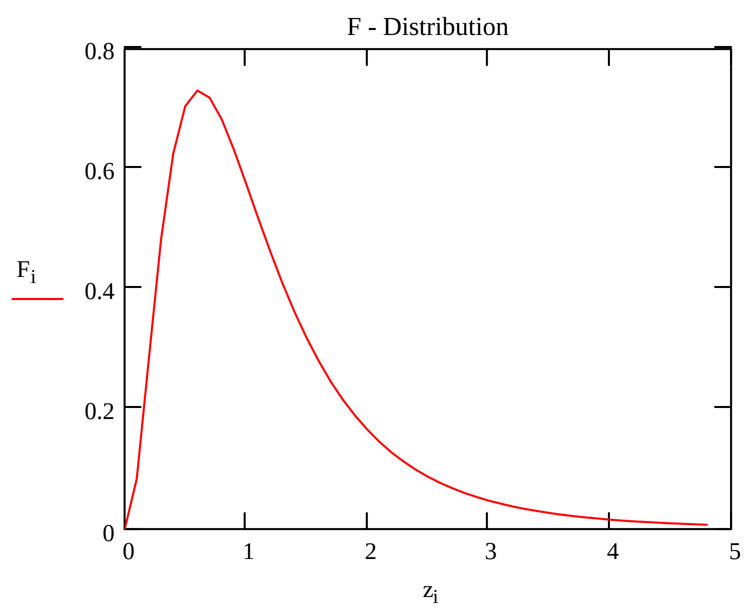
- degrees of freedom (df.F1 & df.F2)
- probability of type I error (α)

$\alpha_F := .1$ **<- set to desired value**

$C := qF[(1 - \alpha_F), df_{F1}, df_{F2}]$

$C = 2.283$

$pF(C, df_{F1}, df_{F2}) = 0.9$ **= (1 - α)**



2005 jw214

ORIGIN ≡ 1

HOTELLING'S T²
jw 214.mcd

Prepared by:
Wm Stein

Read in Data:

Perspiration in n = 20 healthy females (jw p. 214):
Columns:

X_1 = sweat rate
X_2 = Na content
X_3 = K content

X := READPRN("\DATA\T5-1.DAT")

n := rows(X)   p := cols(X)

X =

	1	2	3
1	3.7	48.5	9.3
2	5.7	65.1	8
3	3.8	47.2	10.9
4	3.2	53.2	12
5	3.1	55.5	9.7
6	4.6	36.1	7.9
7	2.4	24.8	14
8	7.2	33.1	7.6
9	6.7	47.4	8.5
10	5.4	54.1	11.3
11	3.9	36.9	12.7
12	4.5	58.8	12.3
13	3.5	27.8	9.8
14	4.5	40.2	8.4
15	1.5	13.5	10.1
16	8.5	56.4	7.1

Summary statistics of the sample:

i := 1 .. n   j := 1 .. p

$l_{n_i} := 1$ I := identity(n)

$X_{bar} := \frac{1}{n} \cdot X^T \cdot l_n$ $X_{bar} = \begin{pmatrix} 4.64 \\ 45.4 \\ 9.965 \end{pmatrix}$

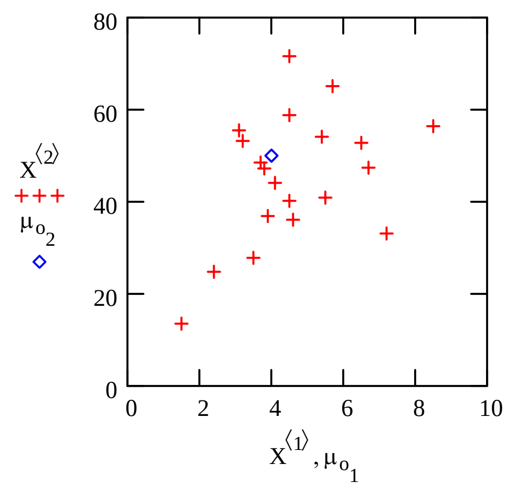
$S := \frac{1}{n-1} \cdot X^T \cdot \left(I - \frac{1}{n} \cdot l_n \cdot l_n^T \right) \cdot X$ $S = \begin{pmatrix} 2.87937 & 10.01 & -1.80905 \\ 10.01 & 199.78842 & -5.64 \\ -1.80905 & -5.64 & 3.62766 \end{pmatrix}$

Specify a test vector μ_0:

$\mu_o := \begin{pmatrix} 4 \\ 50 \\ 10 \end{pmatrix}$

Hotelling's T² statistic (jw Eq. 5-6 p. 212):

$T_{sq} := n \cdot (X_{bar} - \mu_o)^T \cdot S^{-1} \cdot (X_{bar} - \mu_o)$ $T_{sq} = (9.739)$



Hypothesis testing:

$H_0 : \mu = \mu_0$

$H_1 : \mu <> \mu_0$

Assumption: X [X_1,X_2, ... , X_n] rs $N_p(\mu,\Sigma)$

Stringency of the test: α := 0.10   **< set as desired**

If assumptions hold and H_0 is true then:

$C := \frac{(n-1) \cdot p}{(n-p)} \cdot qF(1 - \alpha, p, n-p)$ C = 8.173   **< jw eq. 5-6 p. 212**

NOTE: qF(1-α) is substituted here for F(α) in the text.

Decision Rule: Reject H_0 if T > C

Decision := if(T_{sq_1} > C, 1, 0)   Decision = 1   **< 0 = Do not reject H_0**
1 = Reject H_0

Liklihood Ratio test:

Maximum liklihood estimates of μ & Σ (jw Result 4.11 p. 171) and Σ_0:

$$\mu_{hat} := X_{bar} \qquad \Sigma_{hat} := \frac{(n-1)}{n}\cdot S \qquad \Sigma_{0.hat} := \sum_i \left[\left(X^T\right)^{\langle i\rangle} - \mu_o\right]\cdot\left[\left(X^T\right)^{\langle i\rangle} - \mu_o\right]^T \qquad \left|\Sigma_{0.hat}\right| = 1.226\times10^7$$

$$\left|\Sigma_{hat}\right| = 1.014\times10^3$$

Liklihood Ratio (Λ) & Wilks' lambda (jw Eq. 5-13 p. 217):

$$\Lambda := \left(\frac{\left|n\cdot\Sigma_{hat}\right|}{\left|\Sigma_{0.hat}\right|}\right)^{\frac{n}{2}} \qquad \Lambda = 0.015953 \qquad \Lambda^{\frac{2}{n}} = 0.661128$$

< Λ = Wilks' lambda - value is 1.0 when $X_{bar} = \mu_0$ but decreases as μ_0 increases in distance from the sample mean.

$$\left[1 + \frac{T_{sq}}{(n-1)}\right]^{-1} = (0.661128)$$

< Equivalent value in terms of Hotellings T^2 (jw Result 5.1 p. 218)

Converting to Hotellings T^2:

$$\left[\Lambda^{-\left(\frac{2}{n}\right)} - 1\right]\cdot(n-1) = 9.739 \qquad T_{sq} = (9.739)$$

< Solving for T^2 (same as T^2 above)
< Now run the above test...

Confidence intervals:

$\alpha := 0.05$ **< Set probability of Type 1 error**

The multivariate confidence ellipsoid (jw Eq. 5-18 p. 221):

$$\lambda := \text{reverse}(\text{sort}(\text{eigenvals}(S)))$$

$$\varepsilon^{\langle j\rangle} := \text{eigenvec}\left(S, \lambda_j\right)$$

$$\lambda = \begin{pmatrix} 200.46246 \\ 4.53159 \\ 1.30139 \end{pmatrix} \qquad \varepsilon = \begin{pmatrix} 0.05084 & -0.5737 & 0.81748 \\ 0.99828 & 0.05302 & -0.02488 \\ -0.02907 & 0.81735 & 0.57541 \end{pmatrix}$$

< Coordinates of each column vector of ε gives the directions of confidence elipsoid

$$C := \sqrt{\frac{p\cdot(n-1)}{n\cdot(n-p)}\cdot qF(1-\alpha, p, n-p)} \qquad C = 0.732$$

< C gives the boundary for the confidence ellipsoid for μ - see jw eq. 5-18 p. 221

NOTE: To obtain the values of F(α) reported in the text, qF(1-α) must be used in MathCad here.

$$i := 1 .. p$$

$$L_i := C\cdot\sqrt{\lambda_i}$$

Multivariate simultaneous confidence ellipsoid:

$$X_{bar} = \begin{pmatrix} 4.64 \\ 45.4 \\ 9.965 \end{pmatrix}$$

< Center of ellipsoid

$$L = \begin{pmatrix} 10.36503 \\ 1.5584 \\ 0.83514 \end{pmatrix}$$

< L are half-lengths of the axes of the confidence ellipsoid for μ in the directions of ε

Simultaneous T^2 confidence intervals:

$CI_{lower_i} := X_{bar_i} - C\cdot\sqrt{S_{i,i}}$ $\quad$ $CI_{upper_i} := X_{bar_i} + C\cdot\sqrt{S_{i,i}}$ **< jw eq. 5-24 p. 225 (slightly modified)**

$CI := \text{augment}(CI_{lower}, CI_{upper})$

$$X_{bar} = \begin{pmatrix} 4.64 \\ 45.4 \\ 9.965 \end{pmatrix}$$ **< Mean values**

$$CI = \begin{pmatrix} 3.39777 & 5.88223 \\ 35.05241 & 55.74759 \\ 8.57066 & 11.35934 \end{pmatrix}$$ **< T^2 confidence intervals**

Bonferroni simultaneous confidence intervals:

$c := qt\left(1 - \frac{\alpha}{2\cdot p}, n-1\right)$ $\quad$ $c = 2.625$ **< Critical value c based on t distribution**

NOTE: qt(1 - α/2p) is substituted for t(α/2p) in MathCad as above.

$ci_{lower_i} := X_{bar_i} - c\cdot\sqrt{\frac{S_{i,i}}{n}}$ $\quad$ $ci_{upper_i} := X_{bar_i} + c\cdot\sqrt{\frac{S_{i,i}}{n}}$ **< jw eq. 5-29 p. 232**

$ci := \text{augment}(ci_{lower}, ci_{upper})$

$$X_{bar} = \begin{pmatrix} 4.64 \\ 45.4 \\ 9.965 \end{pmatrix}$$ **< Mean values**

$$ci = \begin{pmatrix} 3.64395 & 5.63605 \\ 37.10308 & 53.69692 \\ 8.84699 & 11.08301 \end{pmatrix}$$ **Bonferroni confidence intervals**

HOTELLING'S T^2
jw 214A.mcd

$ORIGIN \equiv 1$

Prepared by:
Wm Stein

Verifying calculations using jw Example 5.3 p. 221, 226 & 233.
Radiation for Microwave ovens.

Only mean vector and variance-covariance matrix given:

$$X_{bar} := \begin{pmatrix} .564 \\ .603 \end{pmatrix} \qquad S := \begin{pmatrix} .0144 & .0117 \\ .0117 & .0146 \end{pmatrix} \qquad n := 42 \qquad p := 2$$

$$S^{-1} = \begin{pmatrix} 199.046 & -159.509 \\ -159.509 & 196.319 \end{pmatrix}$$

Specify a test vector μ_0:

$$\mu_o := \begin{pmatrix} .562 \\ .589 \end{pmatrix}$$

Hotelling's T^2 statistic (jw Eq. 5-6 p. 212):

$$T_{sq} := n \cdot \left(X_{bar} - \mu_o\right)^T \cdot S^{-1} \cdot \left(X_{bar} - \mu_o\right) \qquad T_{sq} = (1.274)$$

Hypothesis testing:

$H_0 : \mu = \mu_0$

$H_1 : \mu <> \mu_0$

Assumption: X $[X_1, X_2, \ldots, X_n]$ rs $N_p(\mu, \Sigma)$

Stringency of the test: $\alpha := 0.05$ **< set as desired**

If assumptions hold and H_0 is true then:

$$C := \frac{(n-1) \cdot p}{(n-p)} \cdot qF(1 - \alpha, p, n - p) \qquad C = 6.625$$

< jw eq. 5-6 p. 212

NOTE: qF(1-α) is substituted here for F(α) in the text.

Decision Rule: Reject H_0 if T > C

$$Decision := if\left(T_{sq_1} > C, 1, 0\right) \qquad Decision = 0$$

< 0 = Do not reject H_0
1 = Reject H_0

Confidence intervals:

$\alpha := 0.05$ **< Set probability of Type 1 error**

The multivariate confidence ellipsoid (jw Eq. 5-18 p. 221):

$i := 1 .. p$

$\lambda := \text{reverse}(\text{sort}(\text{eigenvals}(S)))$

$\varepsilon^{\langle i \rangle} := \text{eigenvec}(S, \lambda_i)$

$\lambda = \begin{pmatrix} 0.0262 \\ 0.0028 \end{pmatrix}$ $\varepsilon = \begin{pmatrix} 0.70408 & -0.71012 \\ 0.71012 & 0.70408 \end{pmatrix}$

< Coordinates of each column vector of ε gives the directions of confidence ellipsoid

$C := \sqrt{\frac{p \cdot (n-1)}{n \cdot (n-p)} \cdot qF(1-\alpha, p, n-p)}$ $C = 0.397$

< C gives the boundary for the confidence ellipsoid for μ - see jw eq. 5-18 p. 221

$L_i := C \cdot \sqrt{\lambda_i}$

NOTE: To obtain the values of F(α) reported in the text, qF(1-α) must be used in MathCad here.

Multivariate simultaneous confidence ellipsoid:

$X_{bar} = \begin{pmatrix} 0.564 \\ 0.603 \end{pmatrix}$ **< Center of ellipsoid**

$L = \begin{pmatrix} 0.06429 \\ 0.02101 \end{pmatrix}$ **< L are half-lengths of the axes of the confidence ellipsoid for μ in the directions of ε**

Note: calculations are for L_2 is somewhat off what jw calculate on p. 212

Simultaneous T^2 confidence intervals:

$CI_{lower_i} := X_{bar_i} - C \cdot \sqrt{S_{i,i}}$ $CI_{upper_i} := X_{bar_i} + C \cdot \sqrt{S_{i,i}}$ **< jw eq. 5-24 p. 225 (slightly modified)**

$CI := \text{augment}(CI_{lower}, CI_{upper})$

$X_{bar} = \begin{pmatrix} 0.564 \\ 0.603 \end{pmatrix}$ **< Mean values** $CI = \begin{pmatrix} 0.51634 & 0.61166 \\ 0.55501 & 0.65099 \end{pmatrix}$ **< T^2 confidence intervals**

Bonferroni simultaneous confidence intervals:

$c := qt\left(1 - \frac{\alpha}{2 \cdot p}, n-1\right)$ $c = 2.327$ **< Critical value c based on t distribution**

NOTE: qt(1 - α/2p) is substituted for t(α/2p) in MathCad as above.

$ci_{lower_i} := X_{bar_i} - c \cdot \sqrt{\frac{S_{i,i}}{n}}$ $ci_{upper_i} := X_{bar_i} + c \cdot \sqrt{\frac{S_{i,i}}{n}}$ **< jw eq. 5-29 p. 232**

$ci := \text{augment}(ci_{lower}, ci_{upper})$

$X_{bar} = \begin{pmatrix} 0.564 \\ 0.603 \end{pmatrix}$ **< Mean values** $ci = \begin{pmatrix} 0.52092 & 0.60708 \\ 0.55962 & 0.64638 \end{pmatrix}$ **< Bonferroni confidence intervals**

TEST STATS:
stattest.mcd

Prepared by:
Wm Stein

TEST	TEST STATISTIC	CONF. INTERVAL

Student's t

One population:

$X_1, X_2, \ldots, X_n$ ~ rs N(μ,σ)

$$t = \frac{\overline{X} - \mu_o}{\left(\frac{s}{\sqrt{n}}\right)}$$

$$CI = \overline{X} \pm C \cdot \frac{s}{\sqrt{n}}$$

H0: $\mu = \mu_o$

Reject H0 if: $T \geq C$

$t \sim t_{\alpha}(n-1)$

Two populations:

$X_{11}, X_{12}, \ldots, X_{1n}$ ~ rs N(μ_1,σ_1)

$X_{21}, X_{22}, \ldots, X_{2n}$ ~ rs N(μ_2,σ_2)

$\sigma_1^2 = \sigma_2^2$

X_{1j} independent of X_{2j}

$$t = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{sp^2 \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where:

$$sp := \sqrt{\frac{(n_1 - 1) \cdot s_1^2 + (n_2 - 1) \cdot s_2^2}{(n_1 + n_2 - 2)}}$$

$T \sim t_{\alpha}(n1+n2-2)$

H0: $\mu_1 = \mu_2$

Reject H0 if: $T \geq C$

$$CI = \overline{X_1} - \overline{X_2} \pm C \cdot \sqrt{sp^2 \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

Paired populations:

$X_{11}, X_{12}, \ldots, X_{1n}$ ~ rs N(μ_1,σ_1)

$X_{21}, X_{22}, \ldots, X_{2n}$ ~ rs N(μ_2,σ_2)

$d_1, d_2, \ldots, d_n$ ~ rs N(μ_d,σ_d)

where:

$d_j = X_{1j} - X_{2j}$

$$t = \frac{\overline{d}}{\left(\frac{s_d}{\sqrt{n}}\right)}$$

$$CI = \overline{d} \pm C \cdot \frac{s_d}{\sqrt{n}}$$

H0: $\mu d = 0$

$t \sim t_{\alpha}(n-1)$

Reject H0 if: $T \geq C$

TEST	TEST STATISTIC	CONF. Ellipsoid

Hotelling's T^2

One population:

$X = [X_1, X_2, \ldots X_n] \sim rs\ N_p(\mu,\Sigma)$
where:
X_i are vectors,
n = number of observations
p = number of variables

X_{bar} = multivariate mean vector
S = var/cov matrix of X

H0: $\mu = \mu_o$
where: μ and μ_o are vectors

$$T^2 = n(X_{bar} - \mu_o)^T S^{-1} (X_{bar} - \mu_o)$$

$$T^2 \sim \frac{(n-1)\cdot p}{n-p} F_{\alpha}\ p, n-p$$

CE = X_{bar} +/- [sqrt(λ_i) sqrt([p(n-1)/n(n-p)] F_{α} p, n-p] e_i

where: λ_i = i^{th} eigenvalue
e_i = i^{th} eigenvector of length 1 of S

Two populations:

$X_1 = [X_{1,1}, X_{1,2}, \ldots X_{1,n1}] \sim rs\ N_p(\mu_1,\Sigma_1)$
$X_2 = [X_{2,1}, X_{2,2}, \ldots X_{2,n2}] \sim rs\ N_p(\mu_2,\Sigma_2)$
where:
$X_{i,j}$ are vectors,
n_1 = number of observations in population 1
n_2 = number of observations in population 2
p = number of variables

X_1 independent of X_2

If n_1 & n_2 small, then $\Sigma_1 = \Sigma_2$

H0: $\mu_1 - \mu_2 = \delta_o$
where: μ_1, μ_2 & δo are vectors

where: S_1 - var/cov matrix of X_1
S_2 - var/cov matrix of X_2

and: $$S_p = \frac{(n_1-1)\cdot S_1 + (n_2-1)\cdot S_2}{n_1+n_2-2}$$

$$T^2 = (X_{bar-1} - X_{bar-2} - \delta_o)^T [([1/n_1] + [1/n_2]) S_p]^{-1} (X_{bar-1} - X_{bar-2} - \delta_o)$$

$$T^2 \sim \frac{(n_1+n_2-2)\cdot p}{(n_1+n_2-p-1)} F_{\alpha}\ p, n1+n2-p-1$$

CE = (X_{bar-1} - X_{bar-2}) +/- [sqrt(λ_i) sqrt([(1/n_1) + (1/n_2)] F_{α} p, n_1+n_2-p-1)] e_i

where: λ_i = i^{th} eigenvalue
e_i = i^{th} eigenvector of length 1 of S_p

Paired Experimental design:

$X_1 = [X_{1,1}, X_{1,2}, \ldots, X_{1,n1}]$
$X_2 = [X_{2,1}, X_{2,2}, \ldots, X_{2,n2}]$
$D = [\delta_1, \delta_2, \ldots, \delta_n] \sim rs\ N_p(\delta,\Sigma_\delta)$
where: $X_{-,j}$ are paired vectors and
$d_j = X_{1,j} - X_{2,j}$
n = number of paired observations
p = number of variables

D_{bar} = multivariate mean vector of D
S_d = var/cov matrix of D

H0: $\delta = 0$

$$T^2 = n\ d_{bar}^T S_d^{-1} d_{bar}$$

$$T^2 \sim \frac{(n-1)\cdot p}{(n-p)} F_{\alpha}\ p, n-p$$

CE = d_{bar} +/- [sqrt(λ_i) sqrt([p(n-1)/n(n-p)] F_{α} p, n-p)] e_i

where: λ_i = i^{th} eigenvalue
e_i = i^{th} eigenvector of length 1 of S_d

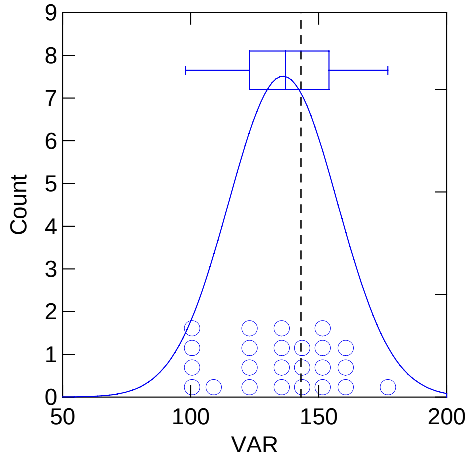
2005 Systat Test output

SYSTAT OUTPUT

IMPORT successfully completed. ONE POPULATON

One-sample t test of VAR with 24 cases;   Ho: Mean =   143.000

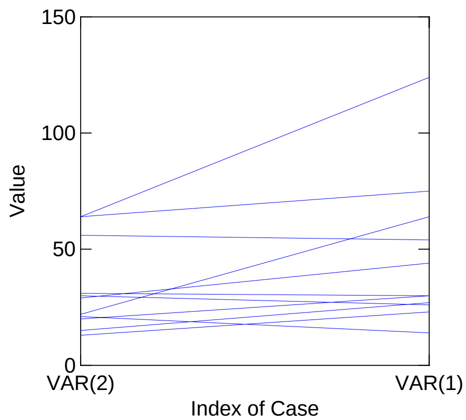
```
  Mean =      136.042           90.00% CI  =   128.608 to    143.475
   SD =       21.249                          t =        -1.604
                                df =    23    Prob =       0.122
```



IMPORT successfully completed. PAIRED OBSERVATIONS

Paired samples t test on VAR(1) vs VAR(2) with 11 cases

```
 Mean VAR(1)      =      46.455
 Mean VAR(2)      =      33.182
  Mean Difference =       13.273   95.00% CI  =     -0.473 to     27.018
   SD Difference =       20.460                     t =        2.152
                                df =    10    Prob =       0.057
```



IMPORT successfully completed. TWO POPULATIONS

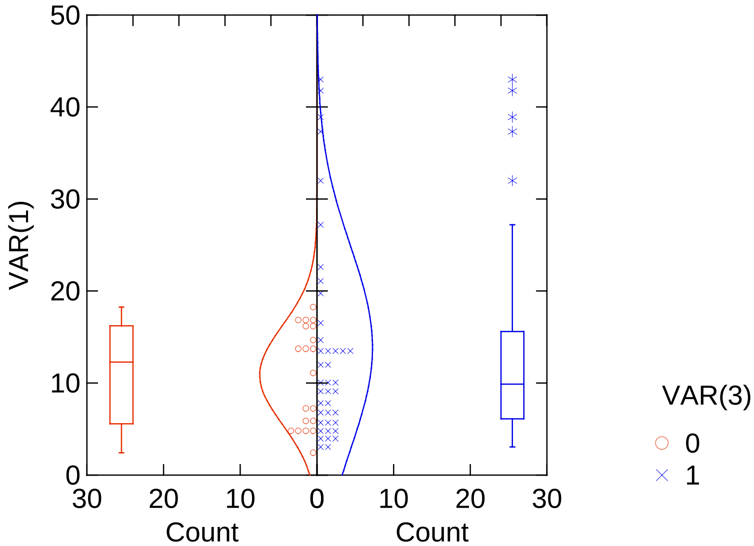
2005 Systat Test output

Two-sample t test on VAR(1) grouped by VAR(3)

Group	N	Mean	SD
0	20	10.875	5.343
1	40	13.840	11.015

Separate Variance t = -1.404 df = 58.0 Prob = 0.166
Difference in Means = -2.965 95.00% CI = -7.193 to 1.263

Pooled Variance t = -1.135 df = 58 Prob = 0.261
Difference in Means = -2.965 95.00% CI = -8.193 to 2.263



UNIVARIATE MEANS TESTS - t tests
ONE POPULATION

$ORIGIN \equiv 1$

Prepared by:
Wm Stein

Read in Data:

Carapase Length for a sample of Female Painter Turtles - Part of jw Table 6.9 p 339):

$X := READPRN("\backslash DATA\backslash T6\text{-}9fem\text{-}len.DAT")$

$n := rows(X)$ $\quad p := cols(X)$ $\quad n = 24$ $\quad p = 1$

X =

	1
1	98
2	103
3	103
4	105
5	109
6	123
7	123
8	133
9	133
10	133
11	134
12	136
13	138
14	138
15	141
16	147

Summary statistics of the sample:

$i := 1..n \quad j := 1..p$

$l_{n_i} := 1 \quad I := identity(n)$

Note that the multivariate approach to summar statistics works for the univariate case as well!

$$X_{bar} := \frac{1}{n} \cdot X^T \cdot l_n \qquad X_{bar} = (136.04167) \qquad mean(X) = 136.042$$

$$S := \frac{1}{n-1} \cdot X^T \cdot \left(I - \frac{1}{n} \cdot l_n \cdot l_n^T\right) \cdot X \qquad S = (451.51993) \qquad Var(X) = 451.52$$

Specify a test value μ_0:

$\mu_o := 143$ **< Set this to test whether the mean is "close enough" to μ_0:**

Student t-test statistic:

$$t := \left| \frac{mean(X) - \mu_o}{\sqrt{\frac{Var(X)}{n}}} \right| \qquad t = 1.604$$

$var(X) = 432.707$

Hypothesis testing:

$H_0 : \mu = \mu_0$

$H_1 : \mu <> \mu_0$ **< two-tailed test!**

Assumption: X [X_1,X_2, ... , X_n] rs $N_p(\mu,\Sigma)$ $\quad p = 1$

Stringency of the test: $\alpha := 0.1$ **< set as desired**

If assumptions hold and H_0 is true then:

$$C := qt\left(1 - \frac{\alpha}{2}, n-1\right) \qquad C = 1.714$$

NOTE: qt(1-α) is used here in the MathCad function for the t distribution & jw Table 2 p. 750

$n = 24$

Decision Rule: Reject H_0 if T > C

$Decision := if(t > C, 1, 0) \qquad Decision = 0$

< 0 = Do not reject H_0
1 = Reject H_0

Confidence intervals:

$$CI := \begin{pmatrix} mean(X) + C \cdot \sqrt{\frac{Var(X)}{n}} \\ mean(X) - C \cdot \sqrt{\frac{Var(X)}{n}} \end{pmatrix} \qquad mean(X) = 136.042 \qquad CI = \begin{pmatrix} 143.475 \\ 128.608 \end{pmatrix}$$

UNIVARIATE MEANS TESTS - t tests
TWO POPULATIONS - PAIRED OBSERVATIONS

Read in Data:

Effluent Data - Suspended Solids compared in eleven split samples sent to two labs - Part of jw Table 6.1 p 275):

$X := \text{READPRN}(\text{"\\DATA\\T6-1pairedSS.DAT"})$

$n := \text{rows}(X)$ $p := \text{cols}(X)$ $n = 11$ $p = 2$

X =

	1	2
1	27	15
2	23	13
3	64	22
4	44	29
5	30	31
6	75	64
7	26	30
8	124	64
9	54	56
10	30	20
11	14	21

Summary statistics of the sample:

$i := 1..n$ $j := 1..p$

$ll_{n_i} := 1$ $I := \text{identity}(n)$

Note: there's a bug in MathCad that prevents me reusing I_n here...

$X_{bar} := \frac{1}{n} \cdot X^T \cdot ll_n$ $X_{bar} = \begin{pmatrix} 46.45455 \\ 33.18182 \end{pmatrix}$ $\text{mean}(X^{\langle 1 \rangle}) = 46.455$ $\text{mean}(X^{\langle 2 \rangle}) = 33.182$

$S := \frac{1}{n-1} \cdot X^T \cdot \left(I - \frac{1}{n} \cdot ll_n \cdot ll_n^T\right) \cdot X$ $S = \begin{pmatrix} 1014.07273 & 479.60909 \\ 479.60909 & 363.76364 \end{pmatrix}$ $\text{Var}(X^{\langle 1 \rangle}) = 1014.0727$ $\text{Var}(X^{\langle 2 \rangle}) = 363.7636$

Calculate differences & Summary statistics:

$D := X^{\langle 1 \rangle} - X^{\langle 2 \rangle}$ $D_{bar} := \frac{1}{n} \cdot D^T \cdot ll_n$ $D_{bar} = (13.273)$

D =

	1
1	12
2	10
3	42
4	15
5	-1
6	11
7	-4
8	60
9	-2
10	10
11	-7

Student t-test statistic:

$\text{mean}(D) = 13.273$

$\text{Var}(D) = 418.618$

$$t := \left| \frac{\text{mean}(D)}{\sqrt{\frac{\text{Var}(D)}{n}}} \right| \qquad t = 2.152$$

Hypothesis testing:

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 <> \mu_2$ **< two-tailed test!**

Assumption: D [$D_1, D_2, ..., D_n$] rs $N_1(\mu, \Sigma)$

Stringency of the test: $\alpha := 0.05$ **< set as desired**

If assumptions hold and H_0 is true then:

$C := qt\left(1 - \frac{\alpha}{2}, n-1\right)$ $C = 2.228$

NOTE: qt(1-α) is used here in the MathCad function for the t distribution & jw Table 2 p. 750

$n = 11$

Decision Rule: Reject H_0 if T > C

$\text{Decision} := \text{if}(t > C, 1, 0)$ $\text{Decision} = 0$

< 0 = Do not reject H_0
1 = Reject H_0

Confidence intervals:

$$CI := \begin{pmatrix} \text{mean}(D) + C \cdot \sqrt{\frac{\text{Var}(D)}{n}} \\ \text{mean}(D) - C \cdot \sqrt{\frac{\text{Var}(D)}{n}} \end{pmatrix} \qquad \text{mean}(D) = 13.273 \qquad CI = \begin{pmatrix} 27.018 \\ -0.473 \end{pmatrix}$$

UNIVARIATE MEANS TESTS - t tests
TWO POPULATIONS WITH EQUAL VARIANCES

Read in Data:

Lizard Data - Mass of individuals of two species C & S
Part of jw Table 6-7 p. 330

$X := \text{READPRN}(\text{"\textbackslash DATA\textbackslash T6-7MassC.DAT"})$ $\quad n_X := \text{rows}(X) \quad p_X := \text{rows}(X)$

$n_X = 20 \quad p_X = 20$

$Y := \text{READPRN}(\text{"\textbackslash DATA\textbackslash T6-7MassS.DAT"})$ $\quad n_Y := \text{rows}(Y) \quad p_Y := \text{rows}(Y)$

$n_Y = 40 \quad p_Y = 40$

X =

	1
1	7.513
2	5.032
3	5.867
4	11.088
5	2.419
6	13.61
7	18.247
8	16.832
9	15.91
10	17.035
11	16.526
12	4.53
13	7.23
14	5.2
15	13.45
16	14.08
17	14.665
18	6.092
19	5.264
20	16.902

Y =

	1
1	13.911
2	5.236
3	37.331
4	41.781
5	31.995
6	3.962
7	4.367
8	3.048
9	4.838
10	6.525
11	22.61
12	13.342
13	4.109
14	12.369
15	7.12
16	21.077
17	42.989
18	27.201
19	38.901
20	19.747
21	14.666
22	4.79
23	5.02
24	5.22
25	5.69
26	6.763
27	9.977
28	8.831
29	9.493
30	7.811
31	6.685
32	11.98
33	16.52
34	13.63
35	13.7
36	10.35
37	7.9
38	9.103
39	13.216
40	9.787

Summary statistics of the sample:

$\text{mean}(X) = 10.875 \qquad \text{mean}(Y) = 13.84$

$\text{Var}(X) = 28.543 \qquad \text{Var}(Y) = 121.339$

Note that the variances are not close so we should not continue with this test without first using a variance-stabilizing transformation of the data. However, as an example, we will continue anyway.

Calculate pooled variance S_p:

$$S_p := \frac{(n_X - 1)\cdot \text{Var}(X) + (n_Y - 1)\cdot \text{Var}(Y)}{(n_X + n_Y - 2)} \qquad S_p = 90.941$$

Student t-test statistic:

$$t := \left| \frac{\text{mean}(X) - \text{mean}(Y)}{\sqrt{S_p \cdot \left(\frac{1}{n_X} + \frac{1}{n_Y}\right)}} \right| \qquad t = 1.135$$

Hypothesis testing:

Assumptions: $X\ [X_1, X_2, \ldots, X_n]$ **rs** $N_1(\mu, \Sigma)$
$Y\ [Y_1, Y_2, \ldots, Y_n]$ **rs** $N_1(\mu, \Sigma)$
Var(X) = Var(Y)
X's are independent of Y's

$H_0 : \mu_1 = \mu_2$ **< ie the difference is zero**

$H_1 : \mu_1 <> \mu_2$ **< two-tailed test!**

Stringency of the test: $\alpha := 0.05$ **< set as desired**

If assumptions hold and H_0 is true then:

$$C := qt\left(1 - \frac{\alpha}{2}, n_X + n_Y - 2\right) \qquad C = 2.002$$

NOTE: qt(1-α) is used here in the MathCad function for the t distribution & jw Table 2 p. 750

$n = 11$

Decision Rule: Reject H_0 if T > C

$\text{Decision} := \text{if}(t > C, 1, 0) \qquad \text{Decision} = 0$ **< 0 = Do not reject H_0**
1 = Reject H_0

Confidence intervals:

$$CI := \begin{bmatrix} \text{mean}(X) - \text{mean}(Y) + C\cdot\sqrt{S_p\cdot\left(\frac{1}{n_X} + \frac{1}{n_Y}\right)} \\ \text{mean}(X) - \text{mean}(Y) - C\cdot\sqrt{S_p\cdot\left(\frac{1}{n_X} + \frac{1}{n_Y}\right)} \end{bmatrix}$$

$\text{Difference} := \text{mean}(X) - \text{mean}(Y)$

$\text{Difference} = -2.965$

$$CI = \begin{pmatrix} 2.263 \\ -8.193 \end{pmatrix}$$

UNIVARIATE MEANS TESTS - t tests
TWO POPULATIONS WITH UNEQUAL VARIANCES
Behrens-Fisher Problem

Read in Data:

Lizard Data - Mass of individuals of two species C & S
Part of jw Table 6-7 p. 330

X := READPRN("\DATA\T6-7MassC.DAT") $n_X := rows(X)$ $p_X := rows(X)$

$n_X = 20$ $p_X = 20$

Y := READPRN("\DATA\T6-7MassS.DAT") $n_Y := rows(Y)$ $p_Y := rows(Y)$

$n_Y = 40$ $p_Y = 40$

X =

	1
1	7.513
2	5.032
3	5.867
4	11.088
5	2.419
6	13.61
7	18.247
8	16.832
9	15.91
10	17.035
11	16.526
12	4.53
13	7.23
14	5.2
15	13.45
16	14.08
17	14.665
18	6.092
19	5.264
20	16.902

Y =

	1
1	13.911
2	5.236
3	37.331
4	41.781
5	31.995
6	3.962
7	4.367
8	3.048
9	4.838
10	6.525
11	22.61
12	13.342
13	4.109
14	12.369
15	7.12
16	21.077
17	42.989
18	27.201
19	38.901
20	19.747
21	14.666
22	4.79
23	5.02
24	5.22
25	5.69
26	6.763
27	9.977
28	8.831
29	9.493
30	7.811
31	6.685
32	11.98
33	16.52
34	13.63
35	13.7
36	10.35
37	7.9
38	9.103
39	13.216
40	9.787

Summary statistics of the sample:

mean(X) = 10.875   mean(Y) = 13.84

Var(X) = 28.543   Var(Y) = 121.339

< Variances are definitely unequal so this approach applies!

Student t-test statistic:

$$t := \left| \frac{mean(X) - mean(Y)}{\sqrt{\frac{Var(X)}{n_X} + \frac{Var(Y)}{n_Y}}} \right|$$

t = 1.404

Revised degrees of freedom for the Behrens-Fisher Problem:

$$dd := \frac{\left(\frac{Var(X)}{n_X} + \frac{Var(Y)}{n_Y}\right)^2}{\frac{\left(\frac{var(X)}{n_X}\right)^2}{n_X} + \frac{\left(\frac{var(Y)}{n_Y}\right)^2}{n_Y}}$$

dd = 64.061

Hypothesis testing:

Assumptions: **X [$X_1, X_2, \ldots, X_n$] rs $N_1(\mu,\Sigma)$**
Y [$Y_1, Y_2, \ldots, Y_n$] rs $N_1(\mu,\Sigma)$
X's are independent of Y's

$H_0 : \mu_1 = \mu_2$ **< ie the difference is zero**

$H_1 : \mu_1 <> \mu_2$ **< two-tailed test!**

Stringency of the test: $\alpha := 0.05$ **< set as desired**

If assumptions hold and H_0 is true then:

$$C := qt\left(1 - \frac{\alpha}{2}, dd\right)$$

C = 1.998

NOTE: qt(1-α) is used here in the MathCad function for the t distribution & jw Table 2 p. 750

Decision Rule: Reject H_0 if T > C

n = 11

Decision := if (t > C, 1, 0)   Decision = 0

< 0 = Do not reject H_0
1 = Reject H_0

Confidence intervals:

$$CI := \begin{pmatrix} mean(X) - mean(Y) + C\cdot\sqrt{\frac{Var(X)}{n_X} + \frac{Var(Y)}{n_Y}} \\ mean(X) - mean(Y) - C\cdot\sqrt{\frac{Var(X)}{n_X} + \frac{Var(Y)}{n_Y}} \end{pmatrix}$$

Difference := mean(X) − mean(Y)

Difference = −2.965

$$CI = \begin{pmatrix} 1.254 \\ -7.184 \end{pmatrix}$$

ORIGIN ≡ 1

ANOVA - ANALYSIS OF VARIANCE - Univariate Case

Prepared by:
Wm Stein

Verifying calcuations using Example 6-6 in jw p297.

Three Populations:

$g := 3$ **<number of populations**

$P_1 := \begin{pmatrix} 9 \\ 6 \\ 9 \end{pmatrix}$ $n_1 := rows(P_1)$ $n_1 = 3$

$P_2 := \begin{pmatrix} 0 \\ 2 \end{pmatrix}$ $n_2 := rows(P_2)$

$P_3 := \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}$ $n_3 := rows(P_3)$ $P := stack(P_1, P_2, P_3)$ $P = \begin{pmatrix} 9 \\ 6 \\ 9 \\ 0 \\ 2 \\ 3 \\ 1 \\ 2 \end{pmatrix}$ $n := rows(P)$

Grand mean:

$i := 1 .. n$

$\mu_i := mean(P)$ $\mu = \begin{pmatrix} 4 \\ 4 \\ 4 \\ 4 \\ 4 \\ 4 \\ 4 \\ 4 \end{pmatrix}$

Sample parameters:

Population means:

$i := 1 .. n_1$

$X_{bar1} := mean(P_1)$ $X_{bar1} = 8$ $\tau 1_i := mean(P_1) - \mu_i$ $\tau 1 = \begin{pmatrix} 4 \\ 4 \\ 4 \end{pmatrix}$

$i := 1 .. n_2$

$X_{bar2} := mean(P_2)$ $X_{bar2} = 1$ $\tau 2_i := mean(P_2) - \mu_i$ $\tau 2 = \begin{pmatrix} -3 \\ -3 \end{pmatrix}$

$i := 1 .. n_3$

$X_{bar3} := mean(P_3)$ $X_{bar3} = 2$ $\tau 3_i := mean(P_3) - \mu_i$ $\tau 3 = \begin{pmatrix} -2 \\ -2 \\ -2 \end{pmatrix}$ $\tau := stack(\tau 1, \tau 2, \tau 3)$ $\tau = \begin{pmatrix} 4 \\ 4 \\ 4 \\ -3 \\ -3 \\ -2 \\ -2 \\ -2 \end{pmatrix}$

Residuals:

$\varepsilon := P - \mu - \tau$

$\varepsilon = \begin{pmatrix} 1 \\ -2 \\ 1 \\ -1 \\ 1 \\ 1 \\ -1 \\ 0 \end{pmatrix}$

2005 ANOVA 1

ANOVA table for comparing population means:

Source of variation	Sum of Squares	Degrees of Freedom		Mean Squares	
TREATMENTS (Between)	$\tau^T \cdot \tau = (78)$	$df_B := g - 1$	$df_B = 2$	$MS_B := \dfrac{\left\lvert \tau^T \cdot \tau \right\rvert}{df_B}$	$MS_B = 39$
RESIDUAL (Within)	$\varepsilon^T \cdot \varepsilon = (10)$	$df_W := n - g$	$df_W = 5$	$MS_W := \dfrac{\left\lvert \varepsilon^T \cdot \varepsilon \right\rvert}{df_W}$	$MS_W = 2$
TOTAL	$(P - \mu)^T \cdot (P - \mu) = (88)$	$df_T := n - 1$	$df_T = 7$	$MS_T := \dfrac{\left\lvert (P - \mu)^T \cdot (P - \mu) \right\rvert}{df_T}$	$MS_T = 12.5714$

Decomposition Model:

$X_{i,j} = \mu + \tau_i + \varepsilon_{i,j}$

Restriction:

$\Sigma n_i \tau_i = 0$

Assumptions:

All populations P_1-P_g rs

$\varepsilon_{i,j}$ **rs $N(0, \sigma^2)$** **< Normality robust if n >15 & n_i>4**

Population variances equal **< Equal variances robust if n_i >4 & all n_i are equal**

Hypothesis testing:

$H_0 : \tau_1 = \tau_2 = ... \tau_g = 0$

H_1 **: at least one τ <> 0**

ANOVA test statistic:

$$F := \frac{MS_B}{MS_W} \qquad F = 19.5$$

Stringency of the test: $\alpha := 0.01$ **< set as desired**

If assumptions hold and H_0 is true then:

$$C := qF(1 - \alpha, df_B, df_W) \qquad C = 13.2739$$

NOTE: qF(1-α) is substituted here for F(α) in the text.

Decision Rule: Reject H_0 if T > C

Decision := if(F > C, 1, 0) Decision = 1

< 0 = Do not reject H_0
1 = Reject H_0

ORIGIN ≡ 1

ANOVA - ANALYSIS OF VARIANCE - Univariate Case

Prepared by:
Wm Stein

Lizard Mass & SVL Original data transformed to natural logs jw Table 6.7 p. 330

Reading the Data:

$X_1 := \text{READPRN}(\text{"\DATA\T6-7Clogs.txt"})$ $\quad p := \text{cols}(X_1)$

$X_2 := \text{READPRN}(\text{"\DATA\T6-7Slogs.txt"})$ $\quad g := 2$

$X_1 =$

	1	2
1	2.0166	4.3041
2	1.6158	4.2413
3	1.7693	4.2767
4	2.4059	4.382
5	0.8834	4.0254
6	2.6108	4.5433
7	2.904	4.5591
8	2.8233	4.6002
9	2.7669	4.5747
10	2.8353	4.5053
11	2.8049	4.5109
12	1.5107	4.2047
13	1.9782	4.3175
14	1.6487	4.2413
15	2.599	4.5163
16	2.6448	4.5109
17	2.6855	4.4998
18	1.807	4.2905
19	1.6609	4.2413
20	2.8274	4.5433

$X_2 =$

	1	2
1	2.6327	4.3438
2	1.6556	4.1271
3	3.6198	4.6821
4	3.7324	4.7449
5	3.4656	4.6634
6	1.3767	4.0254
7	1.4741	4.1026
8	1.1145	3.9512
9	1.5765	4.0943
10	1.8756	4.1589
11	3.1184	4.5643
12	2.5909	4.3758
13	1.4132	4.0164
14	2.5152	4.3175
15	1.9629	4.1667
16	3.0482	4.4716
17	3.7609	4.6913
18	3.3033	4.5643
19	3.661	4.7095
20	2.983	4.4368
21	2.6855	4.382
22	1.5665	4.1271
23	1.6134	4.119
24	1.6525	4.1271
25	1.7387	4.1589
26	1.9115	4.1431
27	2.3003	4.2627
28	2.1783	4.2413
29	2.2506	4.2121
30	2.0555	4.1897
31	1.8999	4.1667
32	2.4832	4.3694
33	2.8046	4.4308
34	2.6123	4.3944
35	2.6174	4.4128
36	2.337	4.3041
37	2.0669	4.2268
38	2.2086	4.2485
39	2.5814	4.3503
40	2.2811	4.2485

Summary statistics of the sample:

$n_1 := \text{rows}(X_1) \quad n_2 := \text{rows}(X_2) \quad n = \begin{pmatrix} 20 \\ 40 \end{pmatrix}$

Mean Vectors & Variance-Covariance matrices:

$i := 1..n_1 \quad ii := 1..n_2 \quad j := 1..p$

$l_{n_i} := 1 \quad I := \text{identity}(n_1)$

$ll_{n_{ii}} := 1 \quad II := \text{identity}(n_2)$

$$X_{bar_1} := \frac{1}{n_1}\cdot X_1^T\cdot l_n \qquad X_{bar_1} = \begin{pmatrix} 2.23992 \\ 4.39443 \end{pmatrix}$$

$$X_{bar_2} := \frac{1}{n_2}\cdot X_2^T\cdot ll_n \qquad X_{bar_2} = \begin{pmatrix} 2.36814 \\ 4.30809 \end{pmatrix}$$

$$d := X_{bar_1} - X_{bar_2} \qquad d = \begin{pmatrix} -0.1282 \\ 0.0863 \end{pmatrix}$$

$$S_1 := \frac{1}{n_1 - 1}\cdot X_1^T\cdot\left(I - \frac{1}{n_1}\cdot l_n\cdot l_n^T\right)\cdot X_1 \qquad S_1 = \begin{pmatrix} 0.353 & 0.0942 \\ 0.0942 & 0.026 \end{pmatrix}$$

$$S_2 := \frac{1}{n_2 - 1}\cdot X_2^T\cdot\left(II - \frac{1}{n_2}\cdot ll_n\cdot ll_n^T\right)\cdot X_2 \qquad S_2 = \begin{pmatrix} 0.5068 & 0.1454 \\ 0.1454 & 0.0426 \end{pmatrix}$$

Total Sample size:

$N := n_1 + n_2 \qquad N = 60$

Grand Mean:

$m := 1..g$

$$X_{barGM} := \frac{1}{N}\cdot\left[\sum_m \left(n_m\cdot X_{bar_m}\right)\right] \qquad X_{barGM} = \begin{pmatrix} 2.3254 \\ 4.3369 \end{pmatrix}$$

Calculation of Univariate Variances:

Residual/Error Sums of Squares SSE

$m := 1..g$

$$W := \sum_m \left(n_m - 1\right)\cdot S_m \qquad SSE_m := W_{m,m} \qquad SSE = \begin{pmatrix} 26.4749 \\ 2.1527 \end{pmatrix}$$

Treatment Sums of Squares SST

$m := 1..g$

$$B := \sum_m n_m\cdot\left(X_{bar_m} - X_{barGM}\right)\cdot\left(X_{bar_m} - X_{barGM}\right)^T \qquad SST_m := B_{m,m} \qquad SST = \begin{pmatrix} 0.2192 \\ 0.0994 \end{pmatrix}$$

2005 ANOVA 2

ANOVA table for comparing population means:

$m := 1 .. g$

Source of variation	Sum of Squares	Degrees of Freedom		Mean Squares	
TREATMENTS (Between)	$SST = \begin{pmatrix} 0.2192 \\ 0.0994 \end{pmatrix}$	$df_{B_m} := g - 1$	$df_B = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$	$MS_{B_m} := \frac{SST_m}{df_{B_m}}$	$MS_B = \begin{pmatrix} 0.2192 \\ 0.0994 \end{pmatrix}$
RESIDUAL (Within)	$SSE = \begin{pmatrix} 26.4749 \\ 2.1527 \end{pmatrix}$	$df_{W_m} := N - g$	$df_W = \begin{pmatrix} 58 \\ 58 \end{pmatrix}$	$MS_{W_m} := \frac{SSE_m}{df_{W_m}}$	$MS_W = \begin{pmatrix} 0.4565 \\ 0.0371 \end{pmatrix}$
TOTAL	$SST + SSE = \begin{pmatrix} 26.6941 \\ 2.2521 \end{pmatrix}$	$df_{T_m} := N - 1$	$df_T = \begin{pmatrix} 59 \\ 59 \end{pmatrix}$	$MS_{T_m} := \frac{SST_m}{df_{T_m}}$	$MS_T = \begin{pmatrix} 0.0037 \\ 0.0017 \end{pmatrix}$

Decomposition Model:

$$X_{i,j} = \mu + \tau_i + \varepsilon_{i,j}$$

Restriction:

$$\Sigma n_i \tau_i = 0$$

Hypothesis testing:

$H_0 : \tau_1 = \tau_2 = ... \tau_g = 0$

H_1 **: at least one** τ **<> 0**

Assumptions:

All populations P_1**-**P_g **rs**

$\varepsilon_{i,j}$ **rs** $N(0, \sigma^2)$ **< Normality robust if n >15 &** n_i**>4**

Population variances equal **< Equal variances robust if** n_i **>4 & all** n_i **are equal**

ANOVA test statistic:

$m := 1 .. g$

$$F_m := \frac{MS_{B_m}}{MS_{W_m}} \qquad F = \begin{pmatrix} 0.4802 \\ 2.6777 \end{pmatrix}$$

Probability:

$$Prob_m := 1 - pF\left(F_m, df_{B_m}, df_{W_m}\right) \qquad Prob = \begin{pmatrix} 0.4911 \\ 0.1072 \end{pmatrix}$$

Stringency of the test: $\alpha := 0.05$ **< set as desired**

If assumptions hold and H_0 **is true then:**

$$C_m := qF\left(1 - \alpha, df_{B_m}, df_{W_m}\right) \qquad C = \begin{pmatrix} 4.0069 \\ 4.0069 \end{pmatrix}$$

NOTE: qF(1-α**) is substituted here for F(**α**) in the text.**

Decision Rule: Reject H_0 **if T > C**

$$Decision_m := if\left(F_m > C_m, 1, 0\right) \qquad Decision = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

< 0 = Do not reject H_0
1 = Reject H_0

2005 Two Way ANOVA

$ORIGIN \equiv 1$

Two Way ANOVA fixed effects - Univariate Case

Prepared by:
Wm Stein

Castle Bakery Sales Data - Neter et al. 1996 *Applied Linear Statistical Models* 4th ed. p. 817ff

Castle Bakery company supplies wrapped Italian bread to supermarkets. An experimental study was made of the effects of:

Factor A: height of shelf display (bottom, middle, top)
Factor B: width of display shelf (regular, wide)

12 supermarkets were chosen with similar sales volume & clientele. Six treatments (3A vs 2B) were assigned to the stores (2 replicates for each treatment) at random.

Reading the Data:

$Y := READPRN(\text{"\textbackslash DATA\textbackslash Castle Bakery.txt"})$

1st col = Response variable
2nd col = Factor A levels
3rd col = Factor B levels

Y =

	1	2	3
1	47	1	1
2	43	1	1
3	62	2	1
4	68	2	1
5	41	3	1
6	39	3	1
7	46	1	2
8	40	1	2
9	67	2	2
10	71	2	2
11	42	3	2
12	46	3	2

Summary Statistics:

$i := 1..3 \quad j := 1..2 \quad k := 1..2 \qquad n := 2$
$a := 3$
$b := 2$

Block Means:

$X_{bar_{1,1}} := mean(Y_{1,1}, Y_{2,1}) \qquad X_{bar_{1,1}} = 45 \qquad Y_{bar_k} := X_{bar_{1,1}}$

$X_{bar_{2,1}} := mean(Y_{3,1}, Y_{4,1}) \qquad X_{bar_{2,1}} = 65 \qquad Y_{bar_{2+k}} := X_{bar_{2,1}}$

$X_{bar_{3,1}} := mean(Y_{5,1}, Y_{6,1}) \qquad X_{bar_{3,1}} = 40 \qquad Y_{bar_{4+k}} := X_{bar_{3,1}}$

$X_{bar_{1,2}} := mean(Y_{7,1}, Y_{8,1}) \qquad X_{bar_{1,2}} = 43 \qquad Y_{bar_{6+k}} := X_{bar_{1,2}}$

$X_{bar_{2,2}} := mean(Y_{9,1}, Y_{10,1}) \qquad X_{bar_{2,2}} = 69 \qquad Y_{bar_{8+k}} := X_{bar_{2,2}}$

$X_{bar_{3,2}} := mean(Y_{11,1}, Y_{12,1}) \qquad X_{bar_{3,2}} = 44 \qquad Y_{bar_{10+k}} := X_{bar_{3,2}}$

$$X_{bar} = \begin{pmatrix} 45 & 43 \\ 65 & 69 \\ 40 & 44 \end{pmatrix}$$

Ybar =

	1
1	45
2	45
3	65
4	65
5	40
6	40
7	43
8	43
9	69
10	69
11	44
12	44

Factor A Means:

$X_{barA_1} := mean(Y_{1,1}, Y_{2,1}, Y_{7,1}, Y_{8,1}) \qquad X_{barA_1} = 44$

$X_{barA_2} := mean(Y_{3,1}, Y_{4,1}, Y_{9,1}, Y_{10,1}) \qquad X_{barA_2} = 67$

$X_{barA_3} := mean(Y_{5,1}, Y_{6,1}, Y_{11,1}, Y_{12,1}) \qquad X_{barA_3} = 42$

$$X_{barA} = \begin{pmatrix} 44 \\ 67 \\ 42 \end{pmatrix}$$

Factor B Means:

$X_{barB_1} := mean(Y_{1,1}, Y_{2,1}, Y_{3,1}, Y_{4,1}, Y_{5,1}, Y_{6,1}) \qquad X_{barB_1} = 50$

$X_{barB_2} := mean(Y_{7,1}, Y_{8,1}, Y_{9,1}, Y_{10,1}, Y_{11,1}, Y_{12,1}) \qquad X_{barB_2} = 52$

$$X_{barB} = \begin{pmatrix} 50 \\ 52 \end{pmatrix}$$

Grand Mean:

$X_{GM} := mean(Y^{\langle 1 \rangle}) \qquad X_{GM} = 51$

2005 Two Way ANOVA

TWO-WAY ANOVA TABLE - fixed factors A & B:

Source of variation **Sum of Squares** **Degrees of Freedom** **Mean Squares**

FACTOR A

$SSA := n \cdot b \cdot \sum_i \left(X_{barA_i} - X_{GM}\right)^2$ $SSA = 1544$ $df_A := a - 1$ $df_A = 2$ $MS_A := \frac{SSA}{df_A}$ $MS_A = 772$

FACTOR B

$SSB := n \cdot a \cdot \sum_j \left(X_{barB_j} - X_{GM}\right)^2$ $SSB = 12$ $df_B := b - 1$ $df_B = 1$ $MS_B := \frac{SSB}{df_B}$ $MS_B = 12$

AB INTERACTION

$SSAB := n \cdot \sum_i \sum_j \left(X_{bar_{i,j}} - X_{barA_i} - X_{barB_j} + X_{GM}\right)^2$

$SSAB = 24$ $df_{AB} := (a-1)\cdot(b-1)$ $df_{AB} = 2$ $MS_{AB} := \frac{SSAB}{df_{AB}}$ $MS_{AB} = 12$

ERROR

$SSE := \left|\left(Y^{\langle 1\rangle} - Y_{bar}\right)^T \cdot \left(Y^{\langle 1\rangle} - Y_{bar}\right)\right|$

$SSE = 62$ $df_E := a \cdot b \cdot (n-1)$ $df_E = 6$ $MS_E := \frac{SSE}{df_E}$ $MS_E = 10.3333$

TOTAL

$SST := \left|\left(Y^{\langle 1\rangle} - X_{GM}\right)^T \cdot \left(Y^{\langle 1\rangle} - X_{GM}\right)\right|$

$SST = 1642$ $df_T := n \cdot a \cdot b - 1$ $df_T = 11$

Decomposition Model:

$$X_{i,j} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{i,j} + \varepsilon_{i,j}$$

Restriction:

$$\Sigma\ \alpha_i = \Sigma_i\ \beta_j = \Sigma_j\ \gamma_{i,j} = \Sigma\ \gamma_{i,j} = 0$$

Assumptions:

All populations P_1-P_g rs

$\varepsilon_{i,j}$ rs $N(0, \sigma^2)$

Population variances equal

Hypothesis testing:

Factor Interaction:

H_0 : all $\gamma_{i,j} = 0$ for all i & j

H_1 : at least one $\gamma_{i,j}$ <> 0

ANOVA test statistics:

Probability:

$F := \frac{MS_{AB}}{MS_E}$ $F = 1.1613$ $Prob := 1 - pF\left(F, df_{AB}, df_E\right)$ $Prob = 0.3747$

2005 Two Way ANOVA

Stringency of the test: $\alpha := 0.05$ **< set as desired**

If assumptions hold and H_0 is true then:

$C := qF(1-\alpha, df_{AB}, df_E)$ $C = 5.1433$ **NOTE: qF(1-α) is substituted here for F(α) in the text.**

Decision Rule: Reject H_0 if T > C

$Decision := if(F > C, 1, 0)$ $Decision = 0$ **< 0 = Do not reject H_0**
1 = Reject H_0

Factor A effect:

$H_0 : \alpha_1 = \alpha_2 = ... \alpha_i = 0$

H_1 : at least one α <> 0

ANOVA test statistics:

Probability:

$F := \frac{MS_A}{MS_E}$ $F = 74.7097$ $Prob := 1 - pF(F, df_A, df_E)$ $Prob = 5.7536 \times 10^{-5}$

Stringency of the test: $\alpha := 0.05$ **< set as desired**

If assumptions hold and H_0 is true then:

$C := qF(1-\alpha, df_A, df_E)$ $C = 5.1433$ **NOTE: qF(1-α) is substituted here for F(α) in the text.**

Decision Rule: Reject H_0 if T > C

$Decision := if(F > C, 1, 0)$ $Decision = 1$ **< 0 = Do not reject H_0**
1 = Reject H_0

Factor B effect:

$H_0 : \beta_1 = \beta_2 = ... \beta_j = 0$

H_1 : at least one β <> 0

ANOVA test statistics:

Probability:

$F := \frac{MS_B}{MS_E}$ $F = 1.1613$ $Prob := 1 - pF(F, df_B, df_E)$ $Prob = 0.3226$

Stringency of the test: $\alpha := 0.05$ **< set as desired**

If assumptions hold and H_0 is true then:

$C := qF(1-\alpha, df_B, df_E)$ $C = 5.9874$ **NOTE: qF(1-α) is substituted here for F(α) in the text.**

Decision Rule: Reject H_0 if T > C

$Decision := if(F > C, 1, 0)$ $Decision = 0$ **< 0 = Do not reject H_0**
1 = Reject H_0

2005 jw272

ORIGIN ≡ 1

HOTELLING'S T^2 - PAIRED DESIGN
jw 272.mcd

Prepared by:
Wm Stein

Read in Data:

Effluent Data jw Table 6-1 p. 275)
Samples taken and split for comparison of technique between:

X = variables BOD & SS measured in a Commercial Lab.
Y = variables BOD & SS measured in a State Lab.

X := READPRN("\DATA\T6-1CL.DAT")

Y := READPRN("\DATA\T6-1SL.DAT")

$n_x := \text{rows}(X)$ $\quad p_x := \text{cols}(X)$ $\quad n_x = 11$ $\quad p_x = 2$

$n_y := \text{rows}(Y)$ $\quad p_y := \text{cols}(Y)$ $\quad n_y = 11$ $\quad p_y = 2$

$n := n_x$ $\quad p := p_x$

X =

	1	2
1	6	27
2	6	23
3	18	64
4	8	44
5	11	30
6	34	75
7	28	26
8	71	124
9	43	54
10	33	30
11	20	14

Y =

	1	2
1	25	15
2	28	13
3	36	22
4	35	29
5	15	31
6	44	64
7	42	30
8	54	64
9	34	56
10	29	20
11	39	21

Construct Paired difference matrix & summary statistics:

$D := X - Y$

$i := 1..n \qquad j := 1..p$

$l_{n_i} := 1 \qquad I := \text{identity}(n)$

$X_{bar} := \frac{1}{n} \cdot X^T \cdot l_n \qquad X_{bar} = \begin{pmatrix} 25.27273 \\ 46.45455 \end{pmatrix}$

$Y_{bar} := \frac{1}{n} \cdot Y^T \cdot l_n \qquad Y_{bar} = \begin{pmatrix} 34.6364 \\ 33.1818 \end{pmatrix}$

$S_X := \frac{1}{n-1} \cdot X^T \cdot \left(I - \frac{1}{n} \cdot l_n \cdot l_n^T\right) \cdot X \qquad S_X = \begin{pmatrix} 387.4182 & 489.3636 \\ 489.3636 & 1014.0727 \end{pmatrix}$

$S_Y := \frac{1}{n-1} \cdot Y^T \cdot \left(I - \frac{1}{n} \cdot l_n \cdot l_n^T\right) \cdot Y \qquad S_Y = \begin{pmatrix} 109.2545 & 120.3727 \\ 120.3727 & 363.7636 \end{pmatrix}$

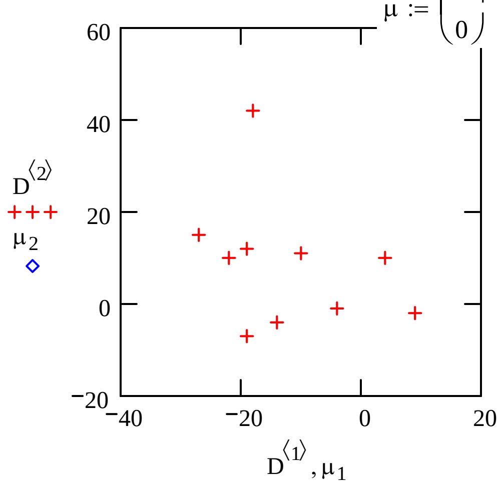
$D_{bar} := \frac{1}{n} \cdot D^T \cdot l_n \qquad D_{bar} = \begin{pmatrix} -9.3636 \\ 13.2727 \end{pmatrix}$

$S_D := \frac{1}{n-1} \cdot D^T \cdot \left(I - \frac{1}{n} \cdot l_n \cdot l_n^T\right) \cdot D \qquad S_D = \begin{pmatrix} 199.2545 & 88.3091 \\ 88.3091 & 418.6182 \end{pmatrix}$

D =

	1	2
1	-19	12
2	-22	10
3	-18	42
4	-27	15
5	-4	-1
6	-10	11
7	-14	-4
8	17	60
9	9	-2
10	4	10
11	-19	-7

$\mu := \begin{pmatrix} 0 \\ 0 \end{pmatrix}$



Hotelling's T^2 statistic (jw Eq. 5-6 p. 212):

$T_{sq} := n \cdot D_{bar}{}^T \cdot S_D{}^{-1} \cdot D_{bar} \qquad T_{sq} = (13.6393)$

Probability:

$K := \frac{(n-1) \cdot p}{(n-p)}$

$\text{Prob} := 1 - pF\left(\frac{|T_{sq}|}{K}, p, n-p\right)$

$\text{Prob} = 0.0208$

Hypothesis testing:

H_0 : μ. = 0

H_1 : μ <> 0

Assumption: D [D_1,D_2, ... , D_n] rs $N_p(\mu,\Sigma)$

Stringency of the test: $\alpha := 0.05$ **< set as desired**

If assumptions hold and H_0 is true then:

$$C := \frac{(n-1)\cdot p}{(n-p)}\cdot qF(1-\alpha,p,n-p) \qquad C = 9.4589$$

< jw eq. 5-6 p. 212

NOTE: qF(1-α) is substituted here for F(α) in the text.

Decision Rule: Reject H_0 if T > C

$$\text{Decision} := \text{if}\left(T_{sq_1} > C,1,0\right) \qquad \text{Decision} = 1$$

< 0 = Do not reject H_0
1 = Reject H_0

Note: jw p. 276 reports a slightly different value for C here, apparently due to rounding...

Liklihood Ratio test:

Maximum liklihood estimates of μ & Σ (jw Result 4.11 p. 171) and Σ_0:

$$\mu_{hat} := D_{bar} \qquad \Sigma_{Dhat} := \frac{(n-1)}{n}\cdot S_D \qquad \Sigma_{D0.hat} := \sum_i \left[\left(D^T\right)^{\langle i\rangle}\right]\cdot\left[\left(D^T\right)^{\langle i\rangle}\right]^T$$

$$\left|\Sigma_{D0.hat}\right| = 1.7874\times 10^7$$

$$\left|\Sigma_{Dhat}\right| = 6.249\times 10^4$$

Liklihood Ratio (Λ) & Wilks' lambda
(Based on analogy with ONE POPULATION situation- See jw Eq. 5-13 p. 217):

$$\Lambda := \left(\frac{\left|n\cdot\Sigma_{Dhat}\right|}{\left|\Sigma_{D0.hat}\right|}\right)^{\frac{n}{2}} \qquad \Lambda = 0.008811 \qquad \Lambda^{\frac{2}{n}} = 0.423024$$

< Λ = Wilks' lambda - value is 1.0 when $X_{bar} = \mu_0$ but decreases as μ_0 increases in distance from the sample mean.

$$\left[1+\frac{T_{sq}}{(n-1)}\right]^{-1} = (0.423024)$$

< Equivalent value in terms of Hotellings T^2 (jw Result 5.1 p. 218)

Converting to Hotellings T^2:

$$\left[\Lambda^{-\left(\frac{2}{n}\right)} - 1\right]\cdot(n-1) = 13.6393 \qquad T_{sq} = (13.6393)$$

< Solving for T^2 (same as T^2 above)
< Now run the above test...

Confidence intervals:

$\alpha := 0.05$ **< Set probability of Type 1 error**

The multivariate confidence ellipsoid (jw Eq. 5-18 p. 221):

$$\lambda := \text{reverse}\left(\text{sort}\left(\text{eigenvals}\left(S_D\right)\right)\right)$$

$$\varepsilon^{\langle j\rangle} := \text{eigenvec}\left(S_D,\lambda_j\right)$$

$$\lambda = \begin{pmatrix}449.75041\\168.12231\end{pmatrix} \qquad \varepsilon = \begin{pmatrix}0.33248 & -0.94311\\0.94311 & 0.33248\end{pmatrix}$$

< Coordinates of each column vector of ε gives the directions of confidence elipsoid

$$C := \sqrt{\frac{p\cdot(n-1)}{n\cdot(n-p)}\cdot qF(1-\alpha,p,n-p)} \qquad C = 0.9273$$

< C gives the boundary for the confidence ellipsoid for μ - see jw eq. 5-18 p. 221

$$i := 1..p$$

$$L_i := C\cdot\sqrt{\lambda_i}$$

NOTE: To obtain the values of F(α) reported in the text, qF(1-α) must be used in MathCad here.

2005 jw272

Multivariate simultaneous confidence ellipsoid:

$D_{bar} = \begin{pmatrix} -9.3636 \\ 13.2727 \end{pmatrix}$ **< Center of ellipsoid**

$L = \begin{pmatrix} 19.66569 \\ 12.02364 \end{pmatrix}$ **< L are half-lengths of the axes of the confidence ellipsoid for μ in the directions of ε**

Simultaneous T^2 confidence intervals:

$CI_{lower_i} := D_{bar_i} - C\cdot\sqrt{S_{D_{i,i}}}$ $\quad CI_{upper_i} := D_{bar_i} + C\cdot\sqrt{S_{D_{i,i}}}$ **< jw eq. 5-24 p. 225 (slightly modified)**

$CI := \text{augment}(CI_{lower}, CI_{upper})$

$D_{bar} = \begin{pmatrix} -9.36364 \\ 13.27273 \end{pmatrix}$ **< Mean values** $\quad CI = \begin{pmatrix} -22.45327 & 3.726 \\ -5.70012 & 32.24557 \end{pmatrix}$ **< T^2 confidence intervals**

Bonferroni simultaneous confidence intervals:

$c := qt\left(1 - \frac{\alpha}{2\cdot p}, n-1\right)$ $\quad c = 2.6338$ **< Critical value c based on t distribution**

NOTE: qt(1 - α/2p) is substituted for t(α/2p) in MathCad as above.

$ci_{lower_i} := D_{bar_i} - c\cdot\sqrt{\frac{S_{D_{i,i}}}{n}}$ $\quad ci_{upper_i} := D_{bar_i} + c\cdot\sqrt{\frac{S_{D_{i,i}}}{n}}$ **< jw eq. 5-29 p. 232**

$ci := \text{augment}(ci_{lower}, ci_{upper})$

$D_{bar} = \begin{pmatrix} -9.36364 \\ 13.27273 \end{pmatrix}$ **< Mean values** $\quad ci = \begin{pmatrix} -20.57311 & 1.84583 \\ -2.9749 & 29.52036 \end{pmatrix}$ **< Bonferroni confidence intervals**

2005 jw278

ORIGIN ≡ 1

HOTELLING'S T^2 PROFILE ANALYSIS FOR ONE POPULATION (=REPEATED MEASURES DESIGN WITH CONTRASTS)

jw 278.mcd

Prepared by:
Wm Stein

Read in Data:

Data on sleeping dogs jw Table 6.2 p. 281. Each of 19 dogs were successively given four treatments involving two levels each of two variables:

$X := \text{READPRN}(\text{"\\DATA\\T6-2.DAT"})$

$n := \text{rows}(X) \qquad q := \text{cols}(X)$

Treatment 1 = High CO_2, no H
Treatment 2 = Low CO_2, no H
Treatment 3 = High CO_2, with H
Treatment 4 = Low CO_2, with H

X =

	1	2	3	4
1	426	609	556	600
2	253	236	392	395
3	359	433	349	357
4	432	431	522	600
5	405	426	513	513
6	324	438	507	539
7	310	312	410	456
8	326	326	350	504
9	375	447	547	548
10	286	286	403	422
11	349	382	473	497
12	429	410	488	547
13	348	377	447	514
14	412	473	472	446
15	347	326	455	468
16	434	458	637	524
17	364	367	432	469
18	420	395	508	531
19	397	556	645	625

Summary statistics of the sample:

$i := 1..n \qquad j := 1..q$

$l_{n_i} := 1 \qquad I := \text{identity}(n)$

$$X_{bar} := \frac{1}{n} \cdot X^T \cdot l_n \qquad X_{bar} = \begin{pmatrix} 368.21053 \\ 404.63158 \\ 479.26316 \\ 502.89474 \end{pmatrix}$$

$$S := \frac{1}{n-1} \cdot X^T \cdot \left(I - \frac{1}{n} \cdot l_n \cdot l_n^T \right) \cdot X$$

$$S = \begin{pmatrix} 2819.28655 & 3568.4152 & 2943.49708 & 2295.35673 \\ 3568.4152 & 7963.1345 & 5303.99123 & 4065.45906 \\ 2943.49708 & 5303.99123 & 6851.31579 & 4499.64035 \\ 2295.35673 & 4065.45906 & 4499.64035 & 4878.9883 \end{pmatrix}$$

Specify Contrast Matrix C:

$$C := \begin{pmatrix} -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

Testing orthogonality of the rows in C:

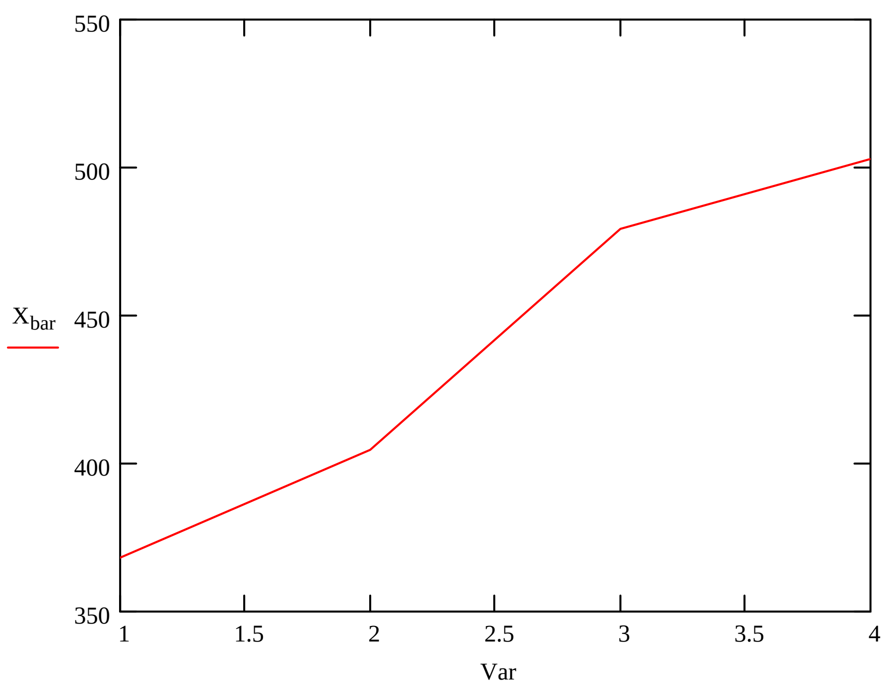
$CC := C^T$

$CC^{\langle 1 \rangle T} \cdot CC^{\langle 2 \rangle} = (0)$

$CC^{\langle 1 \rangle T} \cdot CC^{\langle 3 \rangle} = (0)$

$CC^{\langle 2 \rangle T} \cdot CC^{\langle 3 \rangle} = (0)$ **< products are all zero**

$$Var := \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$$



Linear combinations of mean vector X_{bar} and variance-covariance matrix S with C:

$$C \cdot X_{bar} = \begin{pmatrix} 209.316 \\ -60.053 \\ -12.789 \end{pmatrix} \qquad C \cdot S \cdot C^T = \begin{pmatrix} 9432.22807 & 1098.90643 & 927.59649 \\ 1098.90643 & 5195.83041 & 914.56725 \\ 927.59649 & 914.56725 & 7557.39766 \end{pmatrix}$$

Hotelling's T^2 statistic (jw Eq. 6.16 p. 279):

$$T_{sq} := n \cdot (C \cdot X_{bar})^T \cdot \left(C \cdot S \cdot C^T\right)^{-1} \cdot (C \cdot X_{bar}) \qquad T_{sq} = (116.016)$$

Hypothesis testing:

H_0 : $C\mu = 0$ — **Assumption: Vectors of observations in X [$X_1, X_2, \ldots, X_q$] rs $N_q(\mu,\Sigma)$**

H_1 : $C\mu <> 0$ — **However it is expected that observations within each vector (=subject) will be correlated**

Stringency of the test: $\alpha := 0.05$ **< set as desired**

If assumptions hold and H_0 is true then:

$$c := \frac{(n-1)\cdot(q-1)}{(n-q+1)} \cdot qF(1-\alpha, q-1, n-q+1) \qquad c = 10.931$$

< jw eq. 6.16 p. 279

NOTE: qF(1-α) is substituted here for F(α) in the text.

Decision Rule: Reject H_0 if T > c

$$Decision := if\left(T_{sq_1} > c, 1, 0\right) \qquad Decision = 1$$

< 0 = Do not reject H_0
1 = Reject H_0

Confidence T^2 intervals for the contrasts in C:

$\alpha := 0.05$ **< Set probability of Type 1 error**

$$c := \frac{(n-1)\cdot(q-1)}{(n-q+1)} \cdot qF(1-\alpha, q-1, n-q+1) \qquad c = 10.931$$

< critical value c

For first contrast (first row of C) involving H effect:

$$\phi := \left(C^T\right)^{\langle 1 \rangle}$$

$$CI_{lower} := \phi \cdot X_{bar} - \sqrt{c} \cdot \sqrt{\frac{\phi^T \cdot S \cdot \phi}{n}} \qquad CI_{upper} := \phi \cdot X_{bar} + \sqrt{c} \cdot \sqrt{\frac{\phi^T \cdot S \cdot \phi}{n}} \qquad \sqrt{c} \cdot \sqrt{\frac{\phi^T \cdot S \cdot \phi}{n}} = (73.665)$$

$$CI := augment(CI_{lower}, CI_{upper})$$

$$CI = (135.6503 \quad 282.98128)$$

< T^2 confidence interval for the contrast C_1 = H effect

For second contrast (second row of C) involving CO_2 effect:

$$\phi := \left(C^T\right)^{\langle 2 \rangle}$$

$$CI_{lower} := \phi \cdot X_{bar} - \sqrt{c} \cdot \sqrt{\frac{\phi^T \cdot S \cdot \phi}{n}} \qquad CI_{upper} := \phi \cdot X_{bar} + \sqrt{c} \cdot \sqrt{\frac{\phi^T \cdot S \cdot \phi}{n}} \qquad \sqrt{c} \cdot \sqrt{\frac{\phi^T \cdot S \cdot \phi}{n}} = (54.674)$$

$$CI := \mathrm{augment}\left(CI_{lower}, CI_{upper}\right)$$

$$CI = (-114.72708 \quad -5.37818)$$ **< T^2 confidence interval for the contrast C2 = CO_2 effect**

For third contrast (third row of C) involving H-CO_2 interaction:

$$\phi := \left(C^T\right)^{\langle 3 \rangle}$$

$$CI_{lower} := \phi \cdot X_{bar} - \sqrt{c} \cdot \sqrt{\frac{\phi^T \cdot S \cdot \phi}{n}} \qquad CI_{upper} := \phi \cdot X_{bar} + \sqrt{c} \cdot \sqrt{\frac{\phi^T \cdot S \cdot \phi}{n}} \qquad \sqrt{c} \cdot \sqrt{\frac{\phi^T \cdot S \cdot \phi}{n}} = (65.939)$$

$$CI := \mathrm{augment}\left(CI_{lower}, CI_{upper}\right)$$

$$CI = (-78.72858 \quad 53.14964)$$ **< T^2 confidence interval for the contrast C_3 = H-CO_2 interaction**

2005 jw283

HOTELLING'S T^2 - TWO POPULATIONS Small sample sizes, Equal Σs

jw 283.mcd

Prepared by:
Wm Stein

ORIGIN ≡ 1

Lizard Mass & SVL Original data transformed to natural logs jw Table 6.7 p. 330

Reading the Data:

$X_1 := \text{READPRN}(\text{"\\DATA\\T6-7Clogs.txt"})$ $\quad p := \text{cols}(X_1)$

$X_2 := \text{READPRN}(\text{"\\DATA\\T6-7Slogs.txt"})$

$X_1 =$

	1	2
1	2.0166	4.3041
2	1.6158	4.2413
3	1.7693	4.2767
4	2.4059	4.382
5	0.8834	4.0254
6	2.6108	4.5433
7	2.904	4.5591
8	2.8233	4.6002
9	2.7669	4.5747
10	2.8353	4.5053
11	2.8049	4.5109
12	1.5107	4.2047
13	1.9782	4.3175
14	1.6487	4.2413
15	2.599	4.5163
16	2.6448	4.5109
17	2.6855	4.4998
18	1.807	4.2905
19	1.6609	4.2413
20	2.8274	4.5433

$X_2 =$

	1	2
1	2.6327	4.3438
2	1.6556	4.1271
3	3.6198	4.6821
4	3.7324	4.7449
5	3.4656	4.6634
6	1.3767	4.0254
7	1.4741	4.1026
8	1.1145	3.9512
9	1.5765	4.0943
10	1.8756	4.1589
11	3.1184	4.5643
12	2.5909	4.3758
13	1.4132	4.0164
14	2.5152	4.3175
15	1.9629	4.1667
16	3.0482	4.4716
17	3.7609	4.6913
18	3.3033	4.5643
19	3.661	4.7095
20	2.983	4.4368
21	2.6855	4.382
22	1.5665	4.1271
23	1.6134	4.119
24	1.6525	4.1271
25	1.7387	4.1589
26	1.9115	4.1431
27	2.3003	4.2627
28	2.1783	4.2413
29	2.2506	4.2121
30	2.0555	4.1897
31	1.8999	4.1667
32	2.4832	4.3694
33	2.8046	4.4308
34	2.6123	4.3944
35	2.6174	4.4128
36	2.337	4.3041
37	2.0669	4.2268
38	2.2086	4.2485
39	2.5814	4.3503
40	2.2811	4.2485

Summary statistics of the sample:

$n_1 := \text{rows}(X_1) \quad n_2 := \text{rows}(X_2) \quad n_1 = 20 \quad n_2 = 40$

Mean Vectors & Variance-Covariance matrices:

$i := 1..n_1 \quad ii := 1..n_2 \quad j := 1..p$

$l_{n_i} := 1 \quad I := \text{identity}(n_1)$

$ll_{n_{ii}} := 1 \quad II := \text{identity}(n_2)$

$$X_{bar1} := \frac{1}{n_1}\cdot X_1^T\cdot l_n \qquad X_{bar1} = \begin{pmatrix} 2.23992 \\ 4.39443 \end{pmatrix}$$

$$X_{bar2} := \frac{1}{n_2}\cdot X_2^T\cdot ll_n \qquad X_{bar2} = \begin{pmatrix} 2.36814 \\ 4.30809 \end{pmatrix}$$

$$d := X_{bar1} - X_{bar2} \qquad d = \begin{pmatrix} -0.1282 \\ 0.0863 \end{pmatrix}$$

$$S_1 := \frac{1}{n_1 - 1}\cdot X_1^T\cdot\left(I - \frac{1}{n_1}\cdot l_n\cdot l_n^T\right)\cdot X_1 \qquad S_1 = \begin{pmatrix} 0.353 & 0.0942 \\ 0.0942 & 0.026 \end{pmatrix}$$

$$S_2 := \frac{1}{n_2 - 1}\cdot X_2^T\cdot\left(II - \frac{1}{n_2}\cdot ll_n\cdot ll_n^T\right)\cdot X_2 \qquad S_2 = \begin{pmatrix} 0.5068 & 0.1454 \\ 0.1454 & 0.0426 \end{pmatrix}$$

Calculating Pooled Variance-Covariance matrix (jw Eq 6-21):

$$S_{pooled} := \frac{(n_1 - 1)\cdot S_1 + (n_2 - 1)\cdot S_2}{(n_1 + n_2 - 2)} \qquad S_{pooled} = \begin{pmatrix} 0.46 & 0.13 \\ 0.13 & 0.04 \end{pmatrix}$$

Hotelling's T^2 statistic (jw Eq. 6-23): $\quad \delta_o := \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ **< Set different value for δ_o if a test off zero vector is desired.**

$$T_{sq} := (d - \delta_o)^T\cdot\left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\cdot S_{pooled}\right]^{-1}\cdot(d - \delta_o) \qquad T_{sq} = (228.1898)$$

Hypothesis testing:

$H_0 : \delta_o = \mu_1 - \mu_2 = 0$

$H_1 : \delta_o = \mu_1 - \mu_2 <> 0$

Probability:

$$K := \frac{(n_1 + n_2 - 2)\cdot p}{(n_1 + n_2 - p - 1)}$$

$$Prob := 1 - pF\left(\frac{|T_{sq}|}{K}, p, n_1 + n_2 - p - 1\right)$$

$Prob = 0$

Assumptions: **X_1 [$X_{1,1}$,$X_{1,2}$, ... , $X_{1,n}$] rs $N_p(\mu_1,\Sigma_1)$**
X_2 [$X_{2,1}$,$X_{2,2}$, ... , $X_{2,n}$] rs $N_p(\mu_2,\Sigma_2)$
X_1 independent of X_2
n_1 & n_2 small
$\Sigma_1 = \Sigma_2$

Stringency of the test: $\alpha := 0.05$ **< set as desired**

If assumptions hold and H_0 is true then (jw Result 6.2 & Eq. 6-4):

$$C := \frac{(n_1 + n_2 - 2)\cdot p}{(n_1 + n_2 - p - 1)}\cdot qF(1 - \alpha, p, n_1 + n_2 - p - 1) \qquad C = 6.4285$$

NOTE: qF(1-α) is substituted here for F(α) in the text.

Decision Rule: Reject H_0 if T > C

$$\text{Decision} := \text{if}\left(T_{sq_1} > C, 1, 0\right) \qquad \text{Decision} = 1$$

< 0 = Do not reject H_0
1 = Reject H_0

Coefficient vector for linear combination most responsible for rejection of H_o is proportional to:

$$S_{pooled}^{-1}\cdot d = \begin{pmatrix} -39.5738 \\ 139.4564 \end{pmatrix}$$

< see jw Remark p. 288

Confidence intervals:

$\alpha := 0.05$ **< Set probability of Type 1 error**

The multivariate confidence ellipsoid:

$$\lambda := \text{reverse}\left[\text{sort}\left[\text{eigenvals}\left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\cdot S_{pooled}\right]\right]\right]$$

$$\varepsilon^{\langle j\rangle} := \text{eigenvec}\left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\cdot S_{pooled}, \lambda_j\right]$$

$$\lambda = \begin{pmatrix} 0.03696 \\ 6.10088 \times 10^{-5} \end{pmatrix} \quad \varepsilon = \begin{pmatrix} 0.9624 & -0.27165 \\ 0.27165 & 0.9624 \end{pmatrix}$$

< Coordinates of each column vector of ε gives the directions of confidence ellipsoid

$$C := \sqrt{\frac{(n_1 + n_2 - 2)\cdot p}{(n_1 + n_2 - p - 1)}\cdot qF(1 - \alpha, p, n_1 + n_2 - p - 1)}$$

< C (squared) gives the boundary for the confidence ellipsoid for μ - see jw eq. 5-18 p. 221

$i := 1 .. p$ $\qquad C = 2.5355$

$L_i := C\cdot\sqrt{\lambda_i}$

NOTE: To obtain the values of F(α) reported in the text, qF(1-α) must be used in MathCad here.

Multivariate simultaneous confidence ellipsoid:

$$d = \begin{pmatrix} -0.1282 \\ 0.0863 \end{pmatrix}$$ **< Center of ellipsoid**

$$L = \begin{pmatrix} 0.48742 \\ 0.0198 \end{pmatrix}$$

< L are half-lengths of the axes of the confidence ellipsoid for μ in the directions of ε

2005 jw283

Simultaneous T^2 confidence intervals (jw Result 6.3 p. 287):

$$CI_{lower_i} := d_i - C\cdot\sqrt{\left[\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\cdot S_{pooled}\right]_{i,i}} \qquad CI_{upper_i} := d_i + C\cdot\sqrt{\left[\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\cdot S_{pooled}\right]_{i,i}}$$

$$CI := \text{augment}(CI_{lower}, CI_{upper})$$

$$d = \begin{pmatrix} -0.12822 \\ 0.08634 \end{pmatrix}$$ **< Mean values** $$CI = \begin{pmatrix} -0.59735 & 0.3409 \\ -0.04744 & 0.22011 \end{pmatrix}$$ **< T^2 confidence intervals**

Bonferroni simultaneous confidence intervals:
(extracted from the univariate case)

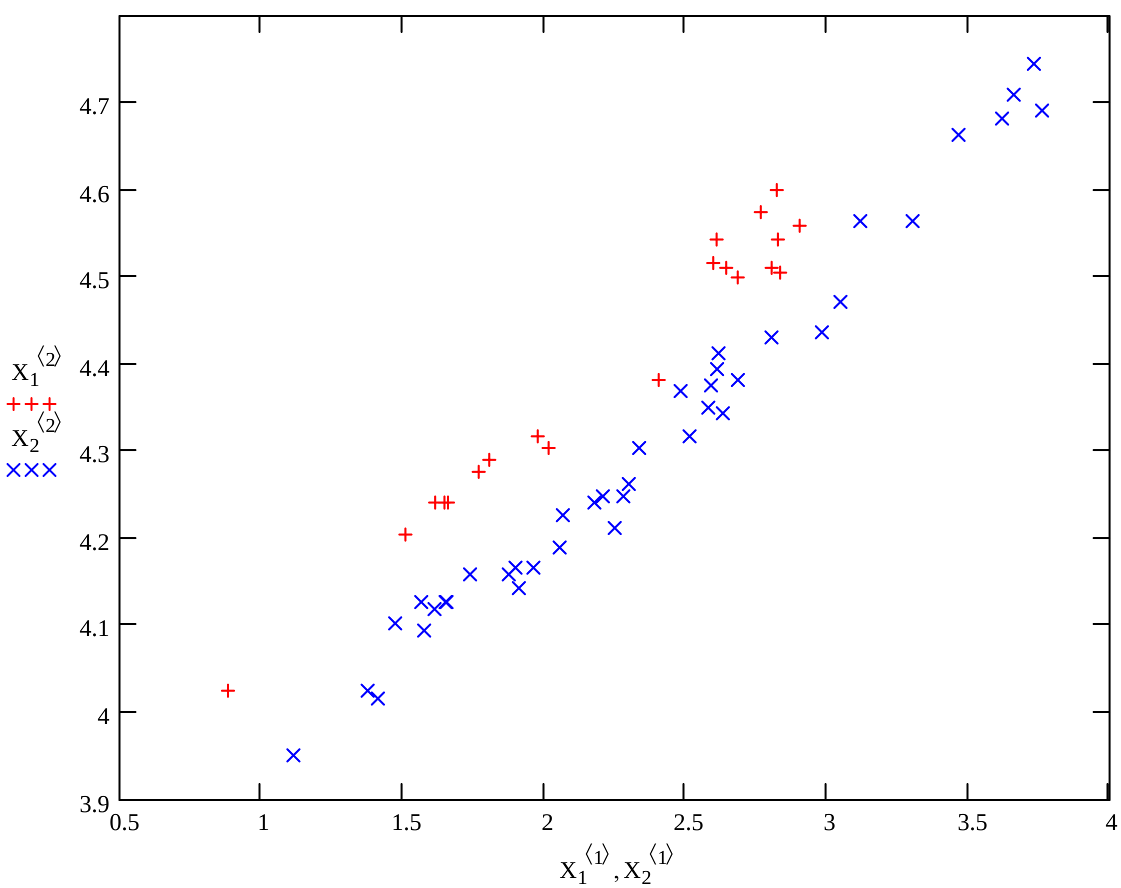
$$c := qt\left(1 - \frac{\alpha}{2\cdot p}, n_1 + n_2 - 2\right) \qquad c = 2.3011$$ **< Critical value c based on t distribution**

$$ci_{lower_i} := d_i - c\cdot\sqrt{\left[\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\cdot S_{pooled}\right]_{i,i}} \qquad ci_{upper_i} := d_i + c\cdot\sqrt{\left[\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\cdot S_{pooled}\right]_{i,i}}$$

NOTE: qt(1 - α/2p) is substituted for t(α/2p) in MathCad as above.

$$ci := \text{augment}(ci_{lower}, ci_{upper})$$

$$d = \begin{pmatrix} -0.12822 \\ 0.08634 \end{pmatrix}$$ **< Mean values** $$ci = \begin{pmatrix} -0.55398 & 0.29754 \\ -0.03507 & 0.20774 \end{pmatrix}$$ **< Bonferroni confidence intervals**



2005 jw283A

Prepared by:
Wm Stein

HOTELLING'S T^2 - TWO POPULATIONS Small sample sizes, Equal Σs
jw 283A.mcd

ORIGIN ≡ 1

Verifying calculations using Using Example 6-4 in jw. Only mean vectors and Variance-Covariance matrices are provided

Mean Vectors & Variance-Covariance matrices:

$$x_{bar1} := \begin{pmatrix} 204.4 \\ 556.6 \end{pmatrix} \qquad x_{bar2} := \begin{pmatrix} 130.0 \\ 355.0 \end{pmatrix} \qquad S_1 := \begin{pmatrix} 13825.3 & 23823.4 \\ 23823.4 & 73107.4 \end{pmatrix} \qquad n_1 := 45 \qquad p := 2$$

$$j := 1..p$$

Observed difference vector d:

$$d := x_{bar1} - x_{bar2} \qquad d = \begin{pmatrix} 74.4 \\ 201.6 \end{pmatrix} \qquad S_2 := \begin{pmatrix} 8632.0 & 19616.7 \\ 19616.7 & 55964.5 \end{pmatrix} \qquad n_2 := 55$$

Calculating Pooled Variance-Covariance matrix (jw Eq 6-21):

$$S_{pooled} := \frac{(n_1 - 1)\cdot S_1 + (n_2 - 1)\cdot S_2}{(n_1 + n_2 - 2)} \qquad S_{pooled} = \begin{pmatrix} 10963.69 & 21505.42 \\ 21505.42 & 63661.31 \end{pmatrix}$$

Hypothesis testing:

H_0 : $\delta_o = \mu_1 - \mu_2 = 0$

H_1 : $\delta_o = \mu_1 - \mu_2$ <> 0

$$\delta_o := \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

< Set different value for δ_o if a test off zero vector is desired.

Assumptions: X_1 $[X_{1,1}, X_{1,2}, \ldots, X_{1,n}]$ rs $N_p(\mu_1, \Sigma_1)$
X_2 $[X_{2,1}, X_{2,2}, \ldots, X_{2,n}]$ rs $N_p(\mu_2, \Sigma_2)$
X_1 independent of X_2
n_1 & n_2 small
$\Sigma_1 = \Sigma_2$

Hotelling's T^2 statistic (jw Eq. 6-23):

$$T_{sq} := (x_{bar1} - x_{bar2} - \delta_o)^T \cdot \left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\cdot S_{pooled}\right]^{-1} \cdot (x_{bar1} - x_{bar2} - \delta_o) \qquad T_{sq} = (16.0662)$$

Stringency of the test: $\alpha := 0.05$ **< set as desired**

If assumptions hold and H_0 is true then (jw Result 6.2 & Eq. 6-4):

$$C := \frac{(n_1 + n_2 - 2)\cdot p}{(n_1 + n_2 - p - 1)} \cdot qF(1 - \alpha, p, n_1 + n_2 - p - 1) \qquad C = 6.2441$$

NOTE: qF(1-α) is substituted here for F(α) in the text.

Decision Rule: Reject H_0 if T > C

$$Decision := if(T_{sq_1} > C, 1, 0) \qquad Decision = 1$$

< 0 = Do not reject H_0
1 = Reject H_0

Coefficient vector for linear combination most responsible for rejection of H_o is proportional to:

$$S_{pooled}^{-1} \cdot d = \begin{pmatrix} 0.0017 \\ 0.0026 \end{pmatrix}$$

< see jw Remark p. 288

Confidence intervals:

$\alpha := 0.05$ **< Set probability of Type 1 error**

The multivariate confidence ellipsoid:

$$\lambda := \text{reverse}\left[\text{sort}\left[\text{eigenvals}\left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\cdot S_{pooled}\right]\right]\right]$$

$$\varepsilon^{\langle j\rangle} := \text{eigenvec}\left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\cdot S_{pooled}, \lambda_j\right]$$

$$\lambda = \begin{pmatrix} 2881.75458 \\ 133.39685 \end{pmatrix} \qquad \varepsilon = \begin{pmatrix} 0.33562 & -0.942 \\ 0.942 & 0.33562 \end{pmatrix}$$

< Coordinates of each column vector of ε gives the directions of confidence elipsoid

$$C := \sqrt{\frac{(n_1 + n_2 - 2)\cdot p}{(n_1 + n_2 - p - 1)}\cdot qF(1 - \alpha, p, n_1 + n_2 - p - 1)}$$

< C (squared) gives the boundary for the confidence ellipsoid for μ - see jw eq. 5-18 p. 221

$i := 1 .. p$

$C = 2.4988$

$L_i := C\cdot\sqrt{\lambda_i}$

NOTE: To obtain the values of F(α) reported in the text, qF(1-α) must be used in MathCad here.

Multivariate simultaneous confidence ellipsoid:

$$d = \begin{pmatrix} 74.4 \\ 201.6 \end{pmatrix}$$ **< Center of ellipsoid**

$$L = \begin{pmatrix} 134.14146 \\ 28.86073 \end{pmatrix}$$ **< L are half-lengths of the axes of the confidence ellipsoid for μ in the directions of ε**

Simultaneous T^2 confidence intervals (jw Result 6.3 p. 287):

$$CI_{lower_i} := d_i - C\cdot\sqrt{\left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\cdot S_{pooled}\right]_{i,i}} \qquad CI_{upper_i} := d_i + C\cdot\sqrt{\left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\cdot S_{pooled}\right]_{i,i}}$$

$CI := \text{augment}(CI_{lower}, CI_{upper})$

$$d = \begin{pmatrix} 74.4 \\ 201.6 \end{pmatrix}$$ **< Mean values**

$$CI = \begin{pmatrix} 21.80733 & 126.99267 \\ 74.86846 & 328.33154 \end{pmatrix}$$ **< T^2 confidence intervals**

Bonferroni simultaneous confidence intervals:
(extracted from the univariate case)

$$c := qt\left(1 - \frac{\alpha}{2\cdot p}, n_1 + n_2 - 2\right) \qquad c = 2.2764$$

< Critical value c based on t distribution

$$ci_{lower_i} := d_i - c\cdot\sqrt{\left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\cdot S_{pooled}\right]_{i,i}} \qquad ci_{upper_i} := d_i + c\cdot\sqrt{\left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\cdot S_{pooled}\right]_{i,i}}$$

NOTE: qt(1 - α/2p) is substituted for t(α/2p) in MathCad as above.

$ci := \text{augment}(ci_{lower}, ci_{upper})$

$$d = \begin{pmatrix} 74.4 \\ 201.6 \end{pmatrix}$$ **< Mean values**

$$ci = \begin{pmatrix} 26.48936 & 122.31064 \\ 86.15065 & 317.04935 \end{pmatrix}$$ **< Bonferroni confidence intervals**

ORIGIN ≡ 1

HOTELLING'S T^2 - TWO POPULATIONS, UNEQUAL Σs for LARGE SAMPLES ONLY
jw 290.mcd

Prepared by:
Wm Stein

Lizard Mass & SVL Original data transformed to natural logs jw Table 6.7 p. 330

Reading the Data:

$X_1 := \text{READPRN}(\text{"\textbackslash DATA\textbackslash T6-7Clogs.txt"})$ $\qquad p := \text{cols}(X_1)$

$X_2 := \text{READPRN}(\text{"\textbackslash DATA\textbackslash T6-7Slogs.txt"})$

Summary statistics of the sample:

$n_1 := \text{rows}(X_1) \quad n_2 := \text{rows}(X_2) \quad n_1 = 20 \quad n_2 = 40$

Mean Vectors & Variance-Covariance matrices:

$i := 1..n_1 \quad ii := 1..n_2 \quad j := 1..p$

$l_{n_i} := 1 \quad I := \text{identity}(n_1)$

$ll_{n_{ii}} := 1 \quad II := \text{identity}(n_2)$

$X_{bar1} := \frac{1}{n_1} \cdot X_1^T \cdot l_n \qquad X_{bar1} = \begin{pmatrix} 2.23992 \\ 4.39443 \end{pmatrix}$

$X_{bar2} := \frac{1}{n_2} \cdot X_2^T \cdot ll_n \qquad X_{bar2} = \begin{pmatrix} 2.36814 \\ 4.30809 \end{pmatrix}$

$d := X_{bar1} - X_{bar2} \qquad d = \begin{pmatrix} -0.1282 \\ 0.0863 \end{pmatrix}$

$S_1 := \frac{1}{n_1 - 1} \cdot X_1^T \cdot \left(I - \frac{1}{n_1} \cdot l_n \cdot l_n^T\right) \cdot X_1 \qquad S_1 = \begin{pmatrix} 0.353 & 0.0942 \\ 0.0942 & 0.026 \end{pmatrix}$

$S_2 := \frac{1}{n_2 - 1} \cdot X_2^T \cdot \left(II - \frac{1}{n_2} \cdot ll_n \cdot ll_n^T\right) \cdot X_2 \qquad S_2 = \begin{pmatrix} 0.5068 & 0.1454 \\ 0.1454 & 0.0426 \end{pmatrix}$

$X_1 =$

	1	2
1	2.0166	4.3041
2	1.6158	4.2413
3	1.7693	4.2767
4	2.4059	4.382
5	0.8834	4.0254
6	2.6108	4.5433
7	2.904	4.5591
8	2.8233	4.6002
9	2.7669	4.5747
10	2.8353	4.5053
11	2.8049	4.5109
12	1.5107	4.2047
13	1.9782	4.3175
14	1.6487	4.2413
15	2.599	4.5163
16	2.6448	4.5109
17	2.6855	4.4998
18	1.807	4.2905
19	1.6609	4.2413
20	2.8274	4.5433

$X_2 =$

	1	2
1	2.6327	4.3438
2	1.6556	4.1271
3	3.6198	4.6821
4	3.7324	4.7449
5	3.4656	4.6634
6	1.3767	4.0254
7	1.4741	4.1026
8	1.1145	3.9512
9	1.5765	4.0943
10	1.8756	4.1589
11	3.1184	4.5643
12	2.5909	4.3758
13	1.4132	4.0164
14	2.5152	4.3175
15	1.9629	4.1667
16	3.0482	4.4716
17	3.7609	4.6913
18	3.3033	4.5643
19	3.661	4.7095
20	2.983	4.4368
21	2.6855	4.382
22	1.5665	4.1271
23	1.6134	4.119
24	1.6525	4.1271
25	1.7387	4.1589
26	1.9115	4.1431
27	2.3003	4.2627
28	2.1783	4.2413
29	2.2506	4.2121
30	2.0555	4.1897
31	1.8999	4.1667
32	2.4832	4.3694
33	2.8046	4.4308
34	2.6123	4.3944
35	2.6174	4.4128
36	2.337	4.3041
37	2.0669	4.2268
38	2.2086	4.2485
39	2.5814	4.3503
40	2.2811	4.2485

Calculating Combined Variance-Covariance matrix (jw Result 6.4 p. 291):

$S_{combined} := \left(\frac{1}{n_1} \cdot S_1 + \frac{1}{n_2} \cdot S_2\right) \qquad S_{combined} = \begin{pmatrix} 0.030324 & 0.008343 \\ 0.008343 & 0.002362 \end{pmatrix}$

$\delta_o := \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ **< Set different value for δ_o if a test off zero vector is desired.**

Hotelling's T^2 statistic (jw Result 6.4 p. 291):

$T_{sq} := (d - \delta_o)^T \cdot (S_{combined})^{-1} \cdot (d - \delta_o) \qquad T_{sq} = (224.796)$

Hypothesis testing:

$H_0 : \delta_o = \mu_1 - \mu_2 = 0$

$H_1 : \delta_o = \mu_1 - \mu_2 <> 0$

Probability:

$Prob := 1 - \text{pchisq}(|T_{sq}|, p)$

$Prob = 0$

Assumptions: $X_1\ [X_{1,1}, X_{1,2}, \ldots, X_{1,n}]$ rs $N_p(\mu_1, \Sigma_1)$
$X_2\ [X_{2,1}, X_{2,2}, \ldots, X_{2,n}]$ rs $N_p(\mu_2, \Sigma_2)$
X_1 independent of X_2
n_1 & n_2 LARGE
$\Sigma_1 <> \Sigma_2$

Stringency of the test: $\alpha := 0.05$ **< set as desired**

If assumptions hold and H_0 is true then (jw Result 6.4):

$C := \text{qchisq}(1 - \alpha, p)$ $C = 5.9915$

NOTE: qchisq(1-α) is substituted here for F(α) in the text.

Decision Rule: Reject H_0 if T > C

$\text{Decision} := \text{if}\left(T_{sq_1} > C, 1, 0\right)$ $\text{Decision} = 1$

< 0 = Do not reject H_0
1 = Reject H_0

Coefficient vector for linear combination most responsible for rejection of H_o is proportional to:

$$S_{combined}^{-1} \cdot d = \begin{pmatrix} -511.5796 \\ 1843.9771 \end{pmatrix}$$

< see jw Remark p. 288 & checks with calculations p. 292

Confidence intervals:

$\alpha := 0.05$ **< Set probability of Type 1 error**

The multivariate confidence ellipsoid:

$$\lambda := \text{reverse}(\text{sort}(\text{eigenvals}(S_{combined})))$$

$$\varepsilon^{\langle j \rangle} := \text{eigenvec}(S_{combined}, \lambda_j)$$

$$\lambda = \begin{pmatrix} 0.03262 \\ 6.13021 \times 10^{-5} \end{pmatrix} \quad \varepsilon = \begin{pmatrix} 0.96403 & -0.26578 \\ 0.26578 & 0.96403 \end{pmatrix}$$

< Coordinates of each column vector of ε gives the directions of confidence elipsoid

$C := \sqrt{\text{qchisq}(1 - \alpha, p)}$ $C = 2.4477$

< C (squared) gives the boundary for the confidence ellipsoid for μ - see jw eq. 5-18 p. 221

$i := 1 .. p$

$L_i := C \cdot \sqrt{\lambda_i}$

NOTE: To obtain the values of chisq(α) reported in the text, qchisq(1-α) must be used in MathCad here.

Multivariate simultaneous confidence ellipsoid:

$d = \begin{pmatrix} -0.1282 \\ 0.0863 \end{pmatrix}$ **< Center of ellipsoid**

$L = \begin{pmatrix} 0.44211 \\ 0.01916 \end{pmatrix}$ **< L are half-lengths of the axes of the confidence ellipsoid for μ in the directions of ε**

2005 jw290

Simultaneous T^2 confidence intervals (jw Result 6.4 p. 291):

$CIlower_i := d_i - C\cdot\sqrt{(S_{combined})_{i,i}}$ $\quad$ $CIupper_i := d_i + C\cdot\sqrt{(S_{combined})_{i,i}}$

$CI := \text{augment}(CI_{lower}, CI_{upper})$

$d = \begin{pmatrix} -0.12822 \\ 0.08634 \end{pmatrix}$ **< Mean values** $CI = \begin{pmatrix} -0.55446 & 0.29802 \\ -0.03261 & 0.20528 \end{pmatrix}$ **< T^2 confidence intervals**

Bonferroni simultaneous confidence intervals:
(extracted from the univariate case)

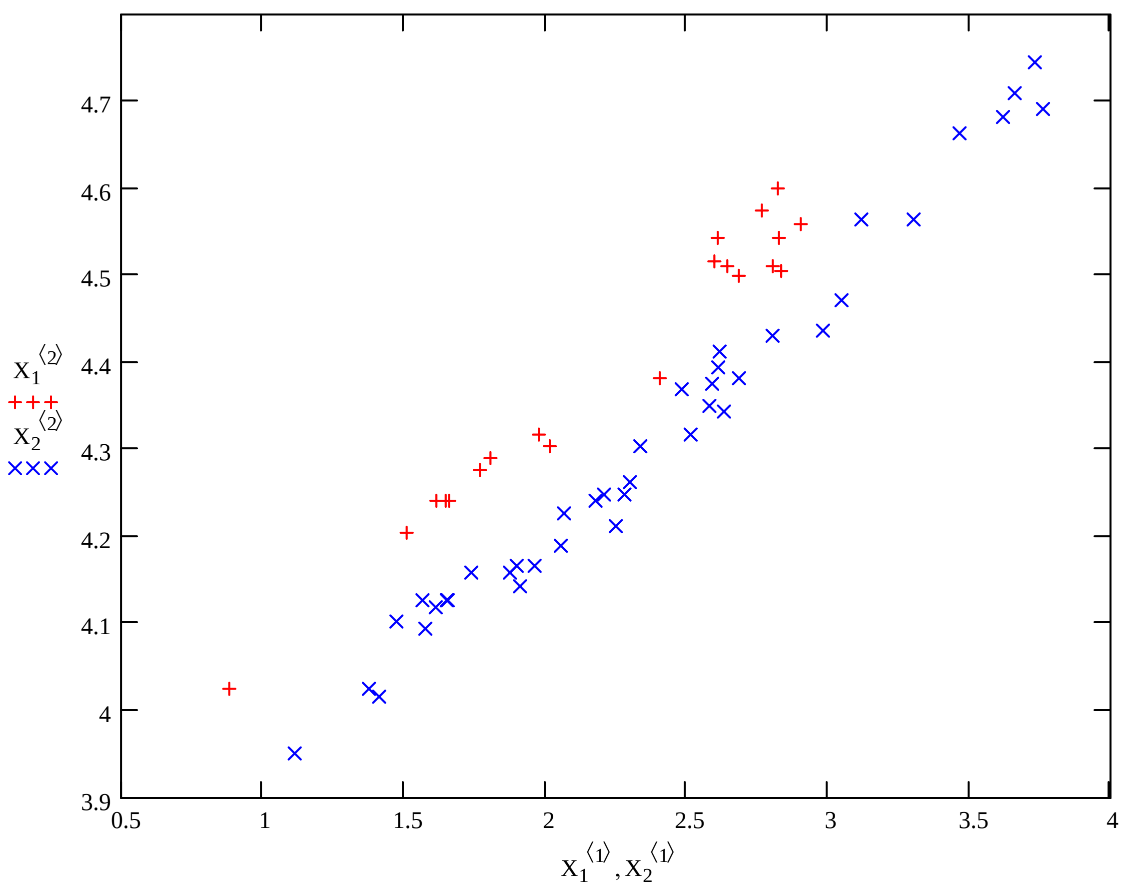
$c := qt\left(1 - \frac{\alpha}{2\cdot p}, n_1 + n_2 - 2\right)$ $\quad c = 2.3011$ **< Critical value c based on t distribution**

$ci_{lower_i} := d_i - c\cdot\sqrt{(S_{combined})_{i,i}}$ $\quad$ $ci_{upper_i} := d_i + c\cdot\sqrt{(S_{combined})_{i,i}}$

$ci := \text{augment}(ci_{lower}, ci_{upper})$

NOTE: qt(1 - α/2p) is substituted for t(α/2p) in MathCad as above.

$d = \begin{pmatrix} -0.12822 \\ 0.08634 \end{pmatrix}$ **< Mean values** $ci = \begin{pmatrix} -0.52892 & 0.27248 \\ -0.02549 & 0.19816 \end{pmatrix}$ **< Bonferroni confidence intervals**



2005 jw290A

ORIGIN ≡ 1 **HOTELLING'S T^2 - TWO POPULATIONS, UNEQUAL Σs for LARGE SAMPLES ONLY**

jw 290A.mcd

Prepared by: Wm Stein

Verifying calcuations using Example 6-4 in jw. Only mean vectors and Variance-Covariance matrices are provided

Mean Vectors & Variance-Covariance matrices:

$$x_{bar1} := \begin{pmatrix} 204.4 \\ 556.6 \end{pmatrix} \quad x_{bar2} := \begin{pmatrix} 130.0 \\ 355.0 \end{pmatrix} \quad S_1 := \begin{pmatrix} 13825.3 & 23823.4 \\ 23823.4 & 73107.4 \end{pmatrix} \quad n_1 := 45 \quad p := 2$$

$$j := 1 .. p$$

Observed difference vector d:

$$d := x_{bar1} - x_{bar2} \quad d = \begin{pmatrix} 74.4 \\ 201.6 \end{pmatrix} \quad S_2 := \begin{pmatrix} 8632.0 & 19616.7 \\ 19616.7 & 55964.5 \end{pmatrix} \quad n_2 := 55$$

Calculating Combined Variance-Covariance matrix (jw Result 6.4 p. 291):

$$S_{combined} := \left(\frac{1}{n_1} \cdot S_1 + \frac{1}{n_2} \cdot S_2 \right) \qquad S_{combined} = \begin{pmatrix} 464.17 & 886.08 \\ 886.08 & 2642.15 \end{pmatrix}$$

Hypothesis testing:

$H_0 : \delta_o = \mu_1 - \mu_2 = 0$

$H_1 : \delta_o = \mu_1 - \mu_2 <> 0$

$$\delta_o := \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

< Set different value for δ_o if a test off zero vector is desired.

Assumptions: X_1 $[X_{1,1}, X_{1,2}, ... , X_{1,n}]$ rs $N_p(\mu_1, \Sigma_1)$

X_2 $[X_{2,1}, X_{2,2}, ... , X_{2,n}]$ rs $N_p(\mu_2, \Sigma_2)$

X_1 independent of X_2

n_1 & n_2 LARGE

$\Sigma_1 <> \Sigma_2$

Hotelling's T^2 statistic (jw Result 6.4 p. 291):

$$T_{sq} := (x_{bar1} - x_{bar2} - \delta_o)^T \cdot (S_{combined})^{-1} \cdot (x_{bar1} - x_{bar2} - \delta_o) \qquad T_{sq} = (15.6585)$$

Stringency of the test: $\alpha := 0.05$ **< set as desired**

If assumptions hold and H_0 is true then (jw Result 6.4):

$$C := \text{qchisq}(1 - \alpha, p) \qquad C = 5.9915$$

NOTE: qchisq(1-α) is substituted here for F(α) in the text.

Decision Rule: Reject H_0 if T > C

$$\text{Decision} := \text{if}\left(T_{sq_1} > C, 1, 0\right) \qquad \text{Decision} = 1$$

< 0 = Do not reject H_0
1 = Reject H_0

Coefficient vector for linear combination most responsible for rejection of H_o is proportional to:

$$S_{combined}^{-1} \cdot d = \begin{pmatrix} 0.0407 \\ 0.0627 \end{pmatrix}$$

< see jw Remark p. 288 & checks with calculations p. 292

Confidence intervals:

$\alpha := 0.05$ **< Set probability of Type 1 error**

The multivariate confidence ellipsoid:

$\lambda := \text{reverse}(\text{sort}(\text{eigenvals}(S_{combined})))$

$\varepsilon^{\langle j \rangle} := \text{eigenvec}(S_{combined}, \lambda_j)$

$\lambda = \begin{pmatrix} 2957.0901 \\ 149.2295 \end{pmatrix}$ $\varepsilon = \begin{pmatrix} 0.33491 & -0.94225 \\ 0.94225 & 0.33491 \end{pmatrix}$

< Coordinates of each column vector of ε gives the directions of confidence elipsoid

$C := \sqrt{\text{qchisq}(1 - \alpha, p)}$ $C = 2.4477$

< C (squared) gives the boundary for the confidence ellipsoid for μ - see jw eq. 5-18 p. 221

NOTE: To obtain the values of chisq(α) reported in the text, qchisq(1-α) must be used in MathCad here.

$i := 1 .. p$

$L_i := C \cdot \sqrt{\lambda_i}$

Multivariate simultaneous confidence ellipsoid:

$d = \begin{pmatrix} 74.4 \\ 201.6 \end{pmatrix}$ **< Center of ellipsoid**

$L = \begin{pmatrix} 133.10635 \\ 29.90156 \end{pmatrix}$ **< L are half-lengths of the axes of the confidence ellipsoid for μ in the directions of ε**

Simultaneous T² confidence intervals (jw Result 6.4 p. 291):

$CI_{lower_i} := d_i - C \cdot \sqrt{(S_{combined})_{i,i}}$ $CI_{upper_i} := d_i + C \cdot \sqrt{(S_{combined})_{i,i}}$

$CI := \text{augment}(CI_{lower}, CI_{upper})$

$d = \begin{pmatrix} 74.4 \\ 201.6 \end{pmatrix}$ **< Mean values** $CI = \begin{pmatrix} 21.66401 & 127.13599 \\ 75.7814 & 327.4186 \end{pmatrix}$ **< T² confidence intervals**

Bonferroni simultaneous confidence intervals:
(extracted from the univariate case)

$c := qt\left(1 - \frac{\alpha}{2 \cdot p}, n_1 + n_2 - 2\right)$ $c = 2.2764$ **< Critical value c based on t distribution**

$ci_{lower_i} := d_i - c \cdot \sqrt{(S_{combined})_{i,i}}$ $ci_{upper_i} := d_i + c \cdot \sqrt{(S_{combined})_{i,i}}$

$ci := \text{augment}(ci_{lower}, ci_{upper})$

NOTE: qt(1 - α/2p) is substituted for t(α/2p) in MathCad as above.

$d = \begin{pmatrix} 74.4 \\ 201.6 \end{pmatrix}$ **< Mean values** $ci = \begin{pmatrix} 25.35645 & 123.44355 \\ 84.5909 & 318.6091 \end{pmatrix}$ **< Bonferroni confidence intervals**

2005 jw290A

Read in Data:

$X := \text{READPRN}(\text{"\\DATA\\T6-1CL.DAT"})$

$Y := \text{READPRN}(\text{"\\DATA\\T6-1SL.DAT"})$

$n_x := \text{rows}(X)$ $\quad p_x := \text{cols}(X)$ $\quad n_x = 11$ $\quad p_x = 2$

$n_y := \text{rows}(Y)$ $\quad p_y := \text{cols}(Y)$ $\quad n_y = 11$ $\quad p_y = 2$

$n := n_x$ $\quad p := p_x$

X =

	1	2
1	6	27
2	6	23
3	18	64
4	8	44
5	11	30
6	34	75
7	28	26
8	71	124
9	43	54
10	33	30
11	20	14

Y =

	1	2
1	25	15
2	28	13
3	36	22
4	35	29
5	15	31
6	44	64
7	42	30
8	54	64
9	34	56
10	29	20
11	39	21

Construct Paired difference matrix & summary statistics:

$D := X - Y$

$i := 1..n$ $\quad j := 1..p$

$l_{n_i} := 1$ $\quad I := \text{identity}(n)$

$X_{bar} := \frac{1}{n} \cdot X^T \cdot l_n$ $\quad X_{bar} = \begin{pmatrix} 25.27273 \\ 46.45455 \end{pmatrix}$

$Y_{bar} := \frac{1}{n} \cdot Y^T \cdot l_n$ $\quad Y_{bar} = \begin{pmatrix} 34.6364 \\ 33.1818 \end{pmatrix}$

$S_X := \frac{1}{n-1} \cdot X^T \cdot \left(I - \frac{1}{n} \cdot l_n \cdot l_n^T\right) \cdot X$ $\quad S_X = \begin{pmatrix} 387.4182 & 489.3636 \\ 489.3636 & 1014.0727 \end{pmatrix}$

$S_Y := \frac{1}{n-1} \cdot Y^T \cdot \left(I - \frac{1}{n} \cdot l_n \cdot l_n^T\right) \cdot Y$ $\quad S_Y = \begin{pmatrix} 109.2545 & 120.3727 \\ 120.3727 & 363.7636 \end{pmatrix}$

$D_{bar} := \frac{1}{n} \cdot D^T \cdot l_n$ $\quad D_{bar} = \begin{pmatrix} -9.3636 \\ 13.2727 \end{pmatrix}$

$S_D := \frac{1}{n-1} \cdot D^T \cdot \left(I - \frac{1}{n} \cdot l_n \cdot l_n^T\right) \cdot D$ $\quad S_D = \begin{pmatrix} 199.2545 & 88.3091 \\ 88.3091 & 418.6182 \end{pmatrix}$

D =

	1	2
1	-19	12
2	-22	10
3	-18	42
4	-27	15
5	-4	-1
6	-10	11
7	-14	-4
8	17	60
9	9	-2
10	4	10
11	-19	-7

2005 jw298

MANOVA - One Way
jw 298.mcd

Prepared by:
Wm Stein

$ORIGIN \equiv 1$

Lizard Mass & SVL Original data transformed to natural logs jw Table 6.7 p. 330

Reading the Data:

$X_1 := READPRN("\backslash DATA\backslash T6\text{-}7Clogs.txt")$ $\quad p := cols(X_1)$

$X_2 := READPRN("\backslash DATA\backslash T6\text{-}7Slogs.txt")$ $\quad g := 2$

Summary statistics of the sample:

$n_1 := rows(X_1) \quad n_2 := rows(X_2) \quad n = \begin{pmatrix} 20 \\ 40 \end{pmatrix}$

Mean Vectors & Variance-Covariance matrices:

$i := 1..n_1 \quad ii := 1..n_2 \quad j := 1..p$

$l_{n_i} := 1 \quad I := identity(n_1)$

$ll_{n_{ii}} := 1 \quad II := identity(n_2)$

$X_{bar_1} := \frac{1}{n_1} \cdot X_1^T \cdot l_n \qquad X_{bar_1} = \begin{pmatrix} 2.23992 \\ 4.39443 \end{pmatrix}$

$X_{bar_2} := \frac{1}{n_2} \cdot X_2^T \cdot ll_n \qquad X_{bar_2} = \begin{pmatrix} 2.36814 \\ 4.30809 \end{pmatrix}$

$d := X_{bar_1} - X_{bar_2} \qquad d = \begin{pmatrix} -0.12822 \\ 0.08634 \end{pmatrix}$

$S_1 := \frac{1}{n_1 - 1} \cdot X_1^T \cdot \left(I - \frac{1}{n_1} \cdot l_n \cdot l_n^T \right) \cdot X_1 \qquad S_1 = \begin{pmatrix} 0.353 & 0.0942 \\ 0.0942 & 0.026 \end{pmatrix}$

$S_2 := \frac{1}{n_2 - 1} \cdot X_2^T \cdot \left(II - \frac{1}{n_2} \cdot ll_n \cdot ll_n^T \right) \cdot X_2 \qquad S_2 = \begin{pmatrix} 0.50684 & 0.14539 \\ 0.14539 & 0.04255 \end{pmatrix}$

$X_1 =$

	1	2
1	2.01663	4.30407
2	1.61582	4.24133
3	1.76934	4.27667
4	2.40586	4.38203
5	0.88335	4.02535
6	2.6108	4.54329
7	2.904	4.55913
8	2.82328	4.60016
9	2.76695	4.57471
10	2.83527	4.50535
11	2.80493	4.51086
12	1.51072	4.20469
13	1.97824	4.31749
14	1.64866	4.24133
15	2.59898	4.51634
16	2.64476	4.51086
17	2.68546	4.49981
18	1.80698	4.29046
19	1.66089	4.24133
20	2.82743	4.54329

=

	1	2
1	2.63268	4.34381
2	1.65556	4.12713
3	3.61982	4.68213
4	3.73244	4.74493
5	3.46558	4.66344
6	1.37675	4.02535
7	1.47408	4.10264
8	1.11449	3.95124
9	1.5765	4.09434
10	1.87564	4.15888
11	3.11839	4.56435
12	2.59092	4.37576
13	1.41318	4.01638
14	2.51519	4.31749
15	1.96291	4.16667
16	3.04818	4.47164
17	3.76094	4.69135
18	3.30325	4.56435
19	3.66102	4.70953
20	2.983	4.43675
21	2.68553	4.38203
22	1.56653	4.12713
23	1.61343	4.11904
24	1.6525	4.12713
25	1.73871	4.15888
26	1.91147	4.14313
27	2.30028	4.26268
28	2.17827	4.24133
29	2.25055	4.21213
30	2.05553	4.18965
31	1.89987	4.16667
32	2.48324	4.36945
33	2.80457	4.43082
34	2.61227	4.39445
35	2.6174	4.4128
36	2.33699	4.30407
37	2.06686	4.22683
38	2.2086	4.2485
39	2.58143	4.35028
40	2.28105	4.2485

Total Sample size:

$N := n_1 + n_2 \qquad N = 60$

Grand Mean:

$m := 1..g$

$X_{barGM} := \frac{1}{N} \cdot \left[\sum_m \left(n_m \cdot X_{bar_m} \right) \right] \qquad X_{barGM} = \begin{pmatrix} 2.3254 \\ 4.33687 \end{pmatrix}$

Residual/Error/Within Matrix W:

$m := 1..g$

$W := \sum_m \left(n_m - 1 \right) \cdot S_m \qquad W = \begin{pmatrix} 26.47486 & 7.45949 \\ 7.45949 & 2.1527 \end{pmatrix}$

Treatment Matrix B:

$m := 1..g$

$B := \sum_m n_m \cdot \left(X_{bar_m} - X_{barGM} \right) \cdot \left(X_{bar_m} - X_{barGM} \right)^T \qquad B = \begin{pmatrix} 0.21921 & -0.1476 \\ -0.1476 & 0.09938 \end{pmatrix}$

2005 jw298

MANOVA Table jw p. 299

Source	SSP Matrix	Degrees of Freedom	
Treatment Matrix B	$B = \begin{pmatrix} 0.21921 & -0.1476 \\ -0.1476 & 0.09938 \end{pmatrix}$	$df_B := g - 1$	$df_B = 1$
Residual/Error Matrix W	$W = \begin{pmatrix} 26.47486 & 7.45949 \\ 7.45949 & 2.1527 \end{pmatrix}$	$df_W := N - g$	$df_W = 58$
TOTAL (mean corrected)	$B + W = \begin{pmatrix} 26.69407 & 7.31189 \\ 7.31189 & 2.25208 \end{pmatrix}$	$df_T := N - 1$	$df_T = 59$

MANOVA tests jw p. 299-300.

Decomposition Model:

$X_{i,j} = \mu + \tau_i + \varepsilon_{i,j}$ **where:** $j = 1$ **to** n_m

$m = 1$ **to** g

Restriction:

$\Sigma n_m \tau_m = 0$

Hypothesis testing:

$H_0 : \tau_1 = \tau_2 = \ldots \tau_g = 0$

H_1 **: at least one** τ **<> 0**

Assumptions:

All populations P_1-P_g **rs**

$\varepsilon_{i,j}$ **rs** $N(0, \sigma^2)$

Population variances equal

Wilk's Lambda Test Statistic:

$\Lambda_s := \dfrac{|W|}{|W + B|}$ $\quad\quad \Lambda_s = 0.20266$

Lawley-Hotelling Trace:

$LHtr := tr\left(B \cdot W^{-1}\right)$ $\quad\quad LHtr = 3.93431$

Pillai trace:

$Ptr := tr\left[B \cdot (B + W)^{-1}\right]$ $\quad\quad Ptr = 0.79734$

Roy's largest root:

$Rrt := max\left[eigenvals\left[W \cdot (B + W)^{-1}\right]\right]$ $\quad\quad Rrt = 1$ $\quad\quad eigenvals\left[W \cdot (B + W)^{-1}\right] = \begin{pmatrix} 1 \\ 0.20266 \end{pmatrix}$

Stringency of the test: $\alpha := 0.01$ **< set as desired**

2005 jw298

If assumptions hold and H_0 is true then jw Table 6.3 & p. 300:

Number of Variables p: **Number of groups g:** **Sampling Distribution Critical Value C:**

p = 1 **g > 1**

$C := qF(1-\alpha, g-1, N-g)$ $C = 7.0931$

Decision Rule: Reject H_0 if K > C:

$K := \left(\frac{N-g}{g-1}\right)\cdot\left(\frac{1-\Lambda_s}{\Lambda_s}\right)$ $K = 228.18983$ **Probability:** $1 - pF(K, g-1, N-g) = 0$

Decision := if(K > C, 1, 0) Decision = 1 **< 0 = Do not reject H_0**
1 = Reject H_0

p = 2 **g > 1**

$C := qF[1-\alpha, 2\cdot(g-1), 2\cdot(N-g-1)]$ $C = 4.79631$

Decision Rule: Reject H_0 if K > C:

$K := \left(\frac{N-g-1}{g-1}\right)\cdot\left(\frac{1-\sqrt{\Lambda_s}}{\sqrt{\Lambda_s}}\right)$ $K = 69.61582$ **Probability:** $1 - pF[K, 2\cdot(g-1), 2\cdot(N-g-1)] = 0$

Decision := if(K > C, 1, 0) Decision = 1 **< 0 = Do not reject H_0**
1 = Reject H_0

p > 0 **g = 2**

$C := qF(1-\alpha, p, N-p-1)$ $C = 4.99811$

Decision Rule: Reject H_0 if K > C:

$K := \left(\frac{N-p-1}{p}\right)\cdot\left(\frac{1-\Lambda_s}{\Lambda_s}\right)$ $K = 112.12776$ **Probability:** $1 - pF(K, p, N-p-1) = 0$

Decision := if(K > C, 1, 0) Decision = 1 **< 0 = Do not reject H_0**
1 = Reject H_0

p > 0 **g = 3**

$C := qF[1-\alpha, 2\cdot p, 2\cdot(N-p-2)]$ $C = 3.49133$

Decision Rule: Reject H_0 if K > C:

$K := \left(\frac{N-p-2}{p}\right)\cdot\left(\frac{1-\sqrt{\Lambda_s}}{\sqrt{\Lambda_s}}\right)$ $K = 34.19724$ **Probability:** $1 - pF[K, 2\cdot p, 2\cdot(N-p-2)] = 0$

Decision := if(K > C, 1, 0) Decision = 1 **< 0 = Do not reject H_0**
1 = Reject H_0

Bartlett's test for Large N

$C := qchisq[1-\alpha, p\cdot(g-1)]$ $C = 9.21034$

Decision Rule: Reject H_0 if K > C:

$K := -\left(N-1-\frac{p+g}{2}\right)\cdot\ln(\Lambda_s)$ $K = 90.9841$ **Probability:** $1 - pchisq[K, p\cdot(g-1)] = 0$

Decision := if(K > C, 1, 0) Decision = 1 **< 0 = Do not reject H_0**
1 = Reject H_0

Simultaneous Bonferroni confidence intervals for treatment effects jw p. 305:

$\alpha := 0.05$ **< Set probability of Type 1 error**

$$c := qt\left[1 - \frac{\alpha}{p \cdot g \cdot (g-1)}, N - g\right] \qquad c = 2.30108$$

Difference in means between first & second populations:

$i := 1 .. p$

$$\kappa_i := \left|\sqrt{\frac{W_{i,i}}{N-g} \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}\right| \qquad \kappa = \begin{pmatrix} 0.18503 \\ 0.05276 \end{pmatrix}$$

$$ci_{lower} := X_{bar_1} - X_{bar_2} - c \cdot \kappa \qquad ci_{upper} := X_{bar_1} - X_{bar_2} + c \cdot \kappa$$

$$ci := augment(ci_{lower}, ci_{upper}) \qquad ci = \begin{pmatrix} -0.55398 & 0.29754 \\ -0.03507 & 0.20774 \end{pmatrix}$$

Note: According to jw Result 6.5, Bonferroni simultaneous confidence intervals apply to ALL variables (=components of mean vectors) simultaneously for ALL pairwise comparisons across groups.

ORIGIN ≡ 1

MANOVA - One Way
jw 298A.mcd

Prepared by:
Wm Stein

Verifying calculations for jw Example 6.9 p. 303.
Wisconsin Nursing Home Data:

$g := 3$ **< # Populations (kinds of homes)**

$p := 4$ **< # Variables (kinds of labor)**

X1 = cost of nursing labor
X2 = cost of dietary labor
X3 = cost of plant operation & maintainance labor
X4 = cost of housekeeping & laundry labor

Data is provided in summary form only:

Private (P1)

$n_1 := 271$

$$X_{bar_1} := \begin{pmatrix} 2.066 \\ .480 \\ .082 \\ .360 \end{pmatrix}$$

$$S_1 := \begin{pmatrix} .291 & -.001 & .002 & .010 \\ -.001 & .011 & .000 & .003 \\ .002 & .000 & .001 & .000 \\ .010 & .003 & .000 & .010 \end{pmatrix}$$

Nonprofit (P2)

$n_2 := 138$

$$X_{bar_2} := \begin{pmatrix} 2.167 \\ .596 \\ .124 \\ .418 \end{pmatrix}$$

$$S_2 := \begin{pmatrix} .561 & .011 & .001 & .037 \\ .011 & .025 & .004 & .007 \\ .001 & .004 & .005 & .002 \\ .037 & .007 & .002 & .019 \end{pmatrix}$$

Government (P3)

$n_3 := 107$

$$X_{bar_3} := \begin{pmatrix} 2.273 \\ .521 \\ .125 \\ .383 \end{pmatrix}$$

$$S_3 := \begin{pmatrix} .261 & .030 & .003 & .018 \\ .030 & .017 & .000 & .006 \\ .003 & .000 & .004 & .001 \\ .018 & .006 & .001 & .013 \end{pmatrix}$$

Total Sample size:

$N := n_1 + n_2 + n_3 \qquad N = 516$

Grand Mean:

$$X_{barGM} := \frac{n_1 \cdot X_{bar_1} + n_2 \cdot X_{bar_2} + n_3 \cdot X_{bar_3}}{n_1 + n_2 + n_3} \qquad X_{barGM} = \begin{pmatrix} 2.1359 \\ 0.5195 \\ 0.1021 \\ 0.3803 \end{pmatrix}$$

Pooled Within (= Residual = Error) Matrix W:

$$W := (n_1 - 1) \cdot S_1 + (n_2 - 1) \cdot S_2 + (n_3 - 1) \cdot S_3 \qquad W = \begin{pmatrix} 183.093 & 4.417 & 0.995 & 9.677 \\ 4.417 & 8.197 & 0.548 & 2.405 \\ 0.995 & 0.548 & 1.379 & 0.38 \\ 9.677 & 2.405 & 0.38 & 6.681 \end{pmatrix}$$

Treatment Matrix B:

$m := 1 .. g$

$$B := \sum_m n_m \cdot (X_{bar_m} - X_{barGM}) \cdot (X_{bar_m} - X_{barGM})^T \qquad B = \begin{pmatrix} 3.4688 & 1.0986 & 0.8107 & 0.586 \\ 1.0986 & 1.2307 & 0.45 & 0.6157 \\ 0.8107 & 0.45 & 0.2318 & 0.2311 \\ 0.586 & 0.6157 & 0.2311 & 0.3086 \end{pmatrix}$$

2005 jw298A

MANOVA Table jw p. 299

Source	SSP Matrix	Degrees of Freedom	
Treatment Matrix B	$B = \begin{pmatrix} 3.4688 & 1.0986 & 0.8107 & 0.586 \\ 1.0986 & 1.2307 & 0.45 & 0.6157 \\ 0.8107 & 0.45 & 0.2318 & 0.2311 \\ 0.586 & 0.6157 & 0.2311 & 0.3086 \end{pmatrix}$	$df_B := g - 1$	$df_B = 2$
Residual/Error Matrix W	$W = \begin{pmatrix} 183.093 & 4.417 & 0.995 & 9.677 \\ 4.417 & 8.197 & 0.548 & 2.405 \\ 0.995 & 0.548 & 1.379 & 0.38 \\ 9.677 & 2.405 & 0.38 & 6.681 \end{pmatrix}$	$df_W := N - g$	$df_W = 513$
TOTAL (mean corrected)	$B + W = \begin{pmatrix} 186.5618 & 5.5156 & 1.8057 & 10.263 \\ 5.5156 & 9.4277 & 0.998 & 3.0207 \\ 1.8057 & 0.998 & 1.6108 & 0.6111 \\ 10.263 & 3.0207 & 0.6111 & 6.9896 \end{pmatrix}$	$df_T := N - 1$	$df_T = 515$

MANOVA tests jw p. 299-300.

Decomposition Model:

$X_{i,j} = \mu + \tau_i + \varepsilon_{i,j}$ **where:** $j = 1$ **to** n_m

$m = 1$ **to** g

Restriction:

$\Sigma n_m \tau_m = 0$

Hypothesis testing:

$H_0 : \tau_1 = \tau_2 = \dots \tau_g = 0$

H_1 **: at least one** τ <> **0**

Assumptions:

All populations P_1-P_g **rs**

$\varepsilon_{i,j}$ **rs** $N(0, \sigma^2)$

Population variances equal

Wilk's Lambda Test Statistic:

$$\Lambda_s := \frac{|W|}{|W + B|} \qquad \Lambda_s = 0.7628$$

Lawley-Hotelling Trace:

$$LHtr := tr\left(B \cdot W^{-1}\right) \qquad LHtr = 0.3001$$

Pillai trace:

$$Ptr := tr\left[B \cdot (B + W)^{-1}\right] \qquad Ptr = 0.2456$$

Roy's largest root:

$$Rrt := max\left[eigenvals\left[W \cdot (B + W)^{-1}\right]\right] \qquad Rrt = 1 \qquad eigenvals\left[W \cdot (B + W)^{-1}\right] = \begin{pmatrix} 1 \\ 0.9593 \\ 0.7951 \\ 1 \end{pmatrix}$$

Stringency of the test: $\alpha := 0.01$ **< set as desired**

2005 jw298A

If assumptions hold and H_0 is true then jw Table 6.3 & p. 300:

Number of Variables p: | **Number of groups g:** | **Sampling Distribution Critical Value C:**

p = 1 **g > 1** $C := qF(1-\alpha, g-1, N-g)$ $C = 4.6468$

Decision Rule: Reject H_0 if $K > C$:

$$K := \left(\frac{N-g}{g-1}\right)\cdot\left(\frac{1-\Lambda_s}{\Lambda_s}\right)$$ $K = 79.7796$

$\text{Decision} := \text{if}(K > C, 1, 0)$ $\text{Decision} = 1$ **< 0 = Do not reject H_0**
1 = Reject H_0

p = 2 **g > 1** $C := qF[1-\alpha, 2\cdot(g-1), 2\cdot(N-g-1)]$ $C = 3.3375$

Decision Rule: Reject H_0 if $K > C$:

$$K := \left(\frac{N-g-1}{g-1}\right)\cdot\left(\frac{1-\sqrt{\Lambda_s}}{\sqrt{\Lambda_s}}\right)$$ $K = 37.1207$

$\text{Decision} := \text{if}(K > C, 1, 0)$ $\text{Decision} = 1$ **< 0 = Do not reject H_0**
1 = Reject H_0

p > 0 **g = 2** $C := qF(1-\alpha, p, N-p-1)$ $C = 3.356$

Decision Rule: Reject H_0 if $K > C$:

$$K := \left(\frac{N-p-1}{p}\right)\cdot\left(\frac{1-\Lambda_s}{\Lambda_s}\right)$$ $K = 39.7343$

$\text{Decision} := \text{if}(K > C, 1, 0)$ $\text{Decision} = 1$ **< 0 = Do not reject H_0**
1 = Reject H_0

p > 0 **g = 3** $C := qF[1-\alpha, 2\cdot p, 2\cdot(N-p-2)]$ $C = 2.5287$

Decision Rule: Reject H_0 if $K > C$:

$$K := \left(\frac{N-p-2}{p}\right)\cdot\left(\frac{1-\sqrt{\Lambda_s}}{\sqrt{\Lambda_s}}\right)$$ $K = 18.4879$

$\text{Decision} := \text{if}(K > C, 1, 0)$ $\text{Decision} = 1$ **< 0 = Do not reject H_0**
1 = Reject H_0

Bartlett's test for Large N $C := qchisq[1-\alpha, p\cdot(g-1)]$ $C = 20.0902$

Decision Rule: Reject H_0 if $K > C$:

$$K := -\left(N-1-\frac{p+g}{2}\right)\cdot\ln(\Lambda_s)$$ $K = 138.5215$

$\text{Decision} := \text{if}(K > C, 1, 0)$ $\text{Decision} = 1$ **< 0 = Do not reject H_0**
1 = Reject H_0

Simultaneous Bonferroni confidence intervals for treatment effects jw p. 305:

$\alpha := 0.05$ **< Set probability of Type 1 error**

$$c := qt\left[1 - \frac{\alpha}{p \cdot g \cdot (g-1)}, N - g\right] \qquad c = 2.8782$$

Difference in means between first & second populations:

$i := 1 .. p$

$$\kappa_i := \left|\sqrt{\frac{W_{i,i}}{N-g} \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}\right| \qquad \kappa = \begin{pmatrix} 0.0625 \\ 0.0132 \\ 0.0054 \\ 0.0119 \end{pmatrix}$$

$$ci_{lower} := X_{bar_1} - X_{bar_2} - c \cdot \kappa \qquad ci_{upper} := X_{bar_1} - X_{bar_2} + c \cdot \kappa$$

$$ci := augment(ci_{lower}, ci_{upper}) \qquad ci = \begin{pmatrix} -0.2808 & 0.0788 \\ -0.154 & -0.078 \\ -0.0576 & -0.0264 \\ -0.0923 & -0.0237 \end{pmatrix}$$

Note: According to jw Result 6.5, Bonferroni simultaneous confidence intervals apply to ALL variables (=components of mean vectors) simultaneously for ALL pairwise comparisons across groups.

Difference in means between first & third populations:

$$\kappa_i := \left|\sqrt{\frac{W_{i,i}}{N-g} \cdot \left(\frac{1}{n_1} + \frac{1}{n_3}\right)}\right| \qquad \kappa = \begin{pmatrix} 0.0682 \\ 0.0144 \\ 0.0059 \\ 0.013 \end{pmatrix}$$

$$ci_{lower} := X_{bar_1} - X_{bar_2} - c \cdot \kappa \qquad ci_{upper} := X_{bar_1} - X_{bar_2} + c \cdot \kappa$$

$$ci := augment(ci_{lower}, ci_{upper}) \qquad ci = \begin{pmatrix} -0.2973 & 0.0953 \\ -0.1575 & -0.0745 \\ -0.059 & -0.025 \\ -0.0955 & -0.0205 \end{pmatrix}$$

Difference in means between second & third populations:

$$\kappa_i := \left|\sqrt{\frac{W_{i,i}}{N-g} \cdot \left(\frac{1}{n_2} + \frac{1}{n_3}\right)}\right| \qquad \kappa = \begin{pmatrix} 0.077 \\ 0.0163 \\ 0.0067 \\ 0.0147 \end{pmatrix}$$

$$ci_{lower} := X_{bar_1} - X_{bar_2} - c \cdot \kappa \qquad ci_{upper} := X_{bar_1} - X_{bar_2} + c \cdot \kappa$$

$$ci := augment(ci_{lower}, ci_{upper}) \qquad ci = \begin{pmatrix} -0.3225 & 0.1205 \\ -0.1629 & -0.0691 \\ -0.0612 & -0.0228 \\ -0.1003 & -0.0157 \end{pmatrix}$$

2005 jw298B

Prepared by:
Wm Stein

MANOVA - One Way
jw 298B.mcd

ORIGIN ≡ 1

Verifying calculation of Standardized Discriminant (Canonical) Coefficients

Reading the Data:

$Y := \text{READPRN}("\backslash DATA\backslash FOOTBALL.DAT")$

$X_2 := \text{READPRN}("\backslash DATA\backslash T6\text{-}7Slogs.txt")$

Head/Helmet measurements in A. C. Rencher *Methods of Multivariate Analysis* 4th edition. p. 306-307

Six head measurement variables
Three groups with 30 measurements each

Summary statistics of the sample:

$p := 6 \qquad g := 3 \qquad n := 30$

Mean Vectors & Variance-Covariance matrices:

$i := 1 .. n \qquad j := 1 .. p$

$l_{n_i} := 1 \qquad I := \text{identity}(n)$

SP (Within/Error partials)

$X := \text{submatrix}(Y, 1, 30, 2, 7)$

$X_{bar_1} := \frac{1}{n} \cdot X^T \cdot l_n \qquad X_{bar_1} = \begin{pmatrix} 15.2 \\ 58.937 \\ 20.10833 \\ 13.08333 \\ 14.73333 \\ 12.26667 \end{pmatrix}$

$SP_1 := X^T \cdot \left(I - \frac{1}{n} \cdot l_n \cdot l_n^T\right) \cdot X$

$X := \text{submatrix}(Y, 31, 60, 2, 7)$

$X_{bar_2} := \frac{1}{n} \cdot X^T \cdot l_n \qquad X_{bar_2} = \begin{pmatrix} 15.42 \\ 57.37967 \\ 19.80333 \\ 10.08 \\ 13.45333 \\ 11.94333 \end{pmatrix}$

$SP_2 := X^T \cdot \left(I - \frac{1}{n} \cdot l_n \cdot l_n^T\right) \cdot X$

$X := \text{submatrix}(Y, 61, 90, 2, 7)$

$X_{bar_3} := \frac{1}{n} \cdot X^T \cdot l_n \qquad X_{bar_3} = \begin{pmatrix} 15.58 \\ 57.77 \\ 19.81 \\ 10.94667 \\ 13.69667 \\ 11.80333 \end{pmatrix}$

$SP_3 := X^T \cdot \left(I - \frac{1}{n} \cdot l_n \cdot l_n^T\right) \cdot X$

Total Sample size:

$N := 3 \cdot n \qquad N = 90$

Grand Mean:

$m := 1 .. g$

$X_{barGM} := \frac{1}{N} \cdot \left[\sum_m \left(n \cdot X_{bar_m}\right)\right] \qquad X_{barGM} = \begin{pmatrix} 15.4 \\ 58.02889 \\ 19.90722 \\ 11.37 \\ 13.96111 \\ 12.00444 \end{pmatrix}$

2005 jw298B

Residual/Error/Within Matrix E:

$m := 1 .. g$

$E := \sum_m SP_m$

$$E = \begin{pmatrix} 37.256 & 50.2812 & 13.739 & 7.29 & 10.836 & 19.836 \\ 50.2812 & 275.01513 & 88.73778 & 56.7853 & 29.58553 & 43.94343 \\ 13.739 & 88.73778 & 47.49708 & 6.68217 & 11.18733 & 13.843 \\ 7.29 & 56.7853 & 6.68217 & 107.22433 & 27.38333 & 3.69467 \\ 10.836 & 29.58553 & 11.18733 & 27.38333 & 53.771 & 0.79433 \\ 19.836 & 43.94343 & 13.843 & 3.69467 & 0.79433 & 32.67 \end{pmatrix}$$

Treatment Matrix H:

$m := 1 .. g$

$H := n \cdot \left[\sum_m \left(X_{bar_m} - X_{barGM} \right) \cdot \left(X_{bar_m} - X_{barGM} \right)^T \right]$

$$H = \begin{pmatrix} 2.184 & -7.2362 & -1.794 & -13.34 & -6.366 & -2.696 \\ -7.2362 & 39.39536 & 8.25744 & 75.0897 & 32.98158 & 9.89601 \\ -1.794 & 8.25744 & 1.82072 & 15.59233 & 7.01294 & 2.35911 \\ -13.34 & 75.0897 & 15.59233 & 143.36467 & 62.70167 & 18.39733 \\ -6.366 & 32.98158 & 7.01294 & 62.70167 & 27.72289 & 8.60122 \\ -2.696 & 9.89601 & 2.35911 & 18.39733 & 8.60122 & 3.38822 \end{pmatrix}$$

MANOVA Table jw p. 299

Source	SSP Matrix	Degrees of Freedom	
Treatment Matrix H (same as B)	$H = \begin{pmatrix} 2.184 & -7.2362 & -1.794 & -13.34 & -6.366 & -2.696 \\ -7.2362 & 39.39536 & 8.25744 & 75.0897 & 32.98158 & 9.89601 \\ -1.794 & 8.25744 & 1.82072 & 15.59233 & 7.01294 & 2.35911 \\ -13.34 & 75.0897 & 15.59233 & 143.36467 & 62.70167 & 18.39733 \\ -6.366 & 32.98158 & 7.01294 & 62.70167 & 27.72289 & 8.60122 \\ -2.696 & 9.89601 & 2.35911 & 18.39733 & 8.60122 & 3.38822 \end{pmatrix}$	$df_H := g - 1$	$df_H = 2$
Residual/Error Matrix E (same as W)	$E = \begin{pmatrix} 37.256 & 50.2812 & 13.739 & 7.29 & 10.836 & 19.836 \\ 50.2812 & 275.01513 & 88.73778 & 56.7853 & 29.58553 & 43.94343 \\ 13.739 & 88.73778 & 47.49708 & 6.68217 & 11.18733 & 13.843 \\ 7.29 & 56.7853 & 6.68217 & 107.22433 & 27.38333 & 3.69467 \\ 10.836 & 29.58553 & 11.18733 & 27.38333 & 53.771 & 0.79433 \\ 19.836 & 43.94343 & 13.843 & 3.69467 & 0.79433 & 32.67 \end{pmatrix}$	$df_E := N - g$	$df_E = 87$
TOTAL (mean corrected)	$H + E = \begin{pmatrix} 39.44 & 43.045 & 11.945 & -6.05 & 4.47 & 17.14 \\ 43.045 & 314.41049 & 96.99522 & 131.875 & 62.56711 & 53.83944 \\ 11.945 & 96.99522 & 49.31781 & 22.2745 & 18.20028 & 16.20211 \\ -6.05 & 131.875 & 22.2745 & 250.589 & 90.085 & 22.092 \\ 4.47 & 62.56711 & 18.20028 & 90.085 & 81.49389 & 9.39556 \\ 17.14 & 53.83944 & 16.20211 & 22.092 & 9.39556 & 36.05822 \end{pmatrix}$	$df_T := N - 1$	$df_T = 89$

MANOVA tests jw p. 299-300.

Decomposition Model:

$X_{i,j} = \mu + \tau_i + \varepsilon_{i,j}$ where: $j = 1$ to n_m, $m = 1$ to g

Restriction:

$\Sigma n_m \tau_m = 0$

Hypothesis testing:

$H_0 : \tau_1 = \tau_2 = \ldots \tau_g = 0$

H_1 : at least one $\tau <> 0$

Assumptions:

All populations P_1-P_g rs

$\varepsilon_{i,j}$ rs $N(0, \sigma^2)$

Population variances equal

Wilk's Lambda Test Statistic:

$$\Lambda_S := \frac{|E|}{|E + H|} \qquad \Lambda_S = 0.30712$$

Lawley-Hotelling Trace:

$$LHtr := tr\left(H \cdot E^{-1}\right) \qquad LHtr = 2.03369$$

Pillai trace:

$$Ptr := tr\left[H \cdot (H + E)^{-1}\right] \qquad Ptr = 0.76116$$

Roy's largest root:

$$Rrt := max\left[eigenvals\left[E \cdot (H + E)^{-1}\right]\right] \qquad Rrt = 1 \qquad eigenvals\left[E \cdot (H + E)^{-1}\right] = \begin{pmatrix} 0.34273 \\ 1 \\ 0.89611 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

Stringency of the test: $\alpha := 0.01$ **< set as desired**

2005 jw298B

If assumptions hold and H_0 is true then jw Table 6.3 & p. 300:

Number of Variables p: **Number of groups g:** **Sampling Distribution Critical Value C:**

p = 1 **g > 1**

$C := qF(1-\alpha, g-1, N-g)$ $C = 4.85777$

Decision Rule: Reject H_0 if K > C:

$K := \left(\frac{N-g}{g-1}\right)\cdot\left(\frac{1-\Lambda_s}{\Lambda_s}\right)$ $K = 98.13705$ **Probability:** $1 - pF(K, g-1, N-g) = 0$

$\text{Decision} := \text{if}(K > C, 1, 0)$ $\text{Decision} = 1$ **< 0 = Do not reject H_0**
1 = Reject H_0

p = 2 **g > 1**

$C := qF[1-\alpha, 2\cdot(g-1), 2\cdot(N-g-1)]$ $C = 3.43011$

Decision Rule: Reject H_0 if K > C:

$K := \left(\frac{N-g-1}{g-1}\right)\cdot\left(\frac{1-\sqrt{\Lambda_s}}{\sqrt{\Lambda_s}}\right)$ $K = 34.59116$ **Probability:** $1 - pF[K, 2\cdot(g-1), 2\cdot(N-g-1)] = 0$

$\text{Decision} := \text{if}(K > C, 1, 0)$ $\text{Decision} = 1$ **< 0 = Do not reject H_0**
1 = Reject H_0

p > 0 **g = 2**

$C := qF(1-\alpha, p, N-p-1)$ $C = 3.0273$

Decision Rule: Reject H_0 if K > C:

$K := \left(\frac{N-p-1}{p}\right)\cdot\left(\frac{1-\Lambda_s}{\Lambda_s}\right)$ $K = 31.20833$ **Probability:** $1 - pF(K, p, N-p-1) = 0$

$\text{Decision} := \text{if}(K > C, 1, 0)$ $\text{Decision} = 1$ **< 0 = Do not reject H_0**
1 = Reject H_0

p > 0 **g = 3**

$C := qF[1-\alpha, 2\cdot p, 2\cdot(N-p-2)]$ $C = 2.29485$

Decision Rule: Reject H_0 if K > C:

$K := \left(\frac{N-p-2}{p}\right)\cdot\left(\frac{1-\sqrt{\Lambda_s}}{\sqrt{\Lambda_s}}\right)$ $K = 10.99409$ **Probability:** $1 - pF[K, 2\cdot p, 2\cdot(N-p-2)] = 0$

$\text{Decision} := \text{if}(K > C, 1, 0)$ $\text{Decision} = 1$ **< 0 = Do not reject H_0**
1 = Reject H_0

Bartlett's test for Large N

$C := qchisq[1-\alpha, p\cdot(g-1)]$ $C = 26.21693$

Decision Rule: Reject H_0 if K > C:

$K := -\left(N - 1 - \frac{p+g}{2}\right)\cdot\ln(\Lambda_s)$ $K = 99.75283$ **Probability:** $1 - pchisq[K, p\cdot(g-1)] = 0$

$\text{Decision} := \text{if}(K > C, 1, 0)$ $\text{Decision} = 1$ **< 0 = Do not reject H_0**
1 = Reject H_0

2005 jw298B

Simultaneous Bonferroni confidence intervals for treatment effects jw p. 305:

$\alpha := 0.05$ **< Set probability of Type 1 error**

$c := qt\left[1 - \frac{\alpha}{p \cdot g \cdot (g-1)}, N - g\right]$ $c = 3.07912$

Note: According to jw Result 6.5, Bonferroni simultaneous confidence intervals apply to ALL variables (=components of mean vectors) simultaneously for ALL pairwise comparisons across groups.

Difference in means between first & second populations:

$i := 1 .. p$

$\kappa_i := \left|\sqrt{\frac{E_{i,i}}{N-g} \cdot \left(\frac{1}{n} + \frac{1}{n}\right)}\right|$

$\kappa = \begin{pmatrix} 0.16896 \\ 0.45906 \\ 0.19078 \\ 0.28664 \\ 0.20299 \\ 0.15822 \end{pmatrix}$

$ci_{lower} := X_{bar_1} - X_{bar_2} - c \cdot \kappa$

$ci_{upper} := X_{bar_1} - X_{bar_2} + c \cdot \kappa$

$ci := \text{augment}(ci_{lower}, ci_{upper})$

$ci = \begin{pmatrix} -0.74026 & 0.30026 \\ 0.14382 & 2.97085 \\ -0.28243 & 0.89243 \\ 2.12072 & 3.88594 \\ 0.65498 & 1.90502 \\ -0.16385 & 0.81052 \end{pmatrix}$

Difference in means between first & third populations:

$\kappa_i := \left|\sqrt{\frac{E_{i,i}}{N-g} \cdot \left(\frac{1}{n} + \frac{1}{n}\right)}\right|$

$\kappa = \begin{pmatrix} 0.16896 \\ 0.45906 \\ 0.19078 \\ 0.28664 \\ 0.20299 \\ 0.15822 \end{pmatrix}$

$ci_{lower} := X_{bar_1} - X_{bar_3} - c \cdot \kappa$

$ci_{upper} := X_{bar_1} - X_{bar_3} + c \cdot \kappa$

$ci := \text{augment}(ci_{lower}, ci_{upper})$

$ci = \begin{pmatrix} -0.90026 & 0.14026 \\ -0.24651 & 2.58051 \\ -0.2891 & 0.88576 \\ 1.25406 & 3.01928 \\ 0.41164 & 1.66169 \\ -0.02385 & 0.95052 \end{pmatrix}$

Difference in means between second & third populations:

$\kappa_i := \left|\sqrt{\frac{E_{i,i}}{N-g} \cdot \left(\frac{1}{n} + \frac{1}{n}\right)}\right|$

$\kappa = \begin{pmatrix} 0.16896 \\ 0.45906 \\ 0.19078 \\ 0.28664 \\ 0.20299 \\ 0.15822 \end{pmatrix}$

$ci_{lower} := X_{bar_2} - X_{bar_3} - c \cdot \kappa$

$ci_{upper} := X_{bar_2} - X_{bar_3} + c \cdot \kappa$

$ci := \text{augment}(ci_{lower}, ci_{upper})$

$ci = \begin{pmatrix} -0.68026 & 0.36026 \\ -1.80385 & 1.02318 \\ -0.5941 & 0.58076 \\ -1.74928 & 0.01594 \\ -0.86836 & 0.38169 \\ -0.34719 & 0.62719 \end{pmatrix}$

Calculating Standardized Discriminant (canonical) Coefficients - Rencher p. 309

$$S_{pl} := \frac{E}{df_E}$$

$$S_{pl} = \begin{pmatrix} 0.42823 & 0.57794 & 0.15792 & 0.08379 & 0.12455 & 0.228 \\ 0.57794 & 3.16109 & 1.01997 & 0.6527 & 0.34006 & 0.5051 \\ 0.15792 & 1.01997 & 0.54594 & 0.07681 & 0.12859 & 0.15911 \\ 0.08379 & 0.6527 & 0.07681 & 1.23246 & 0.31475 & 0.04247 \\ 0.12455 & 0.34006 & 0.12859 & 0.31475 & 0.61806 & 0.00913 \\ 0.228 & 0.5051 & 0.15911 & 0.04247 & 0.00913 & 0.37552 \end{pmatrix}$$

< Pooled Within/Error variance

$$\lambda := \text{reverse}(\text{sort}(\text{eigenvals}(E^{-1} \cdot H))) \qquad \lambda = \begin{pmatrix} 1.91776 \\ 0.11593 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

< Eigenvalues of $E^{-1}H$ p. 304-305

Canonical correlations - Rencher p. 192-193.

$$r_1 := \sqrt{\frac{\lambda_1}{1+\lambda_1}} \qquad r_1 = 0.81072$$

$$r_2 := \sqrt{\frac{\lambda_2}{1+\lambda_2}} \qquad r_2 = 0.32232$$

< Measures degree of association of variables with groups

unit length eigenvectors ε:

$$\varepsilon_1 := \text{eigenvec}(E^{-1} \cdot H, \lambda_1)$$

$$\varepsilon_2 := \text{eigenvec}(E^{-1} \cdot H, \lambda_2)$$

$$\varepsilon_1 = \begin{pmatrix} -0.63093 \\ 0.00242 \\ 0.00428 \\ 0.43073 \\ 0.33552 \\ 0.55118 \end{pmatrix} \qquad \varepsilon_2 = \begin{pmatrix} 0.64503 \\ -0.00024 \\ -0.01312 \\ 0.24772 \\ -0.17603 \\ -0.70101 \end{pmatrix}$$

< Eigenvectors comprise coefficients for the discriminant function describing directions of maximum separation between groups

$$\varepsilon_1^T \cdot \varepsilon_1 = (1)$$

$$\varepsilon_1^T \cdot S_{pl} \cdot \varepsilon_1 = (0.44255)$$

$$\varepsilon_2^T \cdot S_{pl} \cdot \varepsilon_2 = (0.21024)$$

< scaling factors:

$$\kappa := \frac{1}{\sqrt{\left|\varepsilon_1^T \cdot S_{pl} \cdot \varepsilon_1\right|}} \qquad \kappa = 1.50321$$

$$e_1 := \kappa \cdot \varepsilon_1$$

$$\kappa\kappa := \frac{1}{\sqrt{\left|\varepsilon_2^T \cdot S_{pl} \cdot \varepsilon_2\right|}} \qquad \kappa\kappa = 2.18094$$

$$e_2 := \kappa\kappa \cdot \varepsilon_2$$

$$\phi_j := S_{pl_{j,j}}$$

scaled eigenvectors e:

$$e_1 = \begin{pmatrix} -0.94842 \\ 0.00364 \\ 0.00644 \\ 0.64748 \\ 0.50436 \\ 0.82854 \end{pmatrix} \qquad e_2 = \begin{pmatrix} 1.40678 \\ -0.00051 \\ -0.02862 \\ 0.54027 \\ -0.38391 \\ -1.52886 \end{pmatrix}$$

Eigenvectors ε must be scaled in length so that the above equations equal unity

$$\left(\frac{1}{\sqrt{\left|\varepsilon_1^T \cdot S_{pl} \cdot \varepsilon_1\right|}} \cdot \varepsilon_1\right)^T \cdot S_{pl} \cdot \left(\frac{1}{\sqrt{\left|\varepsilon_1^T \cdot S_{pl} \cdot \varepsilon_1\right|}} \cdot \varepsilon_1\right) = (1)$$

$$\left(\frac{1}{\sqrt{\left|\varepsilon_2^T \cdot S_{pl} \cdot \varepsilon_2\right|}} \cdot \varepsilon_2\right)^T \cdot S_{pl} \cdot \left(\frac{1}{\sqrt{\left|\varepsilon_2^T \cdot S_{pl} \cdot \varepsilon_2\right|}} \cdot \varepsilon_2\right) = (1)$$

Scaled eigenvectors e in the above equations now equal unity

$$\text{diag}(\sqrt{\phi}) \cdot e_1 = \begin{pmatrix} -0.62064 \\ 0.00647 \\ 0.00476 \\ 0.71881 \\ 0.39651 \\ 0.50772 \end{pmatrix} \qquad \text{diag}(\sqrt{\phi}) \cdot e_2 = \begin{pmatrix} 0.92058 \\ -0.00091 \\ -0.02115 \\ 0.59979 \\ -0.30182 \\ -0.93687 \end{pmatrix}$$

< See Rencher p. 310. Standardized discriminant (canonical) coefficients are a way to see how much effect a variable has with regard to the differences used to reject H_o

2005 jw313

ORIGIN ≡ 1

Two Way MANOVA - Fixed Effects
jw 313.mcd

Prepared by:
Wm Stein

Plastic Film Data, five replicates per treatment block, and two fixed factors:

Factor A: Rate of Extrusion (low, high) first column of Y
Factor B: Amount of Additive (low, high) second column Y

Reading the Data:

Y := READPRN("\DATA\T6-4.dat")

Data collected for three variables: X1, X2, X3 in columns 4-6 of Y

Y =

	1	2	3	4	5
1	0	0	6.5	9.5	4.4
2	0	0	6.2	9.9	6.4
3	0	0	5.8	9.6	3
4	0	0	6.5	9.6	4.1
5	0	0	6.5	9.2	0.8
6	0	1	6.9	9.1	5.7
7	0	1	7.2	10	2
8	0	1	6.9	9.9	3.9
9	0	1	6.1	9.5	1.9
10	0	1	6.3	9.4	5.7
11	1	0	6.7	9.1	2.8
12	1	0	6.6	9.3	4.1
13	1	0	7.2	8.3	3.8
14	1	0	7.1	8.4	1.6
15	1	0	6.8	8.5	3.4
16	1	1	7.1	9.2	8.4
17	1	1	7	8.8	5.2
18	1	1	7.2	9.7	6.9
19	1	1	7.5	10.1	2.7
20	1	1	7.6	9.2	1.9

Summary statistics of the sample:

$n := 5 \quad a := 2 \quad b := 2 \quad p := 3$

Block Means:

$i := 1..a \quad j := 1..b \quad k := 1..p \quad m := 1..n$

$l_{n_m} := 1 \qquad I := \text{identity}(n)$

Block SSP:

$X := \text{submatrix}(Y,1,5,3,5)$

$Xbar_{1,1} := \frac{1}{n}\cdot X^T\cdot l_n \qquad Xbar_{1,1} = \begin{pmatrix} 6.3 \\ 9.56 \\ 3.74 \end{pmatrix} \qquad SP_1 := X^T\cdot\left(I - \frac{1}{n}\cdot l_n\cdot l_n^T\right)\cdot X$

$X := \text{submatrix}(Y,6,10,3,5)$

$Xbar_{1,2} := \frac{1}{n}\cdot X^T\cdot l_n \qquad Xbar_{1,2} = \begin{pmatrix} 6.68 \\ 9.58 \\ 3.84 \end{pmatrix} \qquad SP_2 := X^T\cdot\left(I - \frac{1}{n}\cdot l_n\cdot l_n^T\right)\cdot X$

$X := \text{submatrix}(Y,11,15,3,5)$

$Xbar_{2,1} := \frac{1}{n}\cdot X^T\cdot l_n \qquad Xbar_{2,1} = \begin{pmatrix} 6.88 \\ 8.72 \\ 3.14 \end{pmatrix} \qquad SP_3 := X^T\cdot\left(I - \frac{1}{n}\cdot l_n\cdot l_n^T\right)\cdot X$

$X := \text{submatrix}(Y,16,20,3,5)$

$Xbar_{2,2} := \frac{1}{n}\cdot X^T\cdot l_n \qquad Xbar_{2,2} = \begin{pmatrix} 7.28 \\ 9.4 \\ 5.02 \end{pmatrix} \qquad SP_4 := X^T\cdot\left(I - \frac{1}{n}\cdot l_n\cdot l_n^T\right)\cdot X$

$Xbar = \begin{pmatrix} \{3,1\} & \{3,1\} \\ \{3,1\} & \{3,1\} \end{pmatrix}$

Factor A Means:

$N_A := a\cdot n \qquad m := 1..N_A \qquad l_{n_m} := 1$

$X := \text{stack}(\text{submatrix}(Y,1,5,3,5), \text{submatrix}(Y,6,10,3,5))$

$XbarA_1 := \frac{1}{N_A}\cdot X^T\cdot l_n \qquad XbarA_1 = \begin{pmatrix} 6.49 \\ 9.57 \\ 3.79 \end{pmatrix}$

$X := \text{stack}(\text{submatrix}(Y,11,15,3,5), \text{submatrix}(Y,16,20,3,5))$

$XbarA_2 := \frac{1}{N_A}\cdot X^T\cdot l_n \qquad XbarA_2 = \begin{pmatrix} 7.08 \\ 9.06 \\ 4.08 \end{pmatrix}$

$XbarA = \begin{pmatrix} \{3,1\} \\ \{3,1\} \end{pmatrix}$

Factor B Means:

$N_B := b \cdot n \qquad m := 1 .. N_B \qquad ml_n := 1$

$X := \text{stack}(\text{submatrix}(Y,1,5,3,5),\text{submatrix}(Y,11,15,3,5))$

$X_{barB_1} := \frac{1}{N_B} \cdot X^T \cdot l_n \qquad X_{barB_1} = \begin{pmatrix} 6.59 \\ 9.14 \\ 3.44 \end{pmatrix}$

$X := \text{stack}(\text{submatrix}(Y,6,10,3,5),\text{submatrix}(Y,16,20,3,5))$

$X_{barB_2} := \frac{1}{N_B} \cdot X^T \cdot l_n \qquad X_{barB_2} = \begin{pmatrix} 6.98 \\ 9.49 \\ 4.43 \end{pmatrix} \qquad X_{barB} = \begin{pmatrix} \{3,1\} \\ \{3,1\} \end{pmatrix}$

Grand Mean:

$N := a \cdot b \cdot n \qquad m := 1 .. N \qquad l_{n_m} := 1$

$X := \text{submatrix}(Y,1,20,3,5)$

$X_{GM} := \frac{1}{N} \cdot X^T \cdot l_n \qquad X_{GM} = \begin{pmatrix} 6.785 \\ 9.315 \\ 3.935 \end{pmatrix}$

Sums of Squares & Products Matrices:

$SSP_A := \sum_i b \cdot n \cdot \left(X_{barA_i} - X_{GM}\right) \cdot \left(X_{barA_i} - X_{GM}\right)^T \qquad SSP_A = \begin{pmatrix} 1.7405 & -1.5045 & 0.8555 \\ -1.5045 & 1.3005 & -0.7395 \\ 0.8555 & -0.7395 & 0.4205 \end{pmatrix}$

$SSP_B := \sum_j a \cdot n \cdot \left(X_{barB_j} - X_{GM}\right) \cdot \left(X_{barB_j} - X_{GM}\right)^T \qquad SSP_B = \begin{pmatrix} 0.7605 & 0.6825 & 1.9305 \\ 0.6825 & 0.6125 & 1.7325 \\ 1.9305 & 1.7325 & 4.9005 \end{pmatrix}$

$SSP_{AB} := \sum_i \sum_j n \cdot \left(X_{bar_{i,j}} - X_{barA_i} - X_{barB_j} + X_{GM}\right) \cdot \left(X_{bar_{i,j}} - X_{barA_i} - X_{barB_j} + X_{GM}\right)^T$

$SSP_{AB} = \begin{pmatrix} 0.0005 & 0.0165 & 0.0445 \\ 0.0165 & 0.5445 & 1.4685 \\ 0.0445 & 1.4685 & 3.9605 \end{pmatrix}$

$m := 1 .. 4$

$SSP_E := \sum_m SP_m \qquad SSP_E = \begin{pmatrix} 1.764 & 0.02 & -3.07 \\ 0.02 & 2.628 & -0.552 \\ -3.07 & -0.552 & 64.924 \end{pmatrix}$

$I := \text{identity}(N)$

$SSP_T := X^T \cdot \left(I - \frac{1}{N} \cdot l_n \cdot l_n^T\right) \cdot X \qquad SSP_T = \begin{pmatrix} 4.2655 & -0.7855 & -0.2395 \\ -0.7855 & 5.0855 & 1.9095 \\ -0.2395 & 1.9095 & 74.2055 \end{pmatrix}$

2005 jw313

MANOVA Table jw p. 310

Source	SSP Matrix	Degrees of Freedom	
Factor A	$SSP_A = \begin{pmatrix} 1.7405 & -1.5045 & 0.8555 \\ -1.5045 & 1.3005 & -0.7395 \\ 0.8555 & -0.7395 & 0.4205 \end{pmatrix}$	$df_A := a - 1$	$df_A = 1$
Factor B	$SSP_B = \begin{pmatrix} 0.7605 & 0.6825 & 1.9305 \\ 0.6825 & 0.6125 & 1.7325 \\ 1.9305 & 1.7325 & 4.9005 \end{pmatrix}$	$df_B := b - 1$	$df_B = 1$
Interaction	$SSP_{AB} = \begin{pmatrix} 0.0005 & 0.0165 & 0.0445 \\ 0.0165 & 0.5445 & 1.4685 \\ 0.0445 & 1.4685 & 3.9605 \end{pmatrix}$	$df_{AB} := (a-1)\cdot(b-1)$	$df_{AB} = 1$
Residual/Error (Matrix E or W)	$SSP_E = \begin{pmatrix} 1.764 & 0.02 & -3.07 \\ 0.02 & 2.628 & -0.552 \\ -3.07 & -0.552 & 64.924 \end{pmatrix}$	$df_E := a\cdot b\cdot(n-1)$	$df_E = 16$
TOTAL (mean corrected)	$SSP_T = \begin{pmatrix} 4.2655 & -0.7855 & -0.2395 \\ -0.7855 & 5.0855 & 1.9095 \\ -0.2395 & 1.9095 & 74.2055 \end{pmatrix}$	$df_T := a\cdot b\cdot n - 1$	$df_T = 19$

MANOVA tests jw p. 311-312

Decomposition Model:

$$X_{i,j,m} = \mu + \alpha_i + \beta_j + \gamma_{i,j} + \varepsilon_{i,j}$$

where: i = 1 to a (factors of A)
j = 1 to b (factors of B)
m = 1 to n (replicates)

Restriction:

$$\Sigma\, \alpha_i = \Sigma_i\, \beta_j = \Sigma_j\, \gamma_{i,j} = \Sigma\, \gamma_{i,j} = 0$$

Assumptions:

All populations P_1-P_g rs
$\varepsilon_{i,j}$ rs $N(0, \sigma^2)$
Population variances equal

Hypothesis testing:

Factor Interaction:

H_0 : all $\gamma_{i,j}$ = 0 for all i & j

H_1 : at least one $\gamma_{i,j}$ <> 0

Wilk's Lambda Test Statistic:

$$\Lambda_s := \frac{\left|SSP_E\right|}{\left|SSP_{AB} + SSP_E\right|} \qquad \Lambda_s = 0.77711$$

Lawley-Hotelling Trace:

$$LHtr := tr\left(SSP_{AB} \cdot SSP_E^{-1}\right) \qquad LHtr = 0.28683$$

Pillai trace:

$$Ptr := tr\left[SSP_{AB} \cdot \left(SSP_{AB} + SSP_E\right)^{-1}\right] \qquad Ptr = 0.22289$$

Stringency of the test: $\alpha := 0.05$ **< set as desired**

Exact F distribution for (a-1)(b-1)=1

$$\nu_1 := \left|(a-1)\cdot(b-1) - p\right| + 1$$

$$\nu_2 := a \cdot b \cdot (n-1) - p + 1$$

$$C := qF\left(1 - \alpha, \nu_1, \nu_2\right) \qquad C = 3.34389$$

$$K := \left[\frac{\frac{[a \cdot b \cdot (n-1) - p + 1]}{2}}{\frac{[\left|(a-1)\cdot(b-1) - p\right| + 1]}{2}}\right] \cdot \left(\frac{1 - \Lambda_s}{\Lambda_s}\right) \qquad K = 1.33852$$

Probability:

$$1 - pF\left(K, \nu_1, \nu_2\right) = 0.30178$$

$$Decision := if\left(K > C, 1, 0\right) \qquad Decision = 0$$

< 0 = Do not reject H_0
1 = Reject H_0

Bartlett's test for Large N

$$C := qchisq\left[1 - \alpha, p \cdot (a-1) \cdot (b-1)\right] \qquad C = 7.81473$$

Decision Rule: Reject H_0 if K > C:

$$K := -\left[a \cdot b \cdot (n-1) - \frac{p + 1 - (a-1)\cdot(b-1)}{2}\right] \cdot \ln\left(\Lambda_s\right) \qquad K = 3.65659$$

Probability:

$$1 - pchisq\left[K, p \cdot (a-1) \cdot (b-1)\right] = 0.30101$$

$$Decision := if\left(K > C, 1, 0\right) \qquad Decision = 0$$

< 0 = Do not reject H_0
1 = Reject H_0

2005 jw313

Factor A Effect:

$H_0 : \alpha_1 = \alpha_2 = ... \alpha_i = 0$

H_1 : at least one α <> 0

Wilk's Lambda Test Statistic:

$$\Lambda_s := \frac{|SSP_E|}{|SSP_A + SSP_E|} \qquad \Lambda_s = 0.38186$$

Lawley-Hotelling Trace:

$$LHtr := tr\left(SSP_A \cdot SSP_E^{-1}\right) \qquad LHtr = 1.61877$$

Pillai trace:

$$Ptr := tr\left[SSP_A \cdot (SSP_A + SSP_E)^{-1}\right] \qquad Ptr = 0.61814$$

Stringency of the test: $\alpha := 0.05$ **< set as desired**

Exact F distribution for (a-1)(b-1)=1

$$\nu_1 := |(a-1) - p| + 1$$

$$\nu_2 := a \cdot b \cdot (n-1) - p + 1$$

$$C := qF(1-\alpha, \nu_1, \nu_2) \qquad C = 3.34389$$

$$K := \left[\frac{\frac{[a \cdot b \cdot (n-1) - p + 1]}{2}}{\frac{[|(a-1)-p| + 1]}{2}}\right] \cdot \left(\frac{1-\Lambda_s}{\Lambda_s}\right) \qquad K = 7.55427$$

Probability:

$$1 - pF(K, \nu_1, \nu_2) = 0.00303$$

$Decision := if(K > C, 1, 0)$ $\qquad Decision = 1$ **< 0 = Do not reject H_0**
1 = Reject H_0

Bartlett's test for Large N

$$C := qchisq[1-\alpha, p \cdot (a-1)] \qquad C = 7.81473$$

Decision Rule: Reject H_0 if $K > C$:

$$K := -\left[a \cdot b \cdot (n-1) - \frac{p + 1 - (a-1)}{2}\right] \cdot \ln(\Lambda_s) \qquad K = 13.95923$$

Probability:

$$1 - pchisq[K, p \cdot (a-1)] = 0.00296$$

$Decision := if(K > C, 1, 0)$ $\qquad Decision = 1$ **< 0 = Do not reject H_0**
1 = Reject H_0

Factor B Effect:

$H_0 : \beta_1 = \beta_2 = ... \beta_j = 0$

H_1 : at least one β <> 0

Wilk's Lambda Test Statistic:

$$\Lambda_s := \frac{|SSP_E|}{|SSP_B + SSP_E|} \qquad \Lambda_s = 0.52303$$

Lawley-Hotelling Trace:

$$LHtr := tr\left(SSP_B \cdot SSP_E^{-1}\right) \qquad LHtr = 0.91192$$

Pillai trace:

$$Ptr := tr\left[SSP_B \cdot \left(SSP_B + SSP_E\right)^{-1}\right] \qquad Ptr = 0.47697$$

Stringency of the test: $\alpha := 0.05$ **< set as desired**

Exact F distribution for (a-1)(b-1)=1

$$\nu_1 := |(b-1) - p| + 1$$

$$\nu_2 := a \cdot b \cdot (n-1) - p + 1$$

$$C := qF\left(1 - \alpha, \nu_1, \nu_2\right) \qquad C = 3.34389$$

$$K := \left[\frac{\frac{[a \cdot b \cdot (n-1) - p + 1]}{2}}{\frac{[|(b-1) - p| + 1]}{2}}\right] \cdot \left(\frac{1 - \Lambda_s}{\Lambda_s}\right) \qquad K = 4.25562$$

Probability:

$$1 - pF\left(K, \nu_1, \nu_2\right) = 0.02475$$

$Decision := if(K > C, 1, 0)$ $\qquad Decision = 1$ **< 0 = Do not reject H_0**
1 = Reject H_0

Bartlett's test for Large N

$$C := qchisq[1 - \alpha, p \cdot (b-1)] \qquad C = 7.81473$$

Decision Rule: Reject H_0 if K > C:

$$K := -\left[a \cdot b \cdot (n-1) - \frac{p + 1 - (b-1)}{2}\right] \cdot \ln\left(\Lambda_s\right) \qquad K = 9.39755$$

Probability:

$$1 - pchisq[K, p \cdot (b-1)] = 0.02445$$

$Decision := if(K > C, 1, 0)$ $\qquad Decision = 1$ **< 0 = Do not reject H_0**
1 = Reject H_0

Simultaneous Bonferroni confidence intervals for treatment effects jw p. 312:
Note: use this only after determining no interaction effects above
- i.e. failure to reject H_o for hypotheses involving γ_{ij}

Note: According to jw Result 6.5, Bonferroni simultaneous confidence intervals apply to ALL variables (=components of mean vectors) simultaneously for ALL pairwise comparisons across groups.

For treatment A effect:

$\alpha := 0.05$ **< Set probability of Type 1 error**

$$c := qt\left[1 - \frac{\alpha}{p\cdot a\cdot(a-1)}, a\cdot b\cdot(n-1)\right] \qquad c = 2.67303$$

Difference in means between first & second populations:

$$\kappa_k := \left|\sqrt{\frac{SSP_{E_{k,k}}}{a\cdot b\cdot(n-1)}\cdot\left(\frac{2}{b\cdot n}\right)}\right| \qquad \kappa = \begin{pmatrix} 0.14849 \\ 0.18125 \\ 0.90086 \end{pmatrix}$$

$$ci_{lower} := X_{barA_1} - X_{barA_2} - c\cdot\kappa \qquad ci_{upper} := X_{barA_1} - X_{barA_2} + c\cdot\kappa$$

$$ci := augment(ci_{lower}, ci_{upper}) \qquad ci = \begin{pmatrix} -0.98693 & -0.19307 \\ 0.02552 & 0.99448 \\ -2.69803 & 2.11803 \end{pmatrix}$$

For treatment B effect:

$\alpha := 0.05$ **< Set probability of Type 1 error**

$$c := qt\left[1 - \frac{\alpha}{p\cdot b\cdot(b-1)}, a\cdot b\cdot(n-1)\right] \qquad c = 2.67303$$

Difference in means between first & second populations:

$$\kappa_k := \left|\sqrt{\frac{SSP_{E_{k,k}}}{a\cdot b\cdot(n-1)}\cdot\left(\frac{2}{a\cdot n}\right)}\right| \qquad \kappa = \begin{pmatrix} 0.14849 \\ 0.18125 \\ 0.90086 \end{pmatrix}$$

$$ci_{lower} := X_{barB_1} - X_{barB_2} - c\cdot\kappa \qquad ci_{upper} := X_{barB_1} - X_{barB_2} + c\cdot\kappa$$

$$ci := augment(ci_{lower}, ci_{upper}) \qquad ci = \begin{pmatrix} -0.78693 & 0.00693 \\ -0.83448 & 0.13448 \\ -3.39803 & 1.41803 \end{pmatrix}$$

2005 jw313

TWO-WAY UNIVARIATE ANOVA TABLE - fixed factors A & B:

Source of variation **Sum of Squares** **Degrees of Freedom** **Mean Squares**

FACTOR A

$$SSA := n \cdot b \cdot \sum_i \left(X_{barA_i} - X_{GM}\right)^2 \quad SSA = \begin{pmatrix} 1.7405 \\ 1.3005 \\ 0.4205 \end{pmatrix} \quad df_A := a - 1 \quad df_A = 1 \quad MS_A := \frac{SSA}{df_A} \quad MS_A = \begin{pmatrix} 1.7405 \\ 1.3005 \\ 0.4205 \end{pmatrix}$$

FACTOR B

$$SSB := n \cdot a \cdot \sum_j \left(X_{barB_j} - X_{GM}\right)^2 \quad SSB = \begin{pmatrix} 0.7605 \\ 0.6125 \\ 4.9005 \end{pmatrix} \quad df_B := b - 1 \quad df_B = 1 \quad MS_B := \frac{SSB}{df_B} \quad MS_B = \begin{pmatrix} 0.7605 \\ 0.6125 \\ 4.9005 \end{pmatrix}$$

AB INTERACTION

$$SSAB := n \cdot \sum_i \sum_j \left(X_{bar_{i,j}} - X_{barA_i} - X_{barB_j} + X_{GM}\right)^2$$

$$SSAB = \begin{pmatrix} 0.0005 \\ 0.5445 \\ 3.9605 \end{pmatrix} \quad df_{AB} := (a - 1) \cdot (b - 1) \quad df_{AB} = 1 \quad MS_{AB} := \frac{SSAB}{df_{AB}} \quad MS_{AB} = \begin{pmatrix} 0.0005 \\ 0.5445 \\ 3.9605 \end{pmatrix}$$

ERROR

$$SSE_k := SSP_{E_{k,k}}$$

$$SSE = \begin{pmatrix} 1.764 \\ 2.628 \\ 64.924 \end{pmatrix} \quad df_E := a \cdot b \cdot (n - 1) \quad df_E = 16 \quad MS_E := \frac{SSE}{df_E} \quad MS_E = \begin{pmatrix} 0.11025 \\ 0.16425 \\ 4.05775 \end{pmatrix}$$

TOTAL

$$SST_k := SSP_{T_{k,k}}$$

$$SST = \begin{pmatrix} 4.2655 \\ 5.0855 \\ 74.2055 \end{pmatrix} \quad df_T := n \cdot a \cdot b - 1 \quad df_T = 19$$

Decomposition Model:

$$X_{i,j} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{i,j} + \varepsilon_{i,j}$$

Restriction:

$$\Sigma\ \alpha_i = \Sigma_i\ \beta_j = \Sigma_j\ \gamma_{i,j} = \Sigma\ \gamma_{i,j} = 0$$

Assumptions:

All populations P_1-P_g rs

$\varepsilon_{i,j}$ rs $N(0, \sigma^2)$

Population variances equal

Hypothesis testing:

Factor Interaction:

H_0 : all $\gamma_{i,j} = 0$ for all i & j

H_1 : at least one $\gamma_{i,j}$ <> 0

ANOVA test statistics:

Probability:

$$F_k := \frac{MS_{AB_k}}{MS_{E_k}} \quad F = \begin{pmatrix} 0.00454 \\ 3.31507 \\ 0.97603 \end{pmatrix} \quad Prob := 1 - pF\left(F, df_{AB}, df_E\right) \quad Prob = \begin{pmatrix} 0.94714 \\ 0.0874 \\ 0.33789 \end{pmatrix}$$

Stringency of the test: $\alpha := 0.05$ **< set as desired**

If assumptions hold and H_0 is true then:

$C := qF(1-\alpha, df_{AB}, df_E)$ $C = 4.494$

NOTE: qF(1-α) is substituted here for F(α) in the text.

Decision Rule: Reject H_0 if T > C

$Decision_k := if(F_k > C, 1, 0)$ $Decision = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ **< 0 = Do not reject H_0**
1 = Reject H_0

Factor A effect:

$H_0 : \alpha_1 = \alpha_2 = ... \alpha_i = 0$

H_1 : at least one α <> 0

ANOVA test statistics:

$F_k := \frac{MS_{A_k}}{MS_{E_k}}$ $F = \begin{pmatrix} 15.78685 \\ 7.91781 \\ 0.10363 \end{pmatrix}$

Probability:

$Prob := 1 - pF(F, df_A, df_E)$ $Prob = \begin{pmatrix} 0.00109 \\ 0.01248 \\ 0.75169 \end{pmatrix}$

Stringency of the test: $\alpha := 0.05$ **< set as desired**

If assumptions hold and H_0 is true then:

$C := qF(1-\alpha, df_A, df_E)$ $C = 4.494$

NOTE: qF(1-α) is substituted here for F(α) in the text.

Decision Rule: Reject H_0 if T > C

$Decision_k := if(F_k > C, 1, 0)$ $Decision = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ **< 0 = Do not reject H_0**
1 = Reject H_0

Factor B effect:

$H_0 : \beta_1 = \beta_2 = ... \beta_j = 0$

H_1 : at least one β <> 0

ANOVA test statistics:

$F_k := \frac{MS_{B_k}}{MS_{E_k}}$ $F = \begin{pmatrix} 6.89796 \\ 3.72907 \\ 1.20769 \end{pmatrix}$

Probability:

$Prob := 1 - pF(F, df_B, df_E)$ $Prob = \begin{pmatrix} 0.01833 \\ 0.07139 \\ 0.28805 \end{pmatrix}$

Stringency of the test: $\alpha := 0.05$ **< set as desired**

If assumptions hold and H_0 is true then:

$C := qF(1-\alpha, df_B, df_E)$ $C = 4.494$

NOTE: qF(1-α) is substituted here for F(α) in the text.

Decision Rule: Reject H_0 if T > C

$Decision_k := if(F_k > C, 1, 0)$ $Decision = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ **< 0 = Do not reject H_0**
1 = Reject H_0

2005 jw313

Calculating Standardized Discriminant (canonical) Coefficients - Rencher p. 309

$S_{pl} := \frac{SSP_E}{df_E}$ $E := SSP_E$

$S_{pl} = \begin{pmatrix} 0.11025 & 0.00125 & -0.19187 \\ 0.00125 & 0.16425 & -0.0345 \\ -0.19187 & -0.0345 & 4.05775 \end{pmatrix}$ **< Pooled Within/Error variance**

For Interaction of Factors A & B:

$H := SSP_{AB}$

$\lambda := \text{reverse}(\text{sort}(\text{eigenvals}(E^{-1} \cdot H)))$ $\lambda = \begin{pmatrix} 0.28683 \\ 0 \\ 0 \end{pmatrix}$ **< Eigenvalues of $E^{-1}H$ p. 304-305**

$\varepsilon_1 := \text{eigenvec}(E^{-1} \cdot H, \lambda_1)$ **< unit length eigenvector ε:**

Canonical Correlations:

$r_1 := \sqrt{\frac{\lambda_1}{1+\lambda_1}}$ $r_1 = 0.47212$

$\kappa := \frac{1}{\sqrt{\left|\varepsilon_1^T \cdot S_{pl} \cdot \varepsilon_1\right|}}$ $\kappa = 2.23515$ **< scaling factor**

$e_1 := \kappa \cdot \varepsilon_1$ $\phi_k := S_{pl_{k,k}}$

$\varepsilon_1 = \begin{pmatrix} 0.24403 \\ 0.96205 \\ 0.12215 \end{pmatrix}$ $e_1 = \begin{pmatrix} 0.54544 \\ 2.15032 \\ 0.27301 \end{pmatrix}$ **< Scaled eigenvector e**

$\text{diag}(\sqrt{\phi}) \cdot e_1 = \begin{pmatrix} 0.18111 \\ 0.87148 \\ 0.54995 \end{pmatrix}$ **< Standardized discriminant coefficients**

For Factor A:

$H := SSP_A$

$\lambda := \text{reverse}(\text{sort}(\text{eigenvals}(E^{-1} \cdot H)))$ $\lambda = \begin{pmatrix} 1.61877 \\ 0 \\ 0 \end{pmatrix}$ **< Eigenvalues of $E^{-1}H$ p. 304-305**

$\varepsilon_1 := \text{eigenvec}(E^{-1} \cdot H, \lambda_1)$ **< unit length eigenvector ε:**

Canonical Correlations:

$r_1 := \sqrt{\frac{\lambda_1}{1+\lambda_1}}$ $r_1 = 0.78622$

$\kappa := \frac{1}{\sqrt{\left|\varepsilon_1^T \cdot S_{pl} \cdot \varepsilon_1\right|}}$ $\kappa = 2.94966$ **< scaling factor**

$e_1 := \kappa \cdot \varepsilon_1$ $\phi_k := S_{pl_{k,k}}$

$\varepsilon_1 = \begin{pmatrix} 0.88708 \\ -0.45904 \\ 0.04869 \end{pmatrix}$ $e_1 = \begin{pmatrix} 2.61658 \\ -1.35401 \\ 0.14362 \end{pmatrix}$ **< Scaled eigenvector e**

$\text{diag}(\sqrt{\phi}) \cdot e_1 = \begin{pmatrix} 0.86881 \\ -0.54875 \\ 0.2893 \end{pmatrix}$ **< Standardized discriminant coefficients**

For Factor B:

$H := SSP_B$

$\lambda := \text{reverse}(\text{sort}(\text{eigenvals}(E^{-1} \cdot H)))$ $\lambda = \begin{pmatrix} 0.91192 \\ 0 \\ 0 \end{pmatrix}$ **< Eigenvalues of $E^{-1}H$ p. 304-305**

$\varepsilon_1 := \text{eigenvec}(E^{-1} \cdot H, \lambda_1)$ **< unit length eigenvector ε:**

Canonical Correlations:

$r_1 := \sqrt{\frac{\lambda_1}{1+\lambda_1}}$ $r_1 = 0.69063$

$\kappa := \frac{1}{\sqrt{\left|\varepsilon_1^T \cdot S_{pl} \cdot \varepsilon_1\right|}}$ $\kappa = 2.85297$ **< scaling factor**

$e_1 := \kappa \cdot \varepsilon_1$ $\phi_k := S_{pl_{k,k}}$

$\varepsilon_1 = \begin{pmatrix} 0.88756 \\ 0.45061 \\ 0.09586 \end{pmatrix}$ $e_1 = \begin{pmatrix} 2.53217 \\ 1.28559 \\ 0.27349 \end{pmatrix}$ **< Scaled eigenvector e**

$\text{diag}(\sqrt{\phi}) \cdot e_1 = \begin{pmatrix} 0.84078 \\ 0.52102 \\ 0.55091 \end{pmatrix}$ **< Standardized discriminant coefficients**

2005 jw318

Prepared by:
Wm Stein

ORIGIN ≡ 1

HOTELLING'S T² PROFILE ANALYSIS FOR TWO GROUPS
jw 318.mcd

Example using Table 6.5 & 6.6 jw p. 323-324

Reading in Data:

Repeated measurement of Calcium at 4 yearly time intervals

$X_1 := READPRN("\backslash DATA\backslash T6\text{-}5.DAT")$

$X_2 := READPRN("\backslash DATA\backslash T6\text{-}6.DAT")$

$X_1 =$

	1	2	3	4
1	87.3	86.9	86.7	75.5
2	59	60.2	60	53.6
3	76.7	76.5	75.7	69.5
4	70.6	76.1	72.1	65.3
5	54.9	55.1	57.2	49
6	78.2	75.3	69.1	67.6
7	73.7	70.8	71.8	74.6
8	61.8	68.7	68.2	57.4
9	85.3	84.4	79.2	67
10	82.3	86.9	79.4	77.4
11	68.6	65.4	72.3	60.8
12	67.8	69.2	66.3	57.9
13	66.2	67	67	56.2
14	81	82.3	86.8	73.9
15	72.3	74.6	75.3	66.1

$X_2 =$

	1	2	3	4
1	83.8	85.5	86.2	81.2
2	65.3	66.9	67	60.6
3	81.2	79.5	84.5	75.2
4	75.4	76.7	74.3	66.7
5	55.3	58.3	59.1	54.2
6	70.3	72.3	70.6	68.6
7	76.5	79.9	80.4	71.6
8	66	70.9	70.3	64.1
9	76.7	79	76.9	70.3
10	77.2	74	77.8	67.9
11	67.3	70.7	68.9	65.9
12	50.3	51.4	53.6	48
13	57.7	57	57.5	51.5
14	74.3	77.7	72.6	68
15	74	74.7	74.5	65.7
16	57.3	56	64.7	53

Summary statistics of the sample:

$p := 4$

$n_1 := rows(X_1) \qquad n_1 = 15$

$n_2 := rows(X_2) \qquad n_2 = 16$

$i := 1..n_1 \qquad j := 1..n_2 \qquad k := 1..p$

$l_{n_i} := 1 \qquad ll_{n_j} := 1$

$I := identity(n_1) \qquad II := identity(n_2)$

$$X_{bar_1} := \frac{1}{n_1}\cdot X_1^T\cdot l_n \qquad X_{bar_1} = \begin{pmatrix} 72.38 \\ 73.293 \\ 72.473 \\ 64.787 \end{pmatrix}$$

$$X_{bar_2} := \frac{1}{n_2}\cdot X_2^T\cdot ll_n \qquad X_{bar_2} = \begin{pmatrix} 69.287 \\ 70.656 \\ 71.181 \\ 64.531 \end{pmatrix}$$

$$S_1 := \frac{1}{n_1 - 1}\cdot X_1^T\cdot\left(I - \frac{1}{n_1}\cdot l_n\cdot l_n^T\right)\cdot X_1 \qquad S_1 = \begin{pmatrix} 92.119 & 86.111 & 73.362 & 74.589 \\ 86.111 & 89.076 & 72.956 & 71.773 \\ 73.362 & 72.956 & 71.891 & 63.592 \\ 74.589 & 71.773 & 63.592 & 75.444 \end{pmatrix}$$

$$S_2 := \frac{1}{n_2 - 1}\cdot X_2^T\cdot\left(II - \frac{1}{n_2}\cdot ll_n\cdot ll_n^T\right)\cdot X_2 \qquad S_2 = \begin{pmatrix} 98.174 & 97.013 & 89.482 & 86.111 \\ 97.013 & 100.596 & 88.142 & 88.209 \\ 89.482 & 88.142 & 86.35 & 80.551 \\ 86.111 & 88.209 & 80.551 & 81.416 \end{pmatrix}$$

$$S_{pooled} := \frac{(n_1 - 1)\cdot S_1 + (n_2 - 1)\cdot S_2}{(n_1 + n_2 - 2)}$$

$$X_{GM} := \frac{n_1}{(n_1 + n_2)}\cdot X_{bar_1} + \frac{n_2}{(n_1 + n_2)}\cdot X_{bar_2}$$

$$X_{GM} = \begin{pmatrix} 70.784 \\ 71.932 \\ 71.806 \\ 64.655 \end{pmatrix} \qquad S_{pooled} = \begin{pmatrix} 95.251 & 91.75 & 81.7 & 80.549 \\ 91.75 & 95.035 & 80.811 & 80.275 \\ 81.7 & 80.811 & 79.369 & 72.364 \\ 80.549 & 80.275 & 72.364 & 78.533 \end{pmatrix}$$

Specify Contrast Matrix C:

$$C := \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix}$$

Testing orthogonality of the rows in C:

$CC := C^T$

$CC^{\langle 1\rangle T}\cdot CC^{\langle 2\rangle} = (-1)$

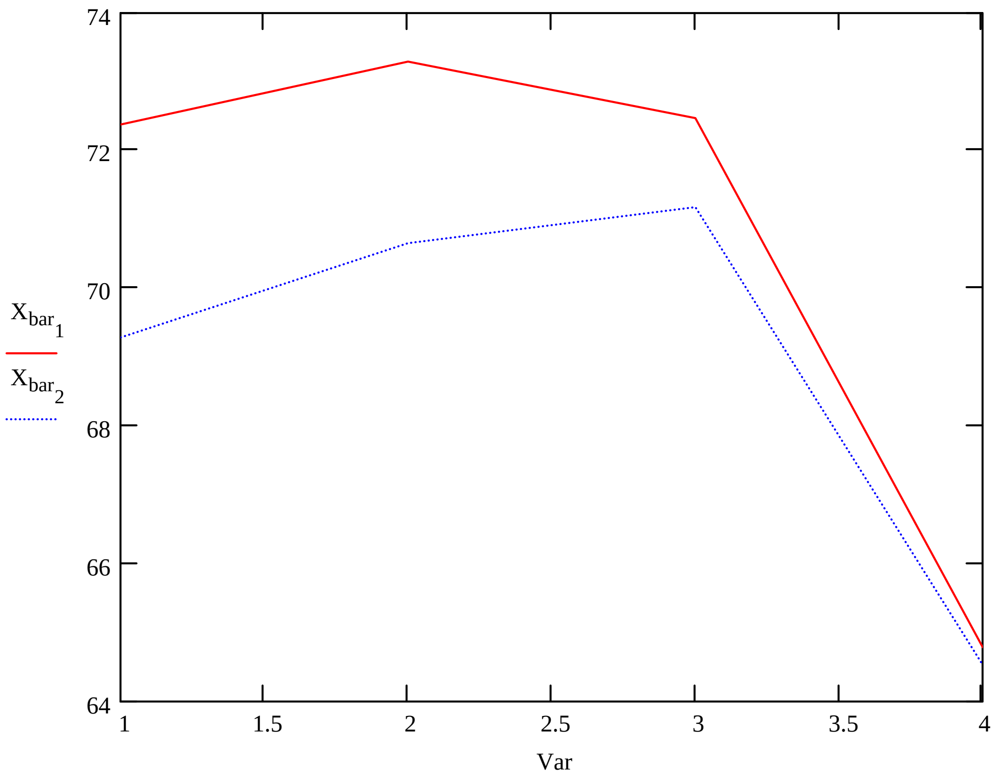
$CC^{\langle 1\rangle T}\cdot CC^{\langle 3\rangle} = (0)$

$CC^{\langle 2\rangle T}\cdot CC^{\langle 3\rangle} = (-1)$

< products are NOT all zero!!

$$Var := \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$$

2005 jw318



Linear combinations of mean difference vector (X_{bar1}-X_{bar2}) and variance-covariance matrix S_{pooled} with C:

$$C\cdot\left(X_{bar_1} - X_{bar_2}\right) = \begin{pmatrix} -0.455 \\ -1.345 \\ -1.037 \end{pmatrix} \qquad C\cdot S_{pooled}\cdot C^T = \begin{pmatrix} 6.78592 & -4.17426 & 0.61525 \\ -4.17426 & 12.78255 & -6.46952 \\ 0.61525 & -6.46952 & 13.17508 \end{pmatrix}$$

Test for PARALLEL Profiles - jw Equation 6-62 p. 319

Hotelling's T^2 statistic:

$$T_{sq} := \left[C\cdot\left(X_{bar_1} - X_{bar_2}\right)\right]^T\cdot\left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\cdot\left(C\cdot S_{pooled}\cdot C^T\right)\right]^{-1}\cdot\left[C\cdot\left(X_{bar_1} - X_{bar_2}\right)\right] \qquad T_{sq} = (5.387)$$

H_{01} : $C\mu_1 = C\mu_2$ **Assumption: Observations in both populations: X [X_1,X_2, ... , X_p] rs $N_p(\mu,\Sigma)$**

H_{11} : $C\mu_1$ <> $C\mu_2$

Stringency of the test: $\alpha := 0.05$ **< set as desired**

If assumptions hold and H_0 is true then:

$$c := \frac{\left(n_1 + n_2 - 2\right)\cdot(p-1)}{\left(n_1 + n_2 - p\right)}\cdot qF\left(1-\alpha, p-1, n_1 + n_2 - p\right) \qquad c = 9.539$$

< jw eq. 6-62 p. 319

NOTE: qF(1-α) is substituted here for F(α) in the text.

Decision Rule: Reject H_0 if T > c

$$\text{Decision} := \text{if}\left(T_{sq_1} > c, 1, 0\right) \qquad \text{Decision} = 0$$

< 0 = Do not reject H_{01} = Parallel Profiles
1 = Reject H_0

Test for COINCIDENT Profiles - jw Equation 6-63 p. 319 - *Only applicable when Profiles are Parallel*

Hotelling's T^2 statistic: $i := 1..p$ $\quad lll_{n_i} := 1$

$$T_{sq} := lll_n^T \cdot (X_{bar_1} - X_{bar_2}) \cdot \left[\left(\frac{1}{n_1} + \frac{1}{n_2} \right) \cdot \left(lll_n^T \cdot S_{pooled} \cdot lll_n \right) \right]^{-1} \cdot \left[lll_n \cdot (X_{bar_1} - X_{bar_2}) \right] \qquad T_{sq} = (0.31)$$

H_{02} : 1'μ_1 = 1'μ_2 **Assumption: Observations in both populations: X [X_1,X_2, ... , X_p] rs $N_p(\mu,\Sigma)$**

H_{12} : 1'μ_1 <> 1'μ_2

Stringency of the test: $\alpha := 0.05$ **< set as desired**

If assumptions hold and H_0 is true then:

< jw eq. 6-62 p. 319

$$c := qF(1 - \alpha, 1, n_1 + n_2 - 2) \qquad c = 4.183$$

NOTE: qF(1-α) is substituted here for F(α) in the text.

Decision Rule: Reject H_0 if T > c

$$\text{Decision} := \text{if}(T_{sq_1} > c, 1, 0) \qquad \text{Decision} = 0$$

< 0 = Do not reject H_0 = Coincident Profiles
1 = Reject H_0

Test for LEVEL Profiles - jw Equation 6-63 p. 319 - *Only applicable when Profiles are Parallel & Coincident*

$X := \text{stack}(X_1, X_2)$ $\quad N := n_1 + n_2$ $\quad i := 1..N$ $\quad l_{B_i} := 1$ $\quad I_B := \text{identity}(N)$

$$X_B := \frac{1}{N} \cdot X^T \cdot l_B$$

$$S := \frac{1}{N-1} \cdot X^T \cdot \left(I_B - \frac{1}{N} \cdot l_B \cdot l_B^T \right) \cdot X$$

$$X_B = \begin{pmatrix} 70.784 \\ 71.932 \\ 71.806 \\ 64.655 \end{pmatrix} \qquad S = \begin{pmatrix} 94.544 & 90.796 & 80.008 & 78.068 \\ 90.796 & 93.662 & 78.996 & 77.773 \\ 80.008 & 78.996 & 77.155 & 70.037 \\ 78.068 & 77.773 & 70.037 & 75.932 \end{pmatrix}$$

Hotelling's Test statisticT_{sq}:

$$T_{sq} := N \cdot (C \cdot X_{GM})^T \cdot (C \cdot S \cdot C^T) \cdot (C \cdot X_{GM}) \qquad T_{sq} = (20250.39)$$

H_{03} : Cμ = 0 **Assumption: All Observations X [X_1,X_2, ... , X_p] rs $N_p(\mu,\Sigma)$**

H_{13} : Cμ_1 <> 0

Stringency of the test: $\alpha := 0.05$ **< set as desired**

If assumptions hold and H_0 is true then:

< jw eq. 6-62 p. 319

$$c := \frac{(N-1) \cdot (p-1)}{(N-p+1)} \cdot qF(1 - \alpha, p-1, N-p+1) \qquad c = 9.471$$

NOTE: qF(1-α) is substituted here for F(α) in the text.

Decision Rule: Reject H_0 if T > c

$$\text{Decision} := \text{if}(T_{sq_1} > c, 1, 0) \qquad \text{Decision} = 1$$

< 0 = Do not reject H_0 = Level Profiles
1 = Reject H_0

2005 jw318A

$ORIGIN \equiv 1$

HOTELLING'S T^2 PROFILE ANALYSIS FOR TWO GROUPS
jw 318A.mcd

Prepared by:
Wm Stein

Verifying calculations in jw Example 6.12 p. 320 - Love & Marriage Survey data

Only Summary Results Given:

Four question variables:
x1 & x2 on a 8 point scale
x3 & x4 on a 5 point scale
were given to 30 wives & 30 husbands

$$X_{bar_1} := \begin{pmatrix} 6.833 \\ 7.033 \\ 3.967 \\ 4.700 \end{pmatrix}$$

< mean vector for husbands

$n_1 := 30$

$$S_{pooled} := \begin{pmatrix} .606 & .262 & .066 & .161 \\ .262 & .637 & .173 & .143 \\ .066 & .173 & .810 & .029 \\ .161 & .143 & .029 & .306 \end{pmatrix}$$

$$X_{bar_2} := \begin{pmatrix} 6.633 \\ 7.000 \\ 4.000 \\ 4.533 \end{pmatrix}$$

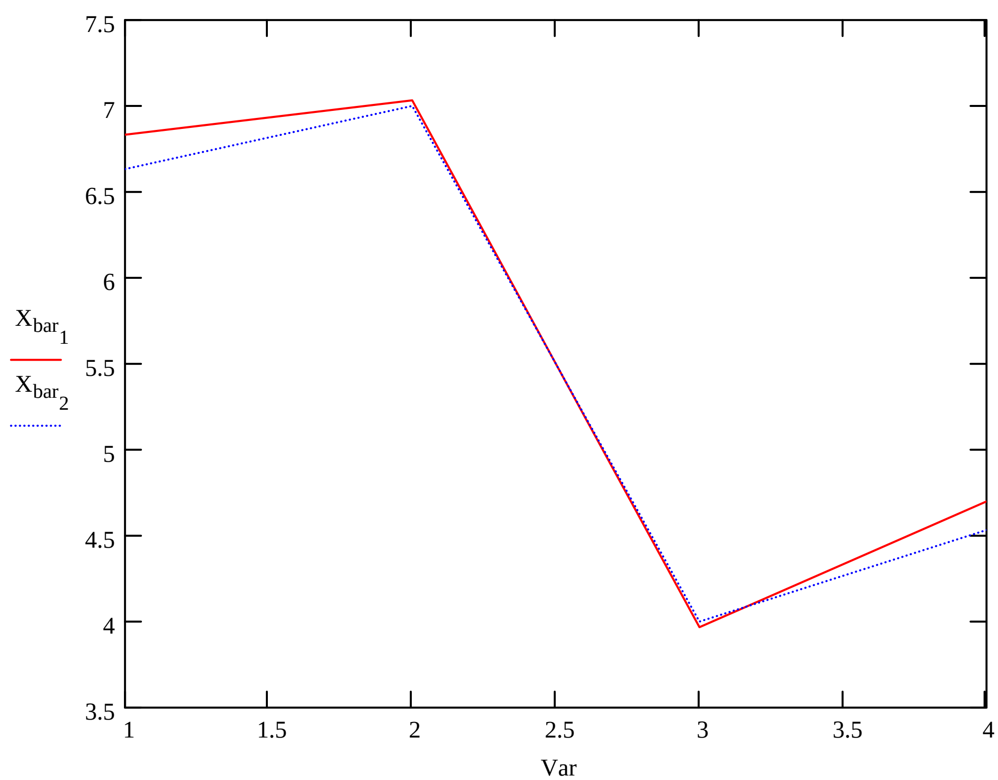
< mean vector for wives

$n_2 := 30$

$$Var := \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$$

$\mathbf{S_{pooled} = [(n_1\text{-}1)S_1+(n_2\text{-}1)S_2]/(n_1+n_2\text{-}2)}$

$p := 4$



Specify Contrast Matrix C:

$$C := \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix}$$

Testing orthogonality of the rows in C:

$CC := C^T$

$CC^{\langle 1 \rangle T} \cdot CC^{\langle 2 \rangle} = (-1)$

$CC^{\langle 1 \rangle T} \cdot CC^{\langle 3 \rangle} = (0)$

$CC^{\langle 2 \rangle T} \cdot CC^{\langle 3 \rangle} = (-1)$

< products are NOT all zero!!

Linear combinations of mean difference vector (X_{bar1}-X_{bar2}) and variance-covariance matrix S_{pooled} with C:

$$C\cdot\left(X_{bar_1} - X_{bar_2}\right) = \begin{pmatrix} -0.167 \\ -0.066 \\ 0.2 \end{pmatrix} \qquad C\cdot S_{pooled}\cdot C^T = \begin{pmatrix} 0.719 & -0.268 & -0.125 \\ -0.268 & 1.101 & -0.751 \\ -0.125 & -0.751 & 1.058 \end{pmatrix}$$

Test for PARALLEL Profiles - jw Equation 6-62 p. 319

Hotelling's T^2 statistic:

$$T_{sq} := \left[C\cdot\left(X_{bar_1} - X_{bar_2}\right)\right]^T\cdot\left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\cdot\left(C\cdot S_{pooled}\cdot C^T\right)\right]^{-1}\cdot\left[C\cdot\left(X_{bar_1} - X_{bar_2}\right)\right] \qquad T_{sq} = (1.005)$$

H_{01} : $C\mu_1 = C\mu_2$ **Assumption: Observations in both populations: X [X_1,X_2, ... , X_p] rs $N_p(\mu,\Sigma)$**

H_{11} : $C\mu_1$ <> $C\mu_2$

Stringency of the test: $\alpha := 0.05$ **< set as desired**

If assumptions hold and H_0 is true then:

$$c := \frac{\left(n_1 + n_2 - 2\right)\cdot(p - 1)}{\left(n_1 + n_2 - p\right)}\cdot qF\left(1 - \alpha, p - 1, n_1 + n_2 - p\right) \qquad c = 8.605$$

< jw eq. 6-62 p. 319

NOTE: qF(1-α) is substituted here for F(α) in the text.

Decision Rule: Reject H_0 if T > c

$$Decision := if\left(T_{sq_1} > c, 1, 0\right) \qquad Decision = 0$$

< 0 = Do not reject H_{01} = Parallel Profiles
1 = Reject H_0

Test for COINCIDENT Profiles - jw Equation 6-63 p. 319 - *Only applicable when Profiles are Parallel*

Hotelling's T^2 statistic: $i := 1 .. p$ $\quad l_{n_i} := 1$

$$T_{sq} := l_n{}^T\cdot\left(X_{bar_1} - X_{bar_2}\right)\cdot\left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\cdot\left(l_n{}^T\cdot S_{pooled}\cdot l_n\right)\right]^{-1}\cdot\left[l_n\cdot\left(X_{bar_1} - X_{bar_2}\right)\right] \qquad T_{sq} = (0.502)$$

H_{02} : $1'\mu_1 = 1'\mu_2$ **Assumption: Observations in both populations: X [X_1,X_2, ... , X_p] rs $N_p(\mu,\Sigma)$**

H_{12} : $1'\mu_1$ <> $1'\mu_2$

Stringency of the test: $\alpha := 0.05$ **< set as desired**

If assumptions hold and H_0 is true then:

$$c := qF\left(1 - \alpha, 1, n_1 + n_2 - 2\right) \qquad c = 4.007$$

< jw eq. 6-62 p. 319

NOTE: qF(1-α) is substituted here for F(α) in the text.

Decision Rule: Reject H_0 if T > c

$$Decision := if\left(T_{sq_1} > c, 1, 0\right) \qquad Decision = 0$$

< 0 = Do not reject H_0 = Coincident Profiles
1 = Reject H_0

2005 jw318B

ORIGIN ≡ 1

HOTELLING'S T^2 PROFILE ANALYSIS FOR TWO GROUPS
jw 318b.mcd

Prepared by:
Wm Stein

Verifying calculations using Table 5.1 in Rencher 4th ed. p. 141, 163-165

Reading in Data:

Observation of 4 psychological variables for 2 groups (males, females) with 19 individuals in each group.

Y := READPRN("\DATA\PSYCH.DAT")

$X_1 := \text{submatrix}(Y,1,32,2,5)$

$X_2 := \text{submatrix}(Y,33,64,2,5)$

$X_1 =$

	1	2	3	4
1	15	17	24	14
2	17	15	32	26
3	15	14	29	23
4	13	12	10	16
5	20	17	26	28
6	15	21	26	21
7	15	13	26	22
8	13	5	22	22
9	14	7	30	17
10	17	15	30	27
11	17	17	26	20
12	17	20	28	24
13	15	15	29	24
14	18	19	32	28
15	18	18	31	27

$X_2 =$

	1	2	3	4
1	13	14	12	21
2	14	12	14	26
3	12	19	21	21
4	12	13	10	16
5	11	20	16	16
6	12	9	14	18
7	10	13	18	24
8	10	8	13	23
9	12	20	19	23
10	11	10	11	27
11	12	18	25	25
12	14	18	13	26
13	14	10	25	28
14	13	16	8	14
15	14	8	13	25

Summary statistics of the sample:

$p := 4$

$n_1 := \text{rows}(X_1) \qquad n_1 = 32$

$n_2 := \text{rows}(X_2) \qquad n_2 = 32$

$i := 1..n_1 \qquad j := 1..n_2 \qquad k := 1..p$

$l_{n_i} := 1 \qquad ll_{n_j} := 1$

$I := \text{identity}(n_1) \qquad II := \text{identity}(n_2)$

$$X_{bar_1} := \frac{1}{n_1}\cdot X_1^T \cdot l_n \qquad X_{bar_1} = \begin{pmatrix} 15.969 \\ 15.906 \\ 27.188 \\ 22.75 \end{pmatrix}$$

$$X_{bar_2} := \frac{1}{n_2}\cdot X_2^T \cdot ll_n \qquad X_{bar_2} = \begin{pmatrix} 12.344 \\ 13.906 \\ 16.656 \\ 21.938 \end{pmatrix}$$

$$S_1 := \frac{1}{n_1 - 1}\cdot X_1^T \cdot \left(I - \frac{1}{n_1}\cdot l_n \cdot l_n^T \right) \cdot X_1 \qquad S_1 = \begin{pmatrix} 5.193 & 4.545 & 6.522 & 5.25 \\ 4.545 & 13.184 & 6.76 & 6.266 \\ 6.522 & 6.76 & 28.673 & 14.468 \\ 5.25 & 6.266 & 14.468 & 16.645 \end{pmatrix}$$

$$S_2 := \frac{1}{n_2 - 1}\cdot X_2^T \cdot \left(II - \frac{1}{n_2}\cdot ll_n \cdot ll_n^T \right) \cdot X_2 \qquad S_2 = \begin{pmatrix} 9.136 & 7.549 & 4.864 & 4.151 \\ 7.549 & 18.604 & 10.225 & 5.446 \\ 4.864 & 10.225 & 30.039 & 13.494 \\ 4.151 & 5.446 & 13.494 & 27.996 \end{pmatrix}$$

$$S_{pooled} := \frac{(n_1 - 1)\cdot S_1 + (n_2 - 1)\cdot S_2}{(n_1 + n_2 - 2)}$$

$$X_{GM} := \frac{n_1}{(n_1 + n_2)}\cdot X_{bar_1} + \frac{n_2}{(n_1 + n_2)}\cdot X_{bar_2}$$

$$X_{GM} = \begin{pmatrix} 14.156 \\ 14.906 \\ 21.922 \\ 22.344 \end{pmatrix} \qquad S_{pooled} = \begin{pmatrix} 7.164 & 6.047 & 5.693 & 4.701 \\ 6.047 & 15.894 & 8.492 & 5.856 \\ 5.693 & 8.492 & 29.356 & 13.981 \\ 4.701 & 5.856 & 13.981 & 22.321 \end{pmatrix}$$

Specify Contrast Matrix C:

$$C := \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix}$$

Testing orthogonality of the rows in C:

$CC := C^T$

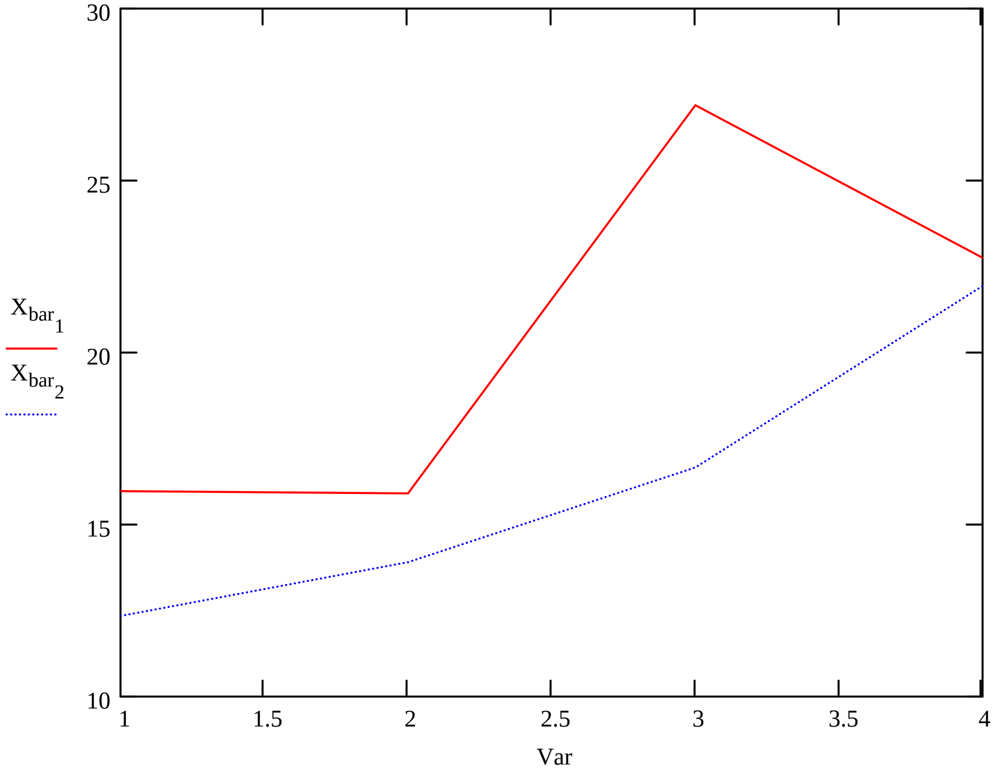
$CC^{\langle 1 \rangle T} \cdot CC^{\langle 2 \rangle} = (-1)$

$CC^{\langle 1 \rangle T} \cdot CC^{\langle 3 \rangle} = (0)$ **< products are NOT all zero!!**

$CC^{\langle 2 \rangle T} \cdot CC^{\langle 3 \rangle} = (-1)$

$$Var := \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$$

2005 jw318B



Linear combinations of mean difference vector (X_{bar1}-X_{bar2}) and variance-covariance matrix S_{pooled} with C:

$$C\cdot\left(X_{bar_1} - X_{bar_2}\right) = \begin{pmatrix} -1.625 \\ 8.531 \\ -9.719 \end{pmatrix} \qquad C\cdot S_{pooled}\cdot C^T = \begin{pmatrix} 10.96371 & -7.04738 & -1.64415 \\ -7.04738 & 28.26562 & -12.73891 \\ -1.64415 & -12.73891 & 23.71522 \end{pmatrix}$$

Test for PARALLEL Profiles - jw Equation 6-62 p. 319

Hotelling's T^2 statistic:

$$T_{sq} := \left[C\cdot\left(X_{bar_1} - X_{bar_2}\right)\right]^T\cdot\left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\cdot\left(C\cdot S_{pooled}\cdot C^T\right)\right]^{-1}\cdot\left[C\cdot\left(X_{bar_1} - X_{bar_2}\right)\right] \qquad T_{sq} = (74.24)$$

H_{01} : $C\mu_1 = C\mu_2$ Assumption: Observations in both populations: X [X_1,X_2, ... , X_p] rs $N_p(\mu,\Sigma)$

H_{11} : $C\mu_1$ <> $C\mu_2$

Stringency of the test: $\alpha := 0.05$ **< set as desired**

If assumptions hold and H_0 is true then:

$$c := \frac{\left(n_1 + n_2 - 2\right)\cdot(p - 1)}{\left(n_1 + n_2 - p\right)}\cdot qF\left(1 - \alpha, p - 1, n_1 + n_2 - p\right) \qquad c = 8.55$$

< jw eq. 6-62 p. 319

NOTE: qF(1-α) is substituted here for F(α) in the text.

Decision Rule: Reject H_0 if T > c

$$\text{Decision} := \text{if}\left(T_{sq_1} > c, 1, 0\right) \qquad \text{Decision} = 1$$

< 0 = Do not reject H_{01} = Parallel Profiles
1 = Reject H_0

Test for COINCIDENT Profiles - jw Equation 6-63 p. 319 - ***Usually applicable when Profiles are Parallel***

Hotelling's T^2 statistic: $i := 1 .. p$ $\quad lll_{n_i} := 1$

$$T_{sq} := lll_n^T \cdot \left(X_{bar_1} - X_{bar_2}\right) \cdot \left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \cdot \left(lll_n^T \cdot S_{pooled} \cdot lll_n\right)\right]^{-1} \cdot \left[lll_n \cdot \left(X_{bar_1} - X_{bar_2}\right)\right] \qquad T_{sq} = (28.044)$$

H_{02} : $1'\mu_1 = 1'\mu_2$ **Assumption: Observations in both populations: X [X_1,X_2, ... , X_p] rs $N_p(\mu,\Sigma)$**

H_{12} : $1'\mu_1$ <> $1'\mu_2$

Stringency of the test: $\alpha := 0.05$ **< set as desired**

If assumptions hold and H_0 is true then:

$$c := qF\left(1 - \alpha, 1, n_1 + n_2 - 2\right) \qquad c = 3.996$$

< jw eq. 6-62 p. 319

NOTE: qF(1-α) is substituted here for F(α) in the text.

Decision Rule: Reject H_0 if T > c

$$Decision := if\left(T_{sq_1} > c, 1, 0\right) \qquad Decision = 1$$

< 0 = Do not reject H_0 = Coincident Profiles
1 = Reject H_0

Test for LEVEL Profiles - jw Equation 6-63 p. 319 - ***Only applicable when Profiles are Parallel & Coincident***

$X := stack(X_1, X_2)$ $\quad N := n_1 + n_2$ $\quad i := 1 .. N$ $\quad l_{B_i} := 1$ $\quad I_B := identity(N)$

$$X_B := \frac{1}{N} \cdot X^T \cdot l_B$$

$$S := \frac{1}{N-1} \cdot X^T \cdot \left(I_B - \frac{1}{N} \cdot l_B \cdot l_B^T\right) \cdot X$$

$$X_B = \begin{pmatrix} 14.156 \\ 14.906 \\ 21.922 \\ 22.344 \end{pmatrix} \qquad S = \begin{pmatrix} 10.388 & 7.793 & 15.298 & 5.374 \\ 7.793 & 16.658 & 13.707 & 6.176 \\ 15.298 & 13.707 & 57.057 & 15.932 \\ 5.374 & 6.176 & 15.932 & 22.134 \end{pmatrix}$$

Hotelling's Test statisticT_{sq}:

$$T_{sq} := N \cdot \left(C \cdot X_B\right)^T \cdot \left(C \cdot S \cdot C^T\right)^{-1} \cdot \left(C \cdot X_B\right) \qquad T_{sq} = (256.362)$$

H_{03} : $C\mu = 0$ **Assumption: All Observations X [X_1,X_2, ... , X_p] rs $N_p(\mu,\Sigma)$**

H_{13} : $C\mu_1$ <> 0

Stringency of the test: $\alpha := 0.05$ **< set as desired**

If assumptions hold and H_0 is true then:

$$c := \frac{(N-1) \cdot (p-1)}{(N-p+1)} \cdot qF\left(1 - \alpha, p-1, N-p+1\right) \qquad c = 8.537$$

< jw eq. 6-62 p. 319

NOTE: qF(1-α) is substituted here for F(α) in the text.

Decision Rule: Reject H_0 if T > c

$$Decision := if\left(T_{sq_1} > c, 1, 0\right) \qquad Decision = 1$$

< 0 = Do not reject H_0 = Level Profiles
1 = Reject H_0

2005 Hotellings Conversion

ORIGIN ≡ 1 **CONVERSION TO HOTELLING'S T^2 FROM TYPICAL COMPUTER OUTPUT OF MANOVA** **Prepared by: Wm Stein**

Lizard Mass & SVL Original data transformed to natural logs jw Table 6.7 p. 330

Reading the Data:

$X_1 := \text{READPRN}(\text{"\DATA\T6-7Clogs.txt"})$ $p := \text{cols}(X_1)$

$X_2 := \text{READPRN}(\text{"\DATA\T6-7Slogs.txt"})$ $g := 2$

$X_1 =$

	1	2
1	2.01663	4.30407
2	1.61582	4.24133
3	1.76934	4.27667
4	2.40586	4.38203
5	0.88335	4.02535
6	2.6108	4.54329
7	2.904	4.55913
8	2.82328	4.60016
9	2.76695	4.57471
10	2.83527	4.50535
11	2.80493	4.51086
12	1.51072	4.20469
13	1.97824	4.31749
14	1.64866	4.24133
15	2.59898	4.51634
16	2.64476	4.51086
17	2.68546	4.49981
18	1.80698	4.29046
19	1.66089	4.24133
20	2.82743	4.54329

=

	1	2
1	2.63268	4.34381
2	1.65556	4.12713
3	3.61982	4.68213
4	3.73244	4.74493
5	3.46558	4.66344
6	1.37675	4.02535
7	1.47408	4.10264
8	1.11449	3.95124
9	1.5765	4.09434
10	1.87564	4.15888
11	3.11839	4.56435
12	2.59092	4.37576
13	1.41318	4.01638
14	2.51519	4.31749
15	1.96291	4.16667
16	3.04818	4.47164
17	3.76094	4.69135
18	3.30325	4.56435
19	3.66102	4.70953
20	2.983	4.43675
21	2.68553	4.38203
22	1.56653	4.12713
23	1.61343	4.11904
24	1.6525	4.12713
25	1.73871	4.15888
26	1.91147	4.14313
27	2.30028	4.26268
28	2.17827	4.24133
29	2.25055	4.21213
30	2.05553	4.18965
31	1.89987	4.16667
32	2.48324	4.36945
33	2.80457	4.43082
34	2.61227	4.39445
35	2.6174	4.4128
36	2.33699	4.30407
37	2.06686	4.22683
38	2.2086	4.2485
39	2.58143	4.35028
40	2.28105	4.2485

Summary statistics of the sample:

$n_1 := \text{rows}(X_1)$ $n_2 := \text{rows}(X_2)$ $n = \begin{pmatrix} 20 \\ 40 \end{pmatrix}$

Mean Vectors & Variance-Covariance matrices:

$i := 1 .. n_1$ $ii := 1 .. n_2$ $j := 1 .. p$

$l_{n_i} := 1$ $I := \text{identity}(n_1)$

$ll_{n_{ii}} := 1$ $II := \text{identity}(n_2)$

$X_{bar_1} := \frac{1}{n_1} \cdot X_1^T \cdot l_n$ $X_{bar_1} = \begin{pmatrix} 2.23992 \\ 4.39443 \end{pmatrix}$

$X_{bar_2} := \frac{1}{n_2} \cdot X_2^T \cdot ll_n$ $X_{bar_2} = \begin{pmatrix} 2.36814 \\ 4.30809 \end{pmatrix}$

$d := X_{bar_1} - X_{bar_2}$ $d = \begin{pmatrix} -0.12822 \\ 0.08634 \end{pmatrix}$

$S_1 := \frac{1}{n_1 - 1} \cdot X_1^T \cdot \left(I - \frac{1}{n_1} \cdot l_n \cdot l_n^T \right) \cdot X_1$ $S_1 = \begin{pmatrix} 0.353 & 0.0942 \\ 0.0942 & 0.026 \end{pmatrix}$

$S_2 := \frac{1}{n_2 - 1} \cdot X_2^T \cdot \left(II - \frac{1}{n_2} \cdot ll_n \cdot ll_n^T \right) \cdot X_2$ $S_2 = \begin{pmatrix} 0.50684 & 0.14539 \\ 0.14539 & 0.04255 \end{pmatrix}$

Total Sample size:

$N := n_1 + n_2$ $N = 60$

Grand Mean:

$m := 1 .. g$

$X_{barGM} := \frac{1}{N} \cdot \left[\sum_m \left(n_m \cdot X_{bar_m} \right) \right]$ $X_{barGM} = \begin{pmatrix} 2.3254 \\ 4.33687 \end{pmatrix}$

Residual/Error/Within Matrix W:

$m := 1 .. g$

$W := \sum_m \left(n_m - 1 \right) \cdot S_m$ $W = \begin{pmatrix} 26.47486 & 7.45949 \\ 7.45949 & 2.1527 \end{pmatrix}$

Treatment Matrix B:

$m := 1 .. g$

$B := \sum_m n_m \cdot \left(X_{bar_m} - X_{barGM} \right) \cdot \left(X_{bar_m} - X_{barGM} \right)^T$ $B = \begin{pmatrix} 0.21921 & -0.1476 \\ -0.1476 & 0.09938 \end{pmatrix}$

2005 Hotellings Conversion

Conversion Table for deriving Hotelling's T^2 from typical computer output of MANOVA

from: A. C. Rencher 1995.
Methods of Multivariate Analysis.
Wiley Interscience, NY. p. 147.

Wilk's Lambda Test Statistic:

$$\Lambda_s := \frac{|W|}{|W + B|} \qquad \Lambda_s = 0.20266$$

$$T_{sq} := (n_1 + n_2 - 2)\cdot\frac{1 - \Lambda_s}{\Lambda_s} \qquad T_{sq} = 228.18983$$

Lawley-Hotelling Trace:

$$LHtr := tr\left(B\cdot W^{-1}\right) \qquad LHtr = 3.93431$$

$$T_{sq} := (n_1 + n_2 - 2)\cdot LHtr \qquad T_{sq} = 228.18983$$

Pillai trace:

$$Ptr := tr\left[B\cdot(B + W)^{-1}\right] \qquad Ptr = 0.79734$$

$$T_{sq} := (n_1 + n_2 - 2)\cdot\frac{Ptr}{1 - Ptr} \qquad T_{sq} = 228.18983$$

Note that the value of T^2 reported here corresponds to the T^2 value calculated in the TWO POPULATIONS case with small sample sizes and hypothesized equal Σ's.

In this case, the variance-covariance matrix is pooled and T^2 is calculated from it. (see jw283.mcd)

Calculating Pooled Variance-Covariance matrix (jw Eq 6-21):

$$S_{pooled} := \frac{(n_1 - 1)\cdot S_1 + (n_2 - 1)\cdot S_2}{(n_1 + n_2 - 2)} \qquad S_{pooled} = \begin{pmatrix} 0.46 & 0.13 \\ 0.13 & 0.04 \end{pmatrix}$$

Hotelling's T^2 statistic (jw Eq. 6-23):

$$\delta_o := \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

< Set different value for δ_o if a test off zero vector is desired.

$$T_{sq} := (d - \delta_o)^T\cdot\left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\cdot S_{pooled}\right]^{-1}\cdot(d - \delta_o) \qquad T_{sq} = (228.18983)$$

PS: I couldn't get the conversion of Roy's largest root (also reported by Rencher 1995) to work. Someday, I may look into this statistic in more detail, but one hardly needs it.

Prepared by:
Wm Stein

ORIGIN ≡ 1

COMPARING ANOVA AND REGRESSION
One-Way Case - jw294.mcd

Example 6.6 p. 294. Three independent samples are tested for differences in treatment effect:

Sample one: $P1 := \begin{pmatrix} 9 \\ 6 \\ 9 \end{pmatrix}$ $m_1 := \text{rows}(P1)$ $j_1 := 1..m_1$

Sample two: $P2 := \begin{pmatrix} 0 \\ 2 \end{pmatrix}$ $m_2 := \text{rows}(P2)$ $j_2 := 1..m_2$

Sample three: $P3 := \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}$ $m_3 := \text{rows}(P3)$ $j_3 := 1..m_3$

Calculating sample means:

$P := \text{stack}(P1, \text{stack}(P2, P3))$ **< P = Vector of all values**

$n := \text{rows}(P)$ $n = 8$ **< n = total number of observations**

$i := 1..n$ $g := 3$ **< g = number of independent samples**

$X_{bar} := \frac{1}{n} \cdot \sum_i P_i$ $X_{bar} = 4$ $M_i := X_{bar}$ **< M = Vector of overall mean**

$P1_{bar} := \frac{1}{m_1} \cdot \sum_{j_1} P1_{j_1}$ $P1_{bar} = 8$ $PB1_{j_1} := P1_{bar}$ **<PB_j's = Vectors of means for each sample**

$P2_{bar} := \frac{1}{m_2} \cdot \sum_{j_2} P2_{j_2}$ $P2_{bar} = 1$ $PB2_{j_2} := P2_{bar}$

$P3_{bar} := \frac{1}{m_3} \cdot \sum_{j_3} P3_{j_3}$ $P3_{bar} = 2$ $PB3_{j_3} := P3_{bar}$

$PB := \text{stack}(PB1, \text{stack}(PB2, PB3))$ **<PB = Single vector of sample means**

$TE := PB - M$ **<TE = Vector of treatment effect**

$RES := P - PB$ **<RES = Vector of residual**

$$PB = \begin{pmatrix} 8 \\ 8 \\ 8 \\ 1 \\ 1 \\ 2 \\ 2 \\ 2 \end{pmatrix}$$

ANOVA model decomposition:

$$P = \begin{pmatrix} 9 \\ 6 \\ 9 \\ 0 \\ 2 \\ 3 \\ 1 \\ 2 \end{pmatrix} \quad M = \begin{pmatrix} 4 \\ 4 \\ 4 \\ 4 \\ 4 \\ 4 \\ 4 \\ 4 \end{pmatrix} \quad TE = \begin{pmatrix} 4 \\ 4 \\ 4 \\ -3 \\ -3 \\ -2 \\ -2 \\ -2 \end{pmatrix} \quad RES = \begin{pmatrix} 1 \\ -2 \\ 1 \\ -1 \\ 1 \\ 1 \\ -1 \\ 0 \end{pmatrix} \quad M + TE + RES = \begin{pmatrix} 9 \\ 6 \\ 9 \\ 0 \\ 2 \\ 3 \\ 1 \\ 2 \end{pmatrix}$$

< Decomposition verified p. 295

ANOVA Table:

Source		SS		df		MS
Treatments:	$SS_{tr} := \left(TE^T \cdot TE\right)_1$	$SS_{tr} = 78$	$df_{tr} := g - 1$	$df_{tr} = 2$		$MS_{tr} := \frac{SS_{tr}}{df_{tr}}$
Residual: (Error)	$SS_{res} := \left(RES^T \cdot RES\right)_1$	$SS_{res} = 10$	$df_{res} := n - g$	$df_{res} = 5$		$MS_{res} := \frac{SS_{res}}{df_{res}}$
Total:	$SS_{cor} := \left[(P - M)^T \cdot (P - M)\right]_1$	$SS_{cor} = 88$	$df_{cor} := n - 1$	$df_{cor} = 7$		$MS_{cor} := \frac{SS_{cor}}{df_{cor}}$

< ANOVA table verified p. 319

Model:

$P_{ij} = \mu + \tau_i + \varepsilon_{ij}$

H_0: $\tau_1 = \tau_2 = \tau_3 = 0$

H_1: **not all τ's are 0**

Test:

$\alpha := 0.01$ **< Set probability level**

Relevant Statistics:

$F := \frac{MS_{tr}}{MS_{res}}$ $F = 19.5$ $qF\left(1 - \alpha, df_{tr}, df_{res}\right) = 13.274$ **< F & qF verified p. 298**

Decision Rule:

Reject H_0 if F > qF

NOW TREATING THE SAME PROBLEM AS A REGRESSION. Example 7.2 p. 357.

$Y := P$

$$Y = \begin{pmatrix} 9 \\ 6 \\ 9 \\ 0 \\ 2 \\ 3 \\ 1 \\ 2 \end{pmatrix} \qquad Z := \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix} \qquad Z^T \cdot Z = \begin{pmatrix} 8 & 3 & 2 & 3 \\ 3 & 3 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 3 & 0 & 0 & 3 \end{pmatrix}$$

^ Note that first column of (Z^TZ) is a linear combination of the other columns

Calculating β_{hat} the normal way:

$\beta_{hat} := \left(Z^T \cdot Z\right)^{-1} \cdot Z \cdot Y$ $\quad rank\left(Z^T \cdot Z\right) = 3$ **< (Z^TZ) is not of full rank - must use generalized inverse**

^ won't work because (Z^TZ) is not of full rank

Calculating the generalized inverse (ZZ_m) - Footnote p. 358 & Exercise 7.6:

$r := \text{rank}(Z^T \cdot Z) - 1 \qquad r = 2 \qquad k := 1 .. r+1$

$\lambda := \text{reverse}(\text{sort}(\text{eigenvals}(Z^T \cdot Z)))$

$e^{\langle k \rangle} := \text{eigenvec}(Z^T \cdot Z, \lambda_k)$

$$\lambda = \begin{pmatrix} 10.772 \\ 3 \\ 2.228 \\ 1.726 \times 10^{-15} \end{pmatrix} \quad \begin{pmatrix} 0.861 & 0 & 0.096 \\ 0.332 & -0.707 & -0.374 \\ 0.196 & 0 & 0.844 \\ 0.332 & 0.707 & -0.374 \end{pmatrix}$$

$$ZZ_m := \sum_k (\lambda_k)^{-1} \cdot e^{\langle k \rangle} \cdot e^{\langle k \rangle T} \qquad ZZ_m = \begin{pmatrix} 0.073 & 0.01 & 0.052 & 0.01 \\ 0.01 & 0.24 & -0.135 & -0.094 \\ 0.052 & -0.135 & 0.323 & -0.135 \\ 0.01 & -0.094 & -0.135 & 0.24 \end{pmatrix}$$

Verifying the generalized inverse:

$$Z^T \cdot Z = \begin{pmatrix} 8 & 3 & 2 & 3 \\ 3 & 3 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 3 & 0 & 0 & 3 \end{pmatrix} \qquad (Z^T \cdot Z) \cdot ZZ_m \cdot (Z^T \cdot Z) = \begin{pmatrix} 8 & 3 & 2 & 3 \\ 3 & 3 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 3 & 0 & 0 & 3 \end{pmatrix}$$

< According to the definition of generalized inverse given in Exercise 7.6 p. 418, these two matrices should be the same...

Back to the regression:

$\beta_{hat} := ZZ_m \cdot Z^T \cdot Y \qquad \beta_{hat} = \begin{pmatrix} 2.75 \\ 5.25 \\ -1.75 \\ -0.75 \end{pmatrix}$

$Y_{hat} := Z \cdot \beta_{hat}$

$\varepsilon_{hat} := Y - Y_{hat}$

$Y_{bar} := \text{mean}(Y)$

$$\beta Z_{slope} := \sum_{k=2}^{g+1} \beta_{hat_k} \cdot Z^{\langle k \rangle}$$

$\beta_{0_i} := \beta_{hat_1}$

$$Y_{hat} = \begin{pmatrix} 8 \\ 8 \\ 8 \\ 1 \\ 1 \\ 2 \\ 2 \\ 2 \end{pmatrix}$$

Regression model decomposition:

$$Y = \begin{pmatrix} 9 \\ 6 \\ 9 \\ 0 \\ 2 \\ 3 \\ 1 \\ 2 \end{pmatrix} \quad \beta_0 = \begin{pmatrix} 2.75 \\ 2.75 \\ 2.75 \\ 2.75 \\ 2.75 \\ 2.75 \\ 2.75 \\ 2.75 \end{pmatrix} \quad \beta Z_{slope} = \begin{pmatrix} 5.25 \\ 5.25 \\ 5.25 \\ -1.75 \\ -1.75 \\ -0.75 \\ -0.75 \\ -0.75 \end{pmatrix} \quad \varepsilon_{hat} = \begin{pmatrix} 1 \\ -2 \\ 1 \\ -1 \\ 1 \\ 1 \\ -1 \\ -1.776 \times 10^{-15} \end{pmatrix} \quad \beta_0 + \beta Z_{slope} + \varepsilon_{hat} = \begin{pmatrix} 9 \\ 6 \\ 9 \\ 0 \\ 2 \\ 3 \\ 1 \\ 2 \end{pmatrix}$$

2005 jw294

Regression sums of squares:

$Total_{ss} := Y^T \cdot Y - n \cdot Y_{bar}^2$ $\quad$ $Total_{ss} = (88)$

$Regression_{ss} := Y_{hat} \cdot Y_{hat} - n \cdot Y_{bar}^2$ $\quad$ $Regression_{ss} = 78$

$Error_{ss} := \varepsilon_{hat}^T \cdot \varepsilon_{hat}$ $\quad$ $Error_{ss} = (10)$

$R_{sq} := \frac{Regression_{ss}}{Total_{ss}}$ $\quad$ $R_{sq} = (0.886)$ $\quad$ **< Coefficient of Determination R^2**

ANOVA for the regression:

Source		SS	df		MS
Regression:	$SS_{reg} := Regression_{ss}$	$SS_{reg} = 78$	$df_{reg} := g - 1$	$df_{reg} = 2$	$MS_{reg} := \frac{SS_{reg}}{df_{reg}}$
Residual: (Error)	$SS_{res} := Error_{ss_1}$	$SS_{res} = 10$	$df_{res} := n - g$	$df_{res} = 5$	$MS_{res} := \frac{SS_{res}}{df_{res}}$
Total:	$SS_{tot} := Total_{ss_1}$	$SS_{tot} = 88$	$df_{tot} := n - 1$	$df_{tot} = 7$	$MS_{cor} := \frac{SS_{tot}}{df_{tot}}$

Model:

$\mathbf{Y_i = \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \beta_3 z_3 + \varepsilon_i}$

$\mathbf{H_0}$**:** $\beta_1 = \beta_2 = \beta_3 = 0$

$\mathbf{H_1}$**:** β_1 , β_2 , β_3 **(slopes) are not all 0**

Test:

$\alpha := 0.01$ $\quad$ **< Set probability level**

Relevant Statistics:

$F := \frac{MS_{reg}}{MS_{res}}$ $\quad$ $F = 19.5$ $\quad$ $qF(1 - \alpha, df_{reg}, df_{res}) = 13.274$

Decision Rule:

Reject H_0 if F > qF

Relationship between the ANOVA & Regression models:

ANOVA:

$$P_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

or

$$P_{ij} - \mu - \varepsilon_{ij} = \tau_i$$

$$P - M = \begin{pmatrix} 5 \\ 2 \\ 5 \\ -4 \\ -2 \\ -1 \\ -3 \\ -2 \end{pmatrix} \quad RES = \begin{pmatrix} 1 \\ -2 \\ 1 \\ -1 \\ 1 \\ 1 \\ -1 \\ 0 \end{pmatrix} \quad TE = \begin{pmatrix} 4 \\ 4 \\ 4 \\ -3 \\ -3 \\ -2 \\ -2 \\ -2 \end{pmatrix}$$

< Solving for treatment effects (τ_i) of mean-centered data (P_{ij} - μ)

Regression:

$$Y_i = \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \beta_3 z_3 + \varepsilon_i$$

or

$$Y_i - \mu - \varepsilon_i = \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \beta_3 z_3 - \mu$$

or

$$Y_i - \mu - \varepsilon_i = \beta_0 + \beta Z_{slope} - \mu$$

$$Y - M = \begin{pmatrix} 5 \\ 2 \\ 5 \\ -4 \\ -2 \\ -1 \\ -3 \\ -2 \end{pmatrix} \quad \varepsilon_{hat} = \begin{pmatrix} 1 \\ -2 \\ 1 \\ -1 \\ 1 \\ 1 \\ -1 \\ -1.776 \times 10^{-15} \end{pmatrix} \quad \beta_0 + \beta Z_{slope} - M = \begin{pmatrix} 4 \\ 4 \\ 4 \\ -3 \\ -3 \\ -2 \\ -2 \\ -2 \end{pmatrix}$$

< Regression components for mean-centered data (Y_i - μ) equal ANOVA treatment effects (τ)

2005 jw368

CLASSICAL LINEAR REGRESSION

jw368.mcd

Prepared by: Wm Stein

ORIGIN ≡ 1

Real estate data gathered from 20 homes in a Milwaukee, Wisconsin neighborhood:

- **z_1 = Total dwelling size (100 ft^2)**
- **z_2 = Assessed value ($ 1000)**
- **Y = Selling price ($ 1000).**

In this study, Y represents the single dependent variable to be predicted by the z's.

Read in data from JW Table 7.1, p. 368:

M := READPRN("DATA\t7-1.DAT")

$n := rows(M) \qquad n = 20$

$$M = \begin{pmatrix} 15.31 & 57.3 & 74.8 \\ 15.2 & 63.8 & 74 \\ 16.25 & 65.4 & 72.9 \\ 14.33 & 57 & 70 \\ 14.57 & 63.8 & 74.9 \\ 17.33 & 63.2 & 76 \\ 14.48 & 60.2 & 72 \\ 14.91 & 57.7 & 73.5 \\ 15.25 & 56.4 & 74.5 \\ 13.89 & 55.6 & 73.5 \\ 15.18 & 62.6 & 71.5 \\ 14.44 & 63.4 & 71 \\ 14.87 & 60.2 & 78.9 \\ 18.63 & 67.2 & 86.5 \\ 15.2 & 57.1 & 68 \\ 25.76 & 89.6 & 102 \\ 19.05 & 68.6 & 84 \\ 15.37 & 60.1 & 69 \\ 18.06 & 66.3 & 88 \\ 16.35 & 65.8 & 76 \end{pmatrix}$$

Constructing variables:

- Dependent variable (Y) is the third data column of M:

$Y := M^{\langle 3 \rangle}$

$$Y = \begin{pmatrix} 74.8 \\ 74 \\ 72.9 \\ 70 \\ 74.9 \\ 76 \\ 72 \\ 73.5 \\ 74.5 \\ 73.5 \\ 71.5 \\ 71 \\ 78.9 \\ 86.5 \\ 68 \\ 102 \\ 84 \\ 69 \\ 88 \\ 76 \end{pmatrix}$$

- Matrix (Z) consisting in independent variables and first column of 1's:

$i := 1 .. n \qquad l_i := 1$

$Z := augment\left(augment\left(l, M^{\langle 1 \rangle}\right), M^{\langle 2 \rangle}\right)$

$$Z = \begin{pmatrix} 1 & 15.31 & 57.3 \\ 1 & 15.2 & 63.8 \\ 1 & 16.25 & 65.4 \\ 1 & 14.33 & 57 \\ 1 & 14.57 & 63.8 \\ 1 & 17.33 & 63.2 \\ 1 & 14.48 & 60.2 \\ 1 & 14.91 & 57.7 \\ 1 & 15.25 & 56.4 \\ 1 & 13.89 & 55.6 \\ 1 & 15.18 & 62.6 \\ 1 & 14.44 & 63.4 \\ 1 & 14.87 & 60.2 \\ 1 & 18.63 & 67.2 \\ 1 & 15.2 & 57.1 \\ 1 & 25.76 & 89.6 \\ 1 & 19.05 & 68.6 \\ 1 & 15.37 & 60.1 \\ 1 & 18.06 & 66.3 \\ 1 & 16.35 & 65.8 \end{pmatrix}$$

$r := cols(Z) - 1 \qquad r = 2$

Calculating least-squares estimator:
(using jw Result 7.1 p. 358)

- Least squares estimate of coefficients [= β_{hat}]:

$\beta_{hat} := \left(Z^T \cdot Z\right)^{-1} \cdot Z^T \cdot Y$

$$\beta_{hat} = \begin{pmatrix} 30.967 \\ 2.634 \\ 0.045 \end{pmatrix}$$

^Values of β_{hat} verified jw p. 369

- Fitted values of Y [= Y_{hat}]:

$Y_{hat} := Z \cdot \beta_{hat}$

- Residuals [= $epsilon_{hat}$]:

$\varepsilon_{hat} := Y - Y_{hat}$

$$\varepsilon_{hat} = \begin{pmatrix} 0.912 \\ 0.108 \\ -3.831 \\ -1.293 \\ 2.668 \\ -3.476 \\ 0.167 \\ 0.647 \\ 0.81 \\ 3.429 \\ -2.285 \\ -0.872 \\ 6.04 \\ 3.418 \\ -5.589 \\ -0.877 \\ -0.251 \\ -5.173 \\ 6.46 \\ -1.012 \end{pmatrix} \qquad Y_{hat} = \begin{pmatrix} 73.888 \\ 73.892 \\ 76.731 \\ 71.293 \\ 72.232 \\ 79.476 \\ 71.833 \\ 72.853 \\ 73.69 \\ 70.071 \\ 73.785 \\ 71.872 \\ 72.86 \\ 83.082 \\ 73.589 \\ 102.877 \\ 84.251 \\ 74.173 \\ 81.54 \\ 77.012 \end{pmatrix}$$

Sums of Squares decomposition jw p. 360

- Total sum of squares about mean:

$Y_{bar} := mean(Y) \qquad Y_{bar} = 76.55$

$Total_{ss} := Y^T \cdot Y - n \cdot Y_{bar}^2 \qquad Total_{ss} = (1237.87)$

- Regression sum of squares:

$Regression_{ss} := Y_{hat}^T \cdot Y_{hat} - n \cdot Y_{bar}^2 \qquad Regression_{ss} = (1032.875)$

- Residual (error) sum of squares:

$Error_{ss} := \varepsilon_{hat}^T \cdot \varepsilon_{hat} \qquad Error_{ss} = (204.995)$

- Verify that $Total_{ss}$ = $Regression_{ss}$ + $Residual_{ss}$:

$Total_{ss} = (1237.87) \qquad Regression_{ss} + Error_{ss} = (1237.87)$

2005 jw368

**Calculating coefficient of determination [= R^2]
and Multiple Correlation Coeffient [R] jw Eq. 7-9, p. 361:**

$$R_{sq} := 1 - \frac{Error_{ss_1}}{Total_{ss_1}} \qquad R_{sq} = 0.834 \qquad R := \sqrt{R_{sq}} \qquad R = 0.913$$

Sampling distribution of least-squares estimate of β (i.e. β_{hat}):

Calculating s^2 where the expected value of $s^2 = \sigma^2$ (JW Result 7.2 p. 363)

$$s_{sq} := \left(\frac{Error_{ss}}{n - r - 1}\right)_1 \qquad s_{sq} = 12.059 \qquad \sqrt{s_{sq}} = 3.473$$ **< s verified jw p. 369**

Expected mean of β_{hat}: **Expected variance/covariance of β_{hat}:**

$$\beta_{hat} = \begin{pmatrix} 30.967 \\ 2.634 \\ 0.045 \end{pmatrix} \qquad SZ := s_{sq}\cdot\left(Z^T\cdot Z\right)^{-1} \qquad SZ = \begin{pmatrix} 62.12921 & 3.06804 & -1.76476 \\ 3.06804 & 0.61717 & -0.2074 \\ -1.76476 & -0.2074 & 0.08133 \end{pmatrix}$$

$$j := 1 .. cols(Z) \qquad SE_j := \sqrt{SZ_{j,j}} \qquad SE = \begin{pmatrix} 7.882 \\ 0.786 \\ 0.285 \end{pmatrix}$$ **< Standard errors verified jw p. 369**

The multivariate confidence ellipsoid (jw Result 7.5 p. 367):

$$ZZ := \left(Z^T\cdot Z\right)^{-1}$$

$\lambda := eigenvals(ZZ)$

$ev := eigenvecs(ZZ)$

$$\lambda = \begin{pmatrix} 5.169 \\ 0.041 \\ 1.163\times 10^{-5} \end{pmatrix} \qquad ev = \begin{pmatrix} 0.99836 & -0.05526 & 0.01515 \\ 0.04973 & 0.96695 & 0.25007 \\ -0.02847 & -0.24891 & 0.96811 \end{pmatrix}$$

$r := rank(Z) - 1 \qquad r = 2$

$n := rows(M) \qquad n = 20$

$p := cols(M) \qquad p = 3$

Eigenvectors (ev) columns give directions of the multivariate confidence ellipsoid

$\alpha := 0.05$ **< Select probability of Type I error**

$$C := \sqrt{(r+1)\cdot s_{sq}\cdot qF(1-\alpha, r+1, n-r-1)} \qquad C = 10.754$$

$$i := 1 .. p \qquad L_i := C\cdot\sqrt{\lambda_i}$$

< Confidence ellipsoid constructed in analogy with that for T^2 intervals e.g. jw 5-19 p. 221

Multivariate simultaneous confidence ellipsoid:

$$\beta_{hat} = \begin{pmatrix} 30.967 \\ 2.634 \\ 0.045 \end{pmatrix}$$ **< Center of ellipsoid**

$$L = \begin{pmatrix} 24.45 \\ 2.179 \\ 0.037 \end{pmatrix}$$ **< L are half-lengths of the axes of the confidence ellipsoid for β_{hat} in the directions of ev**

2005 jw368

Simultaneous confidence intervals (jw Result 7.5 p. 367):

$$\text{var}\beta_{hat_i} := SZ_{i,i} \qquad \text{var}\beta_{hat} = \begin{pmatrix} 62.12921 \\ 0.61717 \\ 0.08133 \end{pmatrix}$$

$$CI_{lower} := \beta_{hat} - \sqrt{\text{var}\beta_{hat}} \cdot \sqrt{(r+1)\cdot qF(1-\alpha, r+1, n-r-1)}$$ **< jw result 7.5 p. 367**

$$CI_{upper} := \beta_{hat} + \sqrt{\text{var}\beta_{hat}} \cdot \sqrt{(r+1)\cdot qF(1-\alpha, r+1, n-r-1)}$$

$$CI := \text{augment}(CI_{lower}, CI_{upper})$$

$$\beta_{hat} = \begin{pmatrix} 30.967 \\ 2.634 \\ 0.045 \end{pmatrix}$$ **< Mean values**

$$CI = \begin{pmatrix} 6.557 & 55.376 \\ 0.202 & 5.067 \\ -0.838 & 0.928 \end{pmatrix}$$ **< Simultaneous confidence intervals**

Confidence intervals based on jw eq. 7-14 p. 368:

$$CI_{lower} := \beta_{hat} - qt\left(1 - \frac{\alpha}{2}, n-r-1\right) \cdot \sqrt{\text{var}\beta_{hat}}$$

$$CI_{upper} := \beta_{hat} + qt\left(1 - \frac{\alpha}{2}, n-r-1\right) \cdot \sqrt{\text{var}\beta_{hat}}$$

$$CI := \text{augment}(CI_{lower}, CI_{upper})$$

$$\beta_{hat} = \begin{pmatrix} 30.967 \\ 2.634 \\ 0.045 \end{pmatrix}$$ **< Mean values**

$$CI = \begin{pmatrix} 14.337 & 47.597 \\ 0.977 & 4.292 \\ -0.556 & 0.647 \end{pmatrix}$$ **< 'Bonferroni' intervals I guess, but not stated as such in jw**

^ Third row CI verified jw p. 370

2005 jw372

ORIGIN ≡ 1

ANOVA AS A REGRESSION
Two-Way Case - jw372.mcd

Prepared by:
Wm Stein

Example 7.5 p. 372. Male and female patrons rated the service in three establishments (locations) of a large restraurant chain. The service ratings were converted into an index.
Note the unbalanced design - Not possible wth a standard two-way ANOVA!

Reading data from separate disk files:

M := READPRN("DATA\t7-2.DAT")

Y := submatrix(M, 1, 18, 3, 3)

Z := READPRN("DATA\jw398z.txt")

Single dependent response variable

$$Y = \begin{pmatrix} 15.2 \\ 21.2 \\ 27.3 \\ 21.2 \\ 21.2 \\ 36.4 \\ 92.4 \\ 27.3 \\ 15.2 \\ 9.1 \\ 18.2 \\ 50 \\ 44 \\ 63.6 \\ 15.2 \\ 30.3 \\ 36.4 \\ 40.9 \end{pmatrix}$$

Independent dummy variables:
3 locations - cols 2-4
2 genders - cols 5-6
interactions - cols 7-12

$$Z = \begin{pmatrix} 1&1&0&0&1&0&1&0&0&0&0&0 \\ 1&1&0&0&1&0&1&0&0&0&0&0 \\ 1&1&0&0&1&0&1&0&0&0&0&0 \\ 1&1&0&0&1&0&1&0&0&0&0&0 \\ 1&1&0&0&1&0&1&0&0&0&0&0 \\ 1&1&0&0&0&1&0&1&0&0&0&0 \\ 1&1&0&0&0&1&0&1&0&0&0&0 \\ 1&0&1&0&1&0&0&0&1&0&0&0 \\ 1&0&1&0&1&0&0&0&1&0&0&0 \\ 1&0&1&0&1&0&0&0&1&0&0&0 \\ 1&0&1&0&1&0&0&0&1&0&0&0 \\ 1&0&1&0&1&0&0&0&1&0&0&0 \\ 1&0&1&0&0&1&0&0&0&1&0&0 \\ 1&0&1&0&0&1&0&0&0&1&0&0 \\ 1&0&0&1&1&0&0&0&0&0&1&0 \\ 1&0&0&1&1&0&0&0&0&0&1&0 \\ 1&0&0&1&0&1&0&0&0&0&0&1 \\ 1&0&0&1&0&1&0&0&0&0&0&1 \end{pmatrix}$$

Calculating β_{hat} the normal way:

$\beta := \left(Z^T \cdot Z\right)^{-1} \cdot Z^T \cdot Y$ < **(Z^TZ) is not of full rank must use generalized inverse**

Calculating the generalized inverse (ZZ_m):

$r := \text{rank}\left(Z^T \cdot Z\right) - 1 \qquad r = 5 \qquad k := 1 .. r + 1$

$\lambda := \text{reverse}\left(\text{sort}\left(\text{eigenvals}\left(Z^T \cdot Z\right)\right)\right)$

$e^{\langle k \rangle} := \text{eigenvec}\left(Z^T \cdot Z, \lambda_k\right)$

$$ZZ_m := \sum_k \left(\lambda_k\right)^{-1} \cdot e^{\langle k \rangle} \cdot e^{\langle k \rangle T}$$

Verifying the generalized inverse:

$$Z^T \cdot Z = \begin{pmatrix} 18&7&7&4&12&6&5&2&5&2&2&2 \\ 7&7&0&0&5&2&5&2&0&0&0&0 \\ 7&0&7&0&5&2&0&0&5&2&0&0 \\ 4&0&0&4&2&2&0&0&0&0&2&2 \\ 12&5&5&2&12&0&5&0&5&0&2&0 \\ 6&2&2&2&0&6&0&2&0&2&0&2 \\ 5&5&0&0&5&0&5&0&0&0&0&0 \\ 2&2&0&0&0&2&0&2&0&0&0&0 \\ 5&0&5&0&5&0&0&0&5&0&0&0 \\ 2&0&2&0&0&2&0&0&0&2&0&0 \\ 2&0&0&2&2&0&0&0&0&0&2&0 \\ 2&0&0&2&0&2&0&0&0&0&0&2 \end{pmatrix}$$

$$\left(Z^T \cdot Z\right) \cdot ZZ_m \cdot \left(Z^T \cdot Z\right) = \begin{pmatrix} 18&7&7&4&12&6&5&2&5&2&2&2 \\ 7&7&0&0&5&2&5&2&0&0&0&0 \\ 7&0&7&0&5&2&0&0&5&2&0&0 \\ 4&0&0&4&2&2&0&0&0&0&2&2 \\ 12&5&5&2&12&0&5&0&5&0&2&0 \\ 6&2&2&2&0&6&0&2&0&2&0&2 \\ 5&5&0&0&5&0&5&0&0&0&0&0 \\ 2&2&0&0&0&2&0&2&0&0&0&0 \\ 5&0&5&0&5&0&0&0&5&0&0&0 \\ 2&0&2&0&0&2&0&0&0&2&0&0 \\ 2&0&0&2&2&0&0&0&0&0&2&0 \\ 2&0&0&2&0&2&0&0&0&0&0&2 \end{pmatrix}$$

^ According to the definition of generalized inverse given in Exercise 7.6 p. 418, these two matrices should be the same.

Back to the regression:

$\beta_{hat} := ZZ_m \cdot Z^T \cdot Y \qquad i := 1 .. \text{rows}(Z)$

$Y_{hat} := Z \cdot \beta_{hat}$

$\varepsilon_{hat} := Y - Y_{hat} \qquad \beta 0_i := \beta_{hat_1}$

$Y_{bar} := \text{mean}(Y)$

2005 jw372

Regression Partials:

$$\beta Z := \sum_{k=2}^{4} \beta_{hat_k} \cdot Z^{\langle k \rangle}$$

< $\beta Z = \beta_1 Z_1 + \beta_2 Z_2 + \beta_3 Z_3$ where β_0 is intercept and Z_i refer to 2^{nd} - p columns in Z

$$\tau Z := \sum_{k=5}^{6} \beta_{hat_k} \cdot Z^{\langle k \rangle}$$

< $\tau Z = \beta_4 Z_4 + \beta_5 Z_5$

$$\gamma Z := \sum_{k=7}^{12} \beta_{hat_k} \cdot Z^{\langle k \rangle}$$

< $\gamma Z = \beta_6 Z_6 + \beta_7 Z_7 + \beta_8 Z_8 + \beta_9 Z_9 + \beta_{10} Z_{10} + \beta_{11} Z_{11}$

Regression model decomposition:

$$YY := \beta_0 + \beta Z + \tau Z + \gamma Z + \varepsilon_{hat}$$

Y =	β_0 =	βZ =	τZ =	γZ =	ε_{hat} =	YY =
15.2	18.732	9.808	–1.749	–5.571	–6.02	15.2
21.2	18.732	9.808	–1.749	–5.571	–0.02	21.2
27.3	18.732	9.808	–1.749	–5.571	6.08	27.3
21.2	18.732	9.808	–1.749	–5.571	–0.02	21.2
21.2	18.732	9.808	–1.749	–5.571	–0.02	21.2
36.4	18.732	9.808	20.481	15.379	–28	36.4
92.4	18.732	9.808	20.481	15.379	28	92.4
27.3	18.732	7.188	–1.749	–0.211	3.34	27.3
15.2	18.732	7.188	–1.749	–0.211	–8.76	15.2
9.1	18.732	7.188	–1.749	–0.211	–14.86	9.1
18.2	18.732	7.188	–1.749	–0.211	–5.76	18.2
50	18.732	7.188	–1.749	–0.211	26.04	50
44	18.732	7.188	20.481	7.399	–9.8	44
63.6	18.732	7.188	20.481	7.399	9.8	63.6
15.2	18.732	1.735	–1.749	4.032	–7.55	15.2
30.3	18.732	1.735	–1.749	4.032	7.55	30.3
36.4	18.732	1.735	20.481	–2.297	–2.25	36.4
40.9	18.732	1.735	20.481	–2.297	2.25	40.9

Regression sums of squares:

$n := \text{rows}(Y)$

$Total_{ss} := Y^T \cdot Y - n \cdot Y_{bar}^{2}$ $\quad Total_{ss} = (7186.609)$ $\quad SSZ_{tot} := Total_{ss_1}$

$Regression_{ss} := Y_{hat}^{T} \cdot Y_{hat} - n \cdot Y_{bar}^{2}$ $\quad Regression_{ss} = (4209.219)$ $\quad SSZ_{reg} := Regression_{ss_1}$

$Error_{ss} := \varepsilon_{hat}^{T} \cdot \varepsilon_{hat}$ $\quad Error_{ss} = (2977.39)$ $\quad SSZ_{res} := Error_{ss_1}$

^ confirmed p. 373

2005 jw372

Maximum Liklihood ratio regression tests for ANOVA parameters:

$Z_1 := \text{submatrix}(Z,1,18,1,6)$ **< Location & Gender (Interaction columns removed) for Z_1**

$Z_2 := \text{submatrix}(Z,1,18,1,4)$ **< Location (Interaction and Gender columns removed) for Z_2**

$Z_3 := \text{augment}(\text{submatrix}(Z,1,18,1,1),\text{submatrix}(Z,1,18,5,6))$ **< Gender (Interaction and Location columns removed) for Z_3**

$$Z_1 = \begin{pmatrix} 1&1&0&0&1&0\\ 1&1&0&0&1&0\\ 1&1&0&0&1&0\\ 1&1&0&0&1&0\\ 1&1&0&0&1&0\\ 1&1&0&0&0&1\\ 1&1&0&0&0&1\\ 1&0&1&0&1&0\\ 1&0&1&0&1&0\\ 1&0&1&0&1&0\\ 1&0&1&0&1&0\\ 1&0&1&0&1&0\\ 1&0&1&0&0&1\\ 1&0&1&0&0&1\\ 1&0&0&1&1&0\\ 1&0&0&1&1&0\\ 1&0&0&1&0&1\\ 1&0&0&1&0&1 \end{pmatrix} \quad Z_2 = \begin{pmatrix} 1&1&0&0\\ 1&1&0&0\\ 1&1&0&0\\ 1&1&0&0\\ 1&1&0&0\\ 1&1&0&0\\ 1&1&0&0\\ 1&0&1&0\\ 1&0&1&0\\ 1&0&1&0\\ 1&0&1&0\\ 1&0&1&0\\ 1&0&1&0\\ 1&0&1&0\\ 1&0&0&1\\ 1&0&0&1\\ 1&0&0&1\\ 1&0&0&1 \end{pmatrix} \quad Z_3 = \begin{pmatrix} 1&1&0\\ 1&1&0\\ 1&1&0\\ 1&1&0\\ 1&1&0\\ 1&0&1\\ 1&0&1\\ 1&1&0\\ 1&1&0\\ 1&1&0\\ 1&1&0\\ 1&1&0\\ 1&0&1\\ 1&0&1\\ 1&1&0\\ 1&1&0\\ 1&0&1\\ 1&0&1 \end{pmatrix}$$

Calculating 'extra' regressions using generalized inverses for each Z_i:

$\lambda 1 := \text{reverse}\left(\text{sort}\left(\text{eigenvals}\left(Z_1^T \cdot Z_1\right)\right)\right)$ $\quad j1 := \text{rows}(\lambda 1)$

$\lambda 2 := \text{reverse}\left(\text{sort}\left(\text{eigenvals}\left(Z_2^T \cdot Z_2\right)\right)\right)$ $\quad j2 := \text{rows}(\lambda 2)$

$\lambda 3 := \text{reverse}\left(\text{sort}\left(\text{eigenvals}\left(Z_3^T \cdot Z_3\right)\right)\right)$ $\quad j3 := \text{rows}(\lambda 3)$

$$\lambda 1 = \begin{pmatrix} 34.745\\ 7.96\\ 7\\ 4.296\\ 0\\ 0 \end{pmatrix} \quad \lambda 2 = \begin{pmatrix} 24.412\\ 7\\ 4.588\\ 0 \end{pmatrix} \quad \lambda 3 = \begin{pmatrix} 28.392\\ 7.608\\ 0 \end{pmatrix}$$

$$ZZ_{1m} := \sum_{i=1}^{\text{rank}\left(Z_1^T \cdot Z_1\right)} \left(\lambda 1_i\right)^{-1} \cdot \left(\text{eigenvec}\left(Z_1^T \cdot Z_1, \lambda 1_i\right)\right) \cdot \text{eigenvec}\left(Z_1^T \cdot Z_1, \lambda 1_i\right)^T$$

$$ZZ_{2m} := \sum_{i=1}^{\text{rank}\left(Z_2^T \cdot Z_2\right)} \left(\lambda 2_i\right)^{-1} \cdot \left(\text{eigenvec}\left(Z_2^T \cdot Z_2, \lambda 2_i\right)\right) \cdot \text{eigenvec}\left(Z_2^T \cdot Z_2, \lambda 2_i\right)^T$$

$$ZZ_{3m} := \sum_{i=1}^{\text{rank}\left(Z_3^T \cdot Z_3\right)} \left(\lambda 3_i\right)^{-1} \cdot \left(\text{eigenvec}\left(Z_3^T \cdot Z_3, \lambda 3_i\right)\right) \cdot \text{eigenvec}\left(Z_3^T \cdot Z_3, \lambda 3_i\right)^T$$

2005 jw372

Verifying generalized inverses:

$$Z_1^T \cdot Z_1 = \begin{pmatrix} 18 & 7 & 7 & 4 & 12 & 6 \\ 7 & 7 & 0 & 0 & 5 & 2 \\ 7 & 0 & 7 & 0 & 5 & 2 \\ 4 & 0 & 0 & 4 & 2 & 2 \\ 12 & 5 & 5 & 2 & 12 & 0 \\ 6 & 2 & 2 & 2 & 0 & 6 \end{pmatrix} \qquad \left(Z_1^T \cdot Z_1\right) \cdot ZZ_{1m} \cdot \left(Z_1^T \cdot Z_1\right) = \begin{pmatrix} 18 & 7 & 7 & 4 & 12 & 6 \\ 7 & 7 & 0 & 0 & 5 & 2 \\ 7 & 0 & 7 & 0 & 5 & 2 \\ 4 & 0 & 0 & 4 & 2 & 2 \\ 12 & 5 & 5 & 2 & 12 & 0 \\ 6 & 2 & 2 & 2 & 0 & 6 \end{pmatrix}$$

$$Z_2^T \cdot Z_2 = \begin{pmatrix} 18 & 7 & 7 & 4 \\ 7 & 7 & 0 & 0 \\ 7 & 0 & 7 & 0 \\ 4 & 0 & 0 & 4 \end{pmatrix} \qquad \left(Z_2^T \cdot Z_2\right) \cdot ZZ_{2m} \cdot \left(Z_2^T \cdot Z_2\right) = \begin{pmatrix} 18 & 7 & 7 & 4 \\ 7 & 7 & 0 & 0 \\ 7 & 0 & 7 & 0 \\ 4 & 0 & 0 & 4 \end{pmatrix}$$

$$Z_3^T \cdot Z_3 = \begin{pmatrix} 18 & 12 & 6 \\ 12 & 12 & 0 \\ 6 & 0 & 6 \end{pmatrix} \qquad \left(Z_3^T \cdot Z_3\right) \cdot ZZ_{3m} \cdot \left(Z_3^T \cdot Z_3\right) = \begin{pmatrix} 18 & 12 & 6 \\ 12 & 12 & 0 \\ 6 & 0 & 6 \end{pmatrix}$$

Calculating 'extra' residual sums of squares:

$\beta1_{hat} := \left(ZZ_{1m}\right) \cdot Z_1^T \cdot Y$ $\quad$ $\beta2_{hat} := \left(ZZ_{2m}\right) \cdot Z_2^T \cdot Y$ $\quad$ $\beta3_{hat} := \left(ZZ_{3m}\right) \cdot Z_3^T \cdot Y$

$Y1_{hat} := Z_1 \cdot \beta1_{hat}$ $\quad$ $Y2_{hat} := Z_2 \cdot \beta2_{hat}$ $\quad$ $Y3_{hat} := Z_3 \cdot \beta3_{hat}$

$\varepsilon1_{hat} := Y - Y1_{hat}$ $\quad$ $\varepsilon2_{hat} := Y - Y2_{hat}$ $\quad$ $\varepsilon3_{hat} := Y - Y3_{hat}$

$\varepsilon1_{hat}^T \cdot \varepsilon1_{hat} = (3419.147)$ $\quad$ $\varepsilon2_{hat}^T \cdot \varepsilon2_{hat} = (7165.826)$ $\quad$ $\varepsilon3_{hat}^T \cdot \varepsilon3_{hat} = (3666.165)$

$SSZ_{1.res} := \left(\varepsilon1_{hat}^T \cdot \varepsilon1_{hat}\right)_1$ $\quad$ $SSZ_{2.res} := \left(\varepsilon2_{hat}^T \cdot \varepsilon2_{hat}\right)_1$ $\quad$ $SSZ_{3.res} := \left(\varepsilon3_{hat}^T \cdot \varepsilon3_{hat}\right)_1$

^ $SSZ_{1.res}$ confirmed p. 399

Liklihood Statistics:

$r := \text{rank}(Z) - 1$

$q_1 := \text{rank}\left(Z_1\right) - 1$ $\quad$ $q_2 := \text{rank}\left(Z_2\right) - 1$ $\quad$ $q_3 := \text{rank}\left(Z_3\right) - 1$

$$F_1 := \frac{\dfrac{SSZ_{1.res} - SSZ_{res}}{\left(r - q_1\right)}}{\dfrac{SSZ_{res}}{(n - r - 1)}} \qquad F_2 := \frac{\dfrac{SSZ_{2.res} - SSZ_{res}}{\left(r - q_2\right)}}{\dfrac{SSZ_{res}}{(n - r - 1)}} \qquad F_3 := \frac{\dfrac{SSZ_{3.res} - SSZ_{res}}{\left(r - q_3\right)}}{\dfrac{SSZ_{res}}{(n - r - 1)}}$$

$F_1 = 0.89$ $\quad$ $F_2 = 5.627$ $\quad$ $F_3 = 0.694$

$qF\left(0.95, r - q_1, n - r - 1\right) = 3.885$ $\quad$ $qF\left(0.95, r - q_2, n - r - 1\right) = 3.49$ $\quad$ $qF\left(0.95, r - q_3, n - r - 1\right) = 3.259$

Decision Rule:

if $F < qF$ then factors of Z dropped from Z_i are not significant and may be ignored

^ Above results verified p. 373 - Only Gender factors in Z are significant.

2005 jw372A

COMPARING ANOVA AND REGRESSION
Two-Way Case - jw372A.mcd

Prepared by:
Wm Stein

ORIGIN ≡ 1

To proceed with a direct comparison of Regression & Two-way ANOVA, we need equal sample sizes in each treatment. So let's consider the peanut data - jw 4th EDITION Table 6.14 p. 375 (deleted in 5th edition).
Two factors (F1 & F2) are tested for effect on Yield (Y).

Reading data from separate disk files:

Y := READPRN("DATA\jw372ay.txt")

Z := READPRN("DATA\jw372az.txt")

$$Y = \begin{pmatrix} 195.3 \\ 194.3 \\ 189.7 \\ 180.4 \\ 203 \\ 195.9 \\ 202.7 \\ 197.6 \\ 193.5 \\ 187 \\ 201.5 \\ 200 \end{pmatrix} \qquad Z = \begin{pmatrix} -1 & 1 & 0 & -1 & 0 \\ -1 & 1 & 0 & -1 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 & -1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ -1 & -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 & -1 \end{pmatrix}$$

< "effect" coding Meyers & Well p. 547 provides a Z matrix of full rank, thus avoiding the generalized inverse

Calculating β_{hat} the normal way:

$r := \text{rank}(Z^T \cdot Z) - 1 \qquad r = 4 \qquad k := 1 .. r + 1$

$\beta_{hat} := (Z^T \cdot Z)^{-1} \cdot Z^T \cdot Y$

$Y_{hat} := Z \cdot \beta_{hat}$

$\varepsilon_{hat} := Y - Y_{hat} - \text{mean}(Y)$

$i := 1 .. \text{rows}(Z)$

$$\beta_{hat} = \begin{pmatrix} 0.24167 \\ -5.15 \\ 4.725 \\ -5.11667 \\ 0.10833 \end{pmatrix}$$

Regression partials:

$\beta_{0_i} := \text{mean}(Y) \qquad \text{mean}(Y) = 195.075$

$$\beta Z := \sum_{k=1}^{1} \beta_{hat_k} \cdot Z^{\langle k \rangle}$$

$$\tau Z := \sum_{k=2}^{3} \beta_{hat_k} \cdot Z^{\langle k \rangle}$$

$$\gamma Z := \sum_{k=4}^{5} \beta_{hat_k} \cdot Z^{\langle k \rangle}$$

$$Y_{hat} = \begin{pmatrix} -0.275 \\ -0.275 \\ -10.025 \\ -10.025 \\ 4.375 \\ 4.375 \\ 5.075 \\ 5.075 \\ -4.825 \\ -4.825 \\ 5.675 \\ 5.675 \end{pmatrix} \qquad Y = \begin{pmatrix} 195.3 \\ 194.3 \\ 189.7 \\ 180.4 \\ 203 \\ 195.9 \\ 202.7 \\ 197.6 \\ 193.5 \\ 187 \\ 201.5 \\ 200 \end{pmatrix} \qquad \text{mean}(Y) + Y_{hat} = \begin{pmatrix} 194.8 \\ 194.8 \\ 185.05 \\ 185.05 \\ 199.45 \\ 199.45 \\ 200.15 \\ 200.15 \\ 190.25 \\ 190.25 \\ 200.75 \\ 200.75 \end{pmatrix}$$

Regression model decomposition:

$YY := \beta_0 + \beta Z + \tau Z + \gamma Z + \varepsilon_{hat}$

$$Y = \begin{pmatrix} 195.3 \\ 194.3 \\ 189.7 \\ 180.4 \\ 203 \\ 195.9 \\ 202.7 \\ 197.6 \\ 193.5 \\ 187 \\ 201.5 \\ 200 \end{pmatrix} \quad \beta_0 = \begin{pmatrix} 195.075 \\ 195.075 \\ 195.075 \\ 195.075 \\ 195.075 \\ 195.075 \\ 195.075 \\ 195.075 \\ 195.075 \\ 195.075 \\ 195.075 \\ 195.075 \end{pmatrix} \quad \beta Z = \begin{pmatrix} -0.24167 \\ -0.24167 \\ 0.24167 \\ 0.24167 \\ -0.24167 \\ -0.24167 \\ 0.24167 \\ 0.24167 \\ -0.24167 \\ -0.24167 \\ 0.24167 \\ 0.24167 \end{pmatrix} \quad \tau Z = \begin{pmatrix} -5.15 \\ -5.15 \\ -5.15 \\ -5.15 \\ 4.725 \\ 4.725 \\ 4.725 \\ 4.725 \\ 0.425 \\ 0.425 \\ 0.425 \\ 0.425 \end{pmatrix} \quad \gamma Z = \begin{pmatrix} 5.11667 \\ 5.11667 \\ -5.11667 \\ -5.11667 \\ -0.10833 \\ -0.10833 \\ 0.10833 \\ 0.10833 \\ -5.00833 \\ -5.00833 \\ 5.00833 \\ 5.00833 \end{pmatrix} \quad \varepsilon_{hat} = \begin{pmatrix} 0.5 \\ -0.5 \\ 4.65 \\ -4.65 \\ 3.55 \\ -3.55 \\ 2.55 \\ -2.55 \\ 3.25 \\ -3.25 \\ 0.75 \\ -0.75 \end{pmatrix} \quad YY = \begin{pmatrix} 195.3 \\ 194.3 \\ 189.7 \\ 180.4 \\ 203 \\ 195.9 \\ 202.7 \\ 197.6 \\ 193.5 \\ 187 \\ 201.5 \\ 200 \end{pmatrix}$$

2005 jw372A

Regression sums of squares:

$n := \text{rows}(Y)$

$Total_{ss} := Y^T \cdot Y - n \cdot \text{mean}(Y)^2$ $\quad Total_{ss} = (506.1225)$ $\quad SSZ_{tot} := Total_{ss_1}$

$Regression_{ss} := Y_{hat}^T \cdot Y_{hat}$ $\quad Regression_{ss} = (401.9175)$ $\quad SSZ_{reg} := Regression_{ss_1}$

$Error_{ss} := \varepsilon_{hat}^T \cdot \varepsilon_{hat}$ $\quad Error_{ss} = (104.205)$ $\quad SSZ_{res} := Error_{ss_1}$

Maximum Liklihood ratio regression tests for ANOVA parameters:

$Z_1 := \text{READPRN}(\text{"DATA\\jw372az1.txt"})$ **F1 & F2 (Interaction columns removed) for Z_1**

$Z_2 := \text{READPRN}(\text{"DATA\\jw372az2.txt"})$ **F1 (Interaction and F2 columns removed) for Z_2**

$Z_3 := \text{READPRN}(\text{"DATA\\jw372az3.txt"})$ **F2 (Interaction and F1 columns removed) for Z_3**

$$Z_1 = \begin{pmatrix} -1 & 0 \\ -1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & -1 \\ 0 & -1 \\ 0 & 1 \\ 0 & 1 \\ 1 & 1 \\ 1 & 1 \\ -1 & -1 \\ -1 & -1 \end{pmatrix} \qquad Z_2 = \begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \\ -1 \\ -1 \\ 1 \\ 1 \\ -1 \\ -1 \\ 1 \\ 1 \end{pmatrix} \qquad Z_3 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ -1 & -1 \\ -1 & -1 \\ -1 & -1 \\ -1 & -1 \end{pmatrix}$$

Calculating 'extra' residual sums of squares:

$\beta1_{hat} := \left(Z_1^T \cdot Z_1\right)^{-1} \cdot Z_1^T \cdot Y$ $\quad \beta2_{hat} := \left(Z_2^T \cdot Z_2\right)^{-1} \cdot Z_2^T \cdot Y$ $\quad \beta3_{hat} := \left(Z_3^T \cdot Z_3\right)^{-1} \cdot Z_3^T \cdot Y$

$Y1_{hat} := Z_1 \cdot \beta1_{hat}$ $\quad Y2_{hat} := Z_2 \cdot \beta2_{hat}$ $\quad Y3_{hat} := Z_3 \cdot \beta3_{hat}$

$\varepsilon1_{hat} := Y - Y1_{hat} - \text{mean}(Y)$ $\quad \varepsilon2_{hat} := Y - Y2_{hat} - \text{mean}(Y)$ $\quad \varepsilon3_{hat} := Y - Y3_{hat} - \text{mean}(Y)$

$\varepsilon1_{hat}^T \cdot \varepsilon1_{hat} = (301.021)$ $\quad \varepsilon2_{hat}^T \cdot \varepsilon2_{hat} = (505.422)$ $\quad \varepsilon3_{hat}^T \cdot \varepsilon3_{hat} = (310.008)$

$SSZ_{1.res} := \left(\varepsilon1_{hat}^T \cdot \varepsilon1_{hat}\right)_1$ $\quad SSZ_{2.res} := \left(\varepsilon2_{hat}^T \cdot \varepsilon2_{hat}\right)_1$ $\quad SSZ_{3.res} := \left(\varepsilon3_{hat}^T \cdot \varepsilon3_{hat}\right)_1$

$\varepsilon_{hat}^T = (0.5 \;\; -0.5 \;\; 4.65 \;\; -4.65 \;\; 3.55 \;\; -3.55 \;\; 2.55 \;\; -2.55 \;\; 3.25 \;\; -3.25 \;\; 0.75 \;\; -0.75)$

$\varepsilon1_{hat}^T = (-4.89167 \;\; -5.89167 \;\; -0.25833 \;\; -9.55833 \;\; 8.03333 \;\; 0.93333 \;\; 7.51667 \;\; 2.41667 \;\; 3.43333 \;\; -3.06667 \;\; 1.41667 \;\; -0.08333)$

$\varepsilon2_{hat}^T = (0.46667 \;\; -0.53333 \;\; -5.61667 \;\; -14.91667 \;\; 8.16667 \;\; 1.06667 \;\; 7.38333 \;\; 2.28333 \;\; -1.33333 \;\; -7.83333 \;\; 6.18333 \;\; 4.68333)$

$\varepsilon3_{hat}^T = (5.375 \;\; 4.375 \;\; -0.225 \;\; -9.525 \;\; 3.2 \;\; -3.9 \;\; 2.9 \;\; -2.2 \;\; -2 \;\; -8.5 \;\; 6 \;\; 4.5)$

2005 jw372A

$$Y - \text{mean}(Y) = \begin{pmatrix} 0.225 \\ -0.775 \\ -5.375 \\ -14.675 \\ 7.925 \\ 0.825 \\ 7.625 \\ 2.525 \\ -1.575 \\ -8.075 \\ 6.425 \\ 4.925 \end{pmatrix} \quad Y_{hat} = \begin{pmatrix} -0.275 \\ -0.275 \\ -10.025 \\ -10.025 \\ 4.375 \\ 4.375 \\ 5.075 \\ 5.075 \\ -4.825 \\ -4.825 \\ 5.675 \\ 5.675 \end{pmatrix} \quad Y1_{hat} = \begin{pmatrix} 5.11667 \\ 5.11667 \\ -5.11667 \\ -5.11667 \\ -0.10833 \\ -0.10833 \\ 0.10833 \\ 0.10833 \\ -5.00833 \\ -5.00833 \\ 5.00833 \\ 5.00833 \end{pmatrix} \quad Y2_{hat} = \begin{pmatrix} -0.24167 \\ -0.24167 \\ 0.24167 \\ 0.24167 \\ -0.24167 \\ -0.24167 \\ 0.24167 \\ 0.24167 \\ -0.24167 \\ -0.24167 \\ 0.24167 \\ 0.24167 \end{pmatrix} \quad Y3_{hat} = \begin{pmatrix} -5.15 \\ -5.15 \\ -5.15 \\ -5.15 \\ 4.725 \\ 4.725 \\ 4.725 \\ 4.725 \\ 0.425 \\ 0.425 \\ 0.425 \\ 0.425 \end{pmatrix}$$

Liklihood Statistics:

$q_1 := \text{rank}(Z_1) \quad q_1 = 2 \qquad q_2 := \text{rank}(Z_2) \quad q_2 = 1 \qquad q_3 := \text{rank}(Z_3) \quad q_3 = 2$

$$F_1 := \frac{\dfrac{SSZ_{1.res} - SSZ_{res}}{(r - q_1)}}{\dfrac{SSZ_{res}}{(n-r-1)}} \qquad F_2 := \frac{\dfrac{SSZ_{2.res} - SSZ_{res}}{(r - q_2)}}{\dfrac{SSZ_{res}}{(n-r-1)}} \qquad F_3 := \frac{\dfrac{SSZ_{3.res} - SSZ_{res}}{(r - q_3)}}{\dfrac{SSZ_{res}}{(n-r-1)}}$$

$F_1 = 6.61058 \qquad F_2 = 8.98395 \qquad F_3 = 6.91242$

$qF(0.95, r - q_1, n - r - 1) = 4.73741 \qquad qF(0.95, r - q_2, n - r - 1) = 4.34683 \qquad qF(0.95, r - q_3, n - r - 1) = 4.73741$

^ Values have not been verified

Decision Rule:

if F < qF then factors of Z dropped from Z_i are not significant and may be ignored

Now performing a standard 2-Way ANOVA:

$g := 2$ **< levels of factor F1 (g)**

$i := 1 .. g$

$b := 3$ **< levels of factor F2 (b)**

$j := 1 .. b$

$n := 2$ **< number of observations in each bin**

$k := 1 .. n$

F1: \ **F2:**	**five**	**six**	**eight**
one	$X_{1,1} := \begin{pmatrix} 195.3 \\ 194.3 \end{pmatrix}$	$X_{1,2} := \begin{pmatrix} 203.0 \\ 195.9 \end{pmatrix}$	$X_{1,3} := \begin{pmatrix} 193.5 \\ 187.0 \end{pmatrix}$
two	$X_{2,1} := \begin{pmatrix} 189.7 \\ 180.4 \end{pmatrix}$	$X_{2,2} := \begin{pmatrix} 202.7 \\ 197.6 \end{pmatrix}$	$X_{2,3} := \begin{pmatrix} 201.5 \\ 200.0 \end{pmatrix}$

Bin means:

$$X_{bar_{i,j}} := \text{mean}(X_{i,j}) \qquad X_{bar} = \begin{pmatrix} 194.8 & 199.45 & 190.25 \\ 185.05 & 200.15 & 200.75 \end{pmatrix}$$

Grand mean:

$X_{GM} := \text{mean}(X_r) \qquad X_{GM} = 195.075$

$I := g \quad J := b \quad K := n$

Row and Column means:

$$X_{r_i} := \text{mean}\left[\left(X_{bar}^T\right)^{\langle i \rangle}\right] \qquad X_r = \begin{pmatrix} 194.83333 \\ 195.31667 \end{pmatrix}$$

$$X_{c_j} := \text{mean}\left(X_{bar}^{\langle j \rangle}\right) \qquad X_c = \begin{pmatrix} 189.925 \\ 199.8 \\ 195.5 \end{pmatrix}$$

2005 jw372A

AVOVA - 2-way:

Source	SS	df	MS
F1	$SS_r := \left[J \cdot K \cdot \left[\left(\left(X_r - X_{GM}\right)\right)^T \cdot \left(X_r - X_{GM}\right)\right]\right]_1$	$df_r := I - 1$	$MS_r := \frac{SS_r}{df_r}$
	$SS_r = 0.70083$	$df_r = 1$	$MS_r = 0.70083$
F2	$SS_c := \left[I \cdot K \cdot \left(X_c - X_{GM}\right)^T \cdot \left(X_c - X_{GM}\right)\right]_1$	$df_c := J - 1$	$MS_c := \frac{SS_c}{df_c}$
	$SS_c = 196.115$	$df_c = 2$	$MS_c = 98.0575$
Interactions	$SS_{int} := K \cdot \sum_{j=1}^{J} \sum_{i=1}^{I} \left(X_{bar_{i,j}} - X_{r_i} - X_{c_j} + X_{GM}\right)^2$	$df_{int} := (I - 1) \cdot (J - 1)$	$MS_{int} := \frac{SS_{int}}{df_{int}}$
	$SS_{int} = 205.10167$	$df_{int} = 2$	$MS_{int} = 102.55083$
Within	$SS_w := \sum_{k=1}^{K} \sum_{j=1}^{J} \sum_{i=1}^{I} \left[\left(X_{i,j}\right)_k - X_{bar_{i,j}}\right]^2$	$df_w := I \cdot J \cdot (K - 1)$	$MS_w := \frac{SS_w}{df_w}$
	$SS_w = 104.205$	$df_w = 6$	$MS_w = 17.3675$
Total	$SS_T := \sum_{k=1}^{K} \sum_{j=1}^{J} \sum_{i=1}^{I} \left[\left(X_{i,j}\right)_k - X_{GM}\right]^2$	$df_T := I \cdot J \cdot K - 1$	$MS_T := \frac{SS_T}{df_T}$
	$SS_T = 506.1225$	$df_T = 11$	$MS_T = 46.01114$

Signifiance level: $\alpha := 0.05$ **< Set desired level of tests**

^ Note that $SS_r + SS_c + SS_{int} = SSZ_{reg}$ above

Model:

$$X_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk}$$

Restrictions:

$$\Sigma\alpha_i = \Sigma\beta_j = \Sigma\gamma_{ij} = 0$$

Assumption:

$\varepsilon_{ijk} \sim N(0, \sigma^2)$ **Robust if K >12**

Hypothesis:	**Hypothesis:**	**Hypothesis:**
H_0: α_i's = 0	**H_0: β_j's = 0**	**H_0: γ_{ij}'s = 0**
H_1: There's a row effect (F1)	**H_1: There's a column effect (F2)**	**H_1: There's an interaction effect**
$F1 := \frac{MS_r}{MS_w}$ $F1 = 0.04035$	$F2 := \frac{MS_c}{MS_w}$ $F2 = 5.64603$	$F3 := \frac{MS_{int}}{MS_w}$ $F3 = 5.90476$
$qF(1 - \alpha, df_r, df_w) = 5.98738$	$qF(1 - \alpha, df_c, df_w) = 5.14325$	$qF(1 - \alpha, df_{int}, df_w) = 5.14325$
$1 - pF(F1, df_r, df_w) = 0.84743$	$1 - pF(F2, df_c, df_w) = 0.04177$	$1 - pF(F3, df_{int}, df_w) = 0.03824$

Decision rule:

Reject H_0 if F > qF in each test.

^ Values verified using Minitab

2005 jw372A

ANOVA (Two-Way) model decomposition:

$$X_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk}$$

Paperwork converting X into a column vector:

$ColA := stack(X_{1,1}, X_{2,1})$

$ColB := stack(X_{1,2}, X_{2,2})$

$ColC := stack(X_{1,3}, X_{2,3})$

$Y := stack(ColA, stack(ColB, ColC))$

$nn := 12 \quad i := 1..nn$

Calculating parameters:

$\mu_i := X_{GM}$ **< setting up a vector of mean**

$X_r - X_{GM} = \begin{pmatrix} -0.24167 \\ 0.24167 \end{pmatrix}$ **< use this to establish values for α**

$X_c - X_{GM} = \begin{pmatrix} -5.15 \\ 4.725 \\ 0.425 \end{pmatrix}$ **< use this to establish values for β**

$i := 1..n$

$\left(Xbar_{i,j} + X_{GM}\right) - X_{r_i} - X_{c_j} = \begin{pmatrix} 5.11667 \\ -5.11667 \\ -0.10833 \\ 0.10833 \\ -5.00833 \\ 5.00833 \end{pmatrix}$ **< use this to establish values for γ**

$X_{1,3} - Xbar_{1,3} = \begin{pmatrix} 3.25 \\ -3.25 \end{pmatrix}$ **< use this to establish values for ε**

$$Y = \begin{pmatrix} 195.3 \\ 194.3 \\ 189.7 \\ 180.4 \\ 203 \\ 195.9 \\ 202.7 \\ 197.6 \\ 193.5 \\ 187 \\ 201.5 \\ 200 \end{pmatrix} \quad \mu = \begin{pmatrix} 195.075 \\ 195.075 \\ 195.075 \\ 195.075 \\ 195.075 \\ 195.075 \\ 195.075 \\ 195.075 \\ 195.075 \\ 195.075 \\ 195.075 \\ 195.075 \end{pmatrix} \quad \alpha := \begin{pmatrix} -0.24167 \\ -0.24167 \\ 0.24167 \\ 0.24167 \\ -0.24167 \\ -0.24167 \\ 0.24167 \\ 0.24167 \\ -0.24167 \\ -0.24167 \\ 0.24167 \\ 0.24167 \end{pmatrix} \quad \beta := \begin{pmatrix} -5.15 \\ -5.15 \\ -5.15 \\ -5.15 \\ 4.725 \\ 4.725 \\ 4.725 \\ 4.725 \\ 0.425 \\ 0.425 \\ 0.425 \\ 0.425 \end{pmatrix} \quad \gamma := \begin{pmatrix} 5.11667 \\ 5.11667 \\ -5.11667 \\ -5.11667 \\ -0.10833 \\ -0.10833 \\ 0.10833 \\ 0.10833 \\ -5.00833 \\ -5.00833 \\ 5.00833 \\ 5.00833 \end{pmatrix} \quad \varepsilon := \begin{pmatrix} 0.5 \\ -0.5 \\ 4.65 \\ -4.65 \\ 3.55 \\ -3.55 \\ 2.55 \\ -2.55 \\ 3.25 \\ -3.25 \\ 0.75 \\ -0.75 \end{pmatrix} \quad \mu + \alpha + \beta + \gamma + \varepsilon = \begin{pmatrix} 195.3 \\ 194.3 \\ 189.7 \\ 180.4 \\ 203 \\ 195.9 \\ 202.7 \\ 197.6 \\ 193.5 \\ 187 \\ 201.5 \\ 200 \end{pmatrix}$$

$\alpha^T \cdot \alpha = (0.70085) \qquad \gamma^T \cdot \gamma = (205.10167) \qquad (Y - \mu)^T \cdot (Y - \mu) = (506.1225)$

$\beta^T \cdot \beta = (196.115) \qquad \varepsilon^T \cdot \varepsilon = (104.205)$

Relationship between ANOVA and Regression models:

Regression:

$$\mathbf{Y_i} - \mu = \Sigma\beta\mathbf{Z} + \Sigma\tau\mathbf{Z} + \Sigma\gamma\mathbf{Z} + \varepsilon_i$$
$$\mathbf{Y_i} - \mu - \varepsilon_i = \Sigma\beta\mathbf{Z} + \Sigma\tau\mathbf{Z} + \Sigma\gamma\mathbf{Z}$$

$$\beta Z = \begin{pmatrix} -0.24167 \\ -0.24167 \\ 0.24167 \\ 0.24167 \\ -0.24167 \\ -0.24167 \\ 0.24167 \\ 0.24167 \\ -0.24167 \\ -0.24167 \\ 0.24167 \\ 0.24167 \end{pmatrix} \quad \tau Z = \begin{pmatrix} -5.15 \\ -5.15 \\ -5.15 \\ -5.15 \\ 4.725 \\ 4.725 \\ 4.725 \\ 4.725 \\ 0.425 \\ 0.425 \\ 0.425 \\ 0.425 \end{pmatrix} \quad \gamma Z = \begin{pmatrix} 5.11667 \\ 5.11667 \\ -5.11667 \\ -5.11667 \\ -0.10833 \\ -0.10833 \\ 0.10833 \\ 0.10833 \\ -5.00833 \\ -5.00833 \\ 5.00833 \\ 5.00833 \end{pmatrix} \quad \beta Z + \tau Z + \gamma Z = \begin{pmatrix} -0.275 \\ -0.275 \\ -10.025 \\ -10.025 \\ 4.375 \\ 4.375 \\ 5.075 \\ 5.075 \\ -4.825 \\ -4.825 \\ 5.675 \\ 5.675 \end{pmatrix}$$

ANOVA (Two-Way):

$$\mathbf{X_{ijk}} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk}$$
$$\mathbf{X_{ijk}} - \mu - \varepsilon_{ijk} = \alpha_i + \beta_j + \gamma_{ij}$$

$$\alpha = \begin{pmatrix} -0.24167 \\ -0.24167 \\ 0.24167 \\ 0.24167 \\ -0.24167 \\ -0.24167 \\ 0.24167 \\ 0.24167 \\ -0.24167 \\ -0.24167 \\ 0.24167 \\ 0.24167 \end{pmatrix} \quad \beta = \begin{pmatrix} -5.15 \\ -5.15 \\ -5.15 \\ -5.15 \\ 4.725 \\ 4.725 \\ 4.725 \\ 4.725 \\ 0.425 \\ 0.425 \\ 0.425 \\ 0.425 \end{pmatrix} \quad \gamma = \begin{pmatrix} 5.11667 \\ 5.11667 \\ -5.11667 \\ -5.11667 \\ -0.10833 \\ -0.10833 \\ 0.10833 \\ 0.10833 \\ -5.00833 \\ -5.00833 \\ 5.00833 \\ 5.00833 \end{pmatrix} \quad \alpha + \beta + \gamma = \begin{pmatrix} -0.275 \\ -0.275 \\ -10.025 \\ -10.025 \\ 4.375 \\ 4.375 \\ 5.075 \\ 5.075 \\ -4.825 \\ -4.825 \\ 5.675 \\ 5.675 \end{pmatrix}$$

NOTE: Equivalent decomposition for Regression and ANOVA only works with a balanced (orthogonal) ANOVA design where there are the same number of observations in each treatment category. In an unbalanced design, α,β,γ are correlated and equivalence of the two procedures is more difficult to demonstrate - use Regression.

ORIGIN ≡ 1 **CLASSICAL LINEAR REGRESSION, CONFIDENCE INTERVALS & PREDICTION** **Prepared by: Wm Stein**

jw376.mcd

Verifying calculations in Example 7.6

Companies considering the purchase of a computer must first assess their future needs in order to determine the proper equipment. A computer scientist collected data from seven similar company sites so that a forcast equation of computer-hardware requirements for inventory management could be developed:

- **z_1 = customer orders in thousands**
- **z_2 = add-delete item counts in thousands**
- **Y = CPU (central processing unit) time in hours.**

In this study, Y represents the single dependent variable to be predicted by the z's.

Read in data from JW Table 7.3, p. 376:

ORIGIN ≡ 1

M := READPRN("DATA\jw376.dta")

n := rows(M) n = 7

M =

z_1	z_2	Y
123.5	2.108	141.5
146.1	9.213	168.9
133.9	1.905	154.8
128.5	0.815	146.5
151.5	1.061	172.8
136.2	8.603	160.1
92	1.125	108.5

Constructing variables:

- Dependent variable (Y) is the third data column of M:

$Y := M^{\langle 3 \rangle}$

- Matrix (Z) consisting in independent variables and first column of 1's:

i := 1 .. n

$l_i := 1$

$Z := \text{augment}\left(\text{augment}\left(l, M^{\langle 1 \rangle}\right), M^{\langle 2 \rangle}\right)$

r := cols(Z) − 1 r = 2

$$Y = \begin{pmatrix} 141.5 \\ 168.9 \\ 154.8 \\ 146.5 \\ 172.8 \\ 160.1 \\ 108.5 \end{pmatrix} \quad l = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad Z = \begin{pmatrix} 1 & 123.5 & 2.108 \\ 1 & 146.1 & 9.213 \\ 1 & 133.9 & 1.905 \\ 1 & 128.5 & 0.815 \\ 1 & 151.5 & 1.061 \\ 1 & 136.2 & 8.603 \\ 1 & 92 & 1.125 \end{pmatrix}$$

Calculating least-squares estimator:
(using jw Result 7.1 p. 358)

- Least squares estimate of coefficients [= β_{hat}]:

$\beta_{hat} := \left(Z^T \cdot Z\right)^{-1} \cdot Z^T \cdot Y$

$$\beta_{hat} = \begin{pmatrix} 8.424 \\ 1.079 \\ 0.42 \end{pmatrix}$$

< Regression function verified p. 376

- Fitted values of Y [= Y_{hat}]:

$Y_{hat} := Z \cdot \beta_{hat}$

$$Y_{hat} = \begin{pmatrix} 142.563 \\ 169.931 \\ 153.699 \\ 147.415 \\ 172.335 \\ 158.993 \\ 108.162 \end{pmatrix}$$

- Residuals [= epsilon_{hat}]:

$\varepsilon_{hat} := Y - Y_{hat}$

$$\varepsilon_{hat} = \begin{pmatrix} -1.063 \\ -1.031 \\ 1.101 \\ -0.915 \\ 0.465 \\ 1.107 \\ 0.338 \end{pmatrix}$$

$$\left(Z^T \cdot Z\right)^{-1} = \begin{pmatrix} 8.1797 & -0.06411 & 0.08831 \\ -0.06411 & 0.00052 & -0.00107 \\ 0.08831 & -0.00107 & 0.0144 \end{pmatrix}$$

^ verified jw p. 376

2005 jw376

Sums of Squares decomposition:

- Total sum of squares about mean:

$Y_{bar} := \text{mean}(Y)$ $\qquad Y_{bar} = 150.443$

$Total_{ss} := Y^T \cdot Y - n \cdot Y_{bar}^2$ $\qquad Total_{ss} = (2807.477)$

- Regression sum of squares:

$Regression_{ss} := Y_{hat}^T \cdot Y_{hat} - n \cdot Y_{bar}^2$ $\qquad Regression_{ss} = (2801.679)$

- Residual (error) sum of squares:

$Error_{ss} := \varepsilon_{hat}^T \cdot \varepsilon_{hat}$ $\qquad Error_{ss} = (5.798)$

- Verify that $Total_{ss}$ = $Regression_{ss}$ + $Residual_{ss}$:

$Total_{ss} = \left(2.807 \times 10^3\right)$ $\qquad Regression_{ss} + Error_{ss} = (2807.477)$

Calculating coefficient of determination [= R^2]
and Multiple Correlation Coeffient [R] jw Eq. 7-9, p. 361

$R_{sq} := 1 - \frac{Error_{ss_1}}{Total_{ss_1}}$ $\qquad R_{sq} = 0.998$ $\qquad R := \sqrt{R_{sq}}$ $\qquad R = 0.999$

Sampling distribution of least-squares estimate of β (i.e. β_{hat}):

Calculating s^2 where the expected value of $s^2 = \sigma^2$ (jw Result 7.2 p. 363)

$s_{sq} := \left(\frac{Error_{ss}}{n - r - 1}\right)_1$ $\qquad s_{sq} = 1.449$ $\qquad s := \sqrt{s_{sq}}$ $\qquad s = 1.204$ $\qquad$ **< s verified p. 377**

Expected mean of β_{hat}:

$$\beta_{hat} = \begin{pmatrix} 8.424 \\ 1.079 \\ 0.42 \end{pmatrix}$$

Expected variance/covariance of β_{hat}:

$$SZ := s_{sq} \cdot \left(Z^T \cdot Z\right)^{-1} \qquad SZ = \begin{pmatrix} 11.85617 & -0.09293 & 0.12801 \\ -0.09293 & 0.00076 & -0.00155 \\ 0.12801 & -0.00155 & 0.02087 \end{pmatrix}$$

The multivariate confidence ellipsoid (jw Result 7.5 p. 367):

$$ZZ := \left(Z^T \cdot Z\right)^{-1} \qquad ZZ = \begin{pmatrix} 8.1797 & -0.06411 & 0.08831 \\ -0.06411 & 0.00052 & -0.00107 \\ 0.08831 & -0.00107 & 0.0144 \end{pmatrix}$$

$\lambda := \text{eigenvals}(ZZ)$

$ev := \text{eigenvecs}(ZZ)$

$$\lambda = \begin{pmatrix} 8.181 \\ 8.257 \times 10^{-6} \\ 0.013 \end{pmatrix}$$

$$ev = \begin{pmatrix} 0.99991 & 0.00753 & -0.01103 \\ -0.00784 & 0.99958 & -0.02804 \\ 0.01081 & 0.02812 & 0.99955 \end{pmatrix}$$

$r := \text{rank}(Z) - 1$ $\qquad r = 2$

$n := \text{rows}(M)$ $\qquad n = 7$

$p := \text{cols}(M)$ $\qquad p = 3$

Eigenvectors (ev) columns give directions of the multivariate confidence ellipsoid

$\alpha := 0.05$ **< Select probability of Type I error**

$C := \sqrt{(r+1)\cdot s_{sq}\cdot qF(1-\alpha, r+1, n-r-1)}$ $C = 5.354$ **< Confidence ellipsoid constructed in analogy with that for T^2 intervals e.g. jw 5-19 p. 221**

Multivariate simultaneous confidence ellipsoid:

$i := 1..p$ $L_i := C\cdot\sqrt{\lambda_i}$

$\beta_{hat} = \begin{pmatrix} 8.424 \\ 1.079 \\ 0.42 \end{pmatrix}$ **< Center of ellipsoid**

$L = \begin{pmatrix} 15.313 \\ 0.015 \\ 0.621 \end{pmatrix}$ **< L are half-lengths of the axes of the confidence ellipsoid for β_{hat} in the directions of ev**

Simultaneous confidence intervals (jw Result 7.5 p. 367):

$var\beta_{hat_i} := SZ_{i,i}$ $var\beta_{hat} = \begin{pmatrix} 11.85617 \\ 0.00076 \\ 0.02087 \end{pmatrix}$

$CI_{lower} := \beta_{hat} - \sqrt{var\beta_{hat}}\cdot\sqrt{(r+1)\cdot qF(1-\alpha, r+1, n-r-1)}$

$CI_{upper} := \beta_{hat} + \sqrt{var\beta_{hat}}\cdot\sqrt{(r+1)\cdot qF(1-\alpha, r+1, n-r-1)}$

$CI := augment(CI_{lower}, CI_{upper})$

$\beta_{hat} = \begin{pmatrix} 8.424 \\ 1.079 \\ 0.42 \end{pmatrix}$ **< Mean values** $CI = \begin{pmatrix} -6.888 & 23.735 \\ 0.957 & 1.201 \\ -0.223 & 1.062 \end{pmatrix}$ **< Confidence interval**

Confidence intervals based on jw eq. 7-14 p. 368:

$CI_{lower} := \beta_{hat} - qt\left(1-\frac{\alpha}{2}, n-r-1\right)\cdot\sqrt{var\beta_{hat}}$

$CI_{upper} := \beta_{hat} + qt\left(1-\frac{\alpha}{2}, n-r-1\right)\cdot\sqrt{var\beta_{hat}}$

$CI := augment(CI_{lower}, CI_{upper})$

$\beta_{hat} = \begin{pmatrix} 8.424 \\ 1.079 \\ 0.42 \end{pmatrix}$ **< Mean values** $CI = \begin{pmatrix} -1.136 & 17.984 \\ 1.003 & 1.155 \\ 0.019 & 0.821 \end{pmatrix}$ **< 'Bonferroni' intervals (I guess, but not stated as such in jw)**

Estimating the regression function at a specified z_o:

$$z_0 := \begin{pmatrix} 1 \\ 130 \\ 7.5 \end{pmatrix}$$ **< Value given jw p. 376**

Expected value of Y (Y_o) given z_o:

$Y_0 := \left(z_0^T \cdot \beta_{hat}\right)_1$ $Y_0 = 151.841$ **< Prediction for y_o verified more-or-less p. 376**

Confidence interval for expected value of Y_o:

$\alpha := 0.05$ **< Set probability of Type I error**

$$CI_{lower} := Y_0 - qt\left(1 - \frac{\alpha}{2}, n - r - 1\right) \cdot \sqrt{\left(z_0^T \cdot ZZ \cdot z_0\right) \cdot s_{sq}} \qquad \sqrt{\left(z_0^T \cdot ZZ \cdot z_0\right) \cdot s_{sq}} = (0.732)$$

$$CI_{upper} := Y_0 + qt\left(1 - \frac{\alpha}{2}, n - r - 1\right) \cdot \sqrt{\left(z_0^T \cdot ZZ \cdot z_0\right) \cdot s_{sq}} \qquad qt\left(1 - \frac{\alpha}{2}, n - r - 1\right) = 2.776$$

$CI := augment\left(CI_{lower}, CI_{upper}\right)$

Confidence interval for expected value:

$Y_0 = 151.841$ **< prediction** $CI = (149.807 \quad 153.874)$ **< Confidence interval for expected value verified p. 377**

Prediction of Y_o:

$$CI_{lower} := Y_0 - qt\left(1 - \frac{\alpha}{2}, n - r - 1\right) \cdot \sqrt{\left(1 + z_0^T \cdot ZZ \cdot z_0\right) \cdot s_{sq}} \qquad \sqrt{\left(1 + z_0^T \cdot ZZ \cdot z_0\right) \cdot s_{sq}} = (1.409)$$

$$CI_{upper} := Y_0 + qt\left(1 - \frac{\alpha}{2}, n - r - 1\right) \cdot \sqrt{\left(1 + z_0^T \cdot ZZ \cdot z_0\right) \cdot s_{sq}} \qquad qt\left(1 - \frac{\alpha}{2}, n - r - 1\right) = 2.776$$

$CI := augment\left(CI_{lower}, CI_{upper}\right)$

Confidence interval for predicted new observation:

$Y_0 = 151.841$ **< prediction** $CI = (147.928 \quad 155.753)$ **< Confidence interval for new observation verified p. 377**

2005 jw383

Prepared by: Wm Stein

MULTIVARIATE AND UNIVARIATE MULTIPLE LINEAR REGRESSION

ORIGIN ≡ 1

jw383.mcd

Template for Multivariate multiple regression followed by separate univariate muliple regressions for each dependent variable.
Using Amitriptyline Data Table 7.6 p. 423
Two dependent variables: TOT & AMI - Columns 1 & 2
Five independent variables: GEN, AMT, PR, DIAP, QRS - Cols 3-7

Read in data from jw Table 7.6, p. 423:

M := READPRN("DATA\t7-6.DAT")

n := rows(M) n = 17

M =

	1	2	3	4	5	6	7
1	3389	3149	1	7500	220	0	140
2	1101	653	1	1975	200	0	100
3	1131	810	0	3600	205	60	111
4	596	448	1	675	160	60	120
5	896	844	1	750	185	70	83
6	1767	1450	1	2500	180	60	80
7	807	493	1	350	154	80	98
8	1111	941	0	1500	200	70	93
9	645	547	1	375	137	60	105
10	628	392	1	1050	167	60	74
11	1360	1283	1	3000	180	60	80
12	652	458	1	450	160	64	60
13	860	722	1	1750	135	90	79
14	500	384	0	2000	160	60	80
15	781	501	0	4500	180	0	100
16	1070	405	0	1500	170	90	120
17	1754	1520	1	3000	180	0	129

Constructing variables:

- Dependent variable matrix (Y) is the first & second column of M:

Y := submatrix(M, 1, 17, 1, 2)

- Matrix (Z) consisting in independent variables and first column of 1's:

$i := 1 .. n \quad l_i := 1$

Z := augment(l, submatrix(M, 1, 17, 3, 7))

r := cols(Z) − 1 r = 5

Calculating least-squares estimator:
(using jw Eq. 7-28 p. 384)

$$\beta_{hat} := \left(Z^T \cdot Z\right)^{-1} \cdot Z^T \cdot Y$$

$$\beta_{hat} = \begin{pmatrix} -2879.478 & -2728.709 \\ 675.651 & 763.03 \\ 0.285 & 0.306 \\ 10.272 & 8.896 \\ 7.251 & 7.206 \\ 7.598 & 4.987 \end{pmatrix}$$

- Fitted values of Y [= Y_{hat}]:

$$Y_{hat} := Z \cdot \beta_{hat}$$

- Residuals [= ε_{hat}]:

$$\varepsilon_{hat} := Y - Y_{hat}$$

$$Y = \begin{pmatrix} 3389 & 3149 \\ 1101 & 653 \\ 1131 & 810 \\ 596 & 448 \\ 896 & 844 \\ 1767 & 1450 \\ 807 & 493 \\ 1111 & 941 \\ 645 & 547 \\ 628 & 392 \\ 1360 & 1283 \\ 652 & 458 \\ 860 & 722 \\ 500 & 384 \\ 781 & 501 \\ 1070 & 405 \\ 1754 & 1520 \end{pmatrix} \quad Z = \begin{pmatrix} 1 & 1 & 7500 & 220 & 0 & 140 \\ 1 & 1 & 1975 & 200 & 0 & 100 \\ 1 & 0 & 3600 & 205 & 60 & 111 \\ 1 & 1 & 675 & 160 & 60 & 120 \\ 1 & 1 & 750 & 185 & 70 & 83 \\ 1 & 1 & 2500 & 180 & 60 & 80 \\ 1 & 1 & 350 & 154 & 80 & 98 \\ 1 & 0 & 1500 & 200 & 70 & 93 \\ 1 & 1 & 375 & 137 & 60 & 105 \\ 1 & 1 & 1050 & 167 & 60 & 74 \\ 1 & 1 & 3000 & 180 & 60 & 80 \\ 1 & 1 & 450 & 160 & 64 & 60 \\ 1 & 1 & 1750 & 135 & 90 & 79 \\ 1 & 0 & 2000 & 160 & 60 & 80 \\ 1 & 0 & 4500 & 180 & 0 & 100 \\ 1 & 0 & 1500 & 170 & 90 & 120 \\ 1 & 1 & 3000 & 180 & 0 & 129 \end{pmatrix}$$

$$Y_{hat} = \begin{pmatrix} 3256.178 & 2987.472 \\ 1173.004 & 917.353 \\ 1530.248 & 1183.852 \\ 978.847 & 695.295 \\ 1048.391 & 828.212 \\ 1400.214 & 1232.868 \\ 802.5 & 576.742 \\ 816.443 & 478.276 \\ 543.159 & 323.964 \\ 808.052 & 643.054 \\ 1542.639 & 1386.055 \\ 487.866 & 355.96 \\ 934.266 & 813.939 \\ 376.695 & 138.728 \\ 1011.159 & 749.992 \\ 858.455 & 490.152 \\ 1479.883 & 1198.087 \end{pmatrix} \quad \varepsilon_{hat} = \begin{pmatrix} 132.822 & 161.528 \\ -72.004 & -264.353 \\ -399.248 & -373.852 \\ -382.847 & -247.295 \\ -152.391 & 15.788 \\ 366.786 & 217.132 \\ 4.5 & -83.742 \\ 294.557 & 462.724 \\ 101.841 & 223.036 \\ -180.052 & -251.054 \\ -182.639 & -103.055 \\ 164.134 & 102.04 \\ -74.266 & -91.939 \\ 123.305 & 245.272 \\ -230.159 & -248.992 \\ 211.545 & -85.152 \\ 274.117 & 321.913 \end{pmatrix}$$

Sums of Squares decomposition jw Eq. 7-34 p. 385:

$j := 1..\text{cols}(Y) \qquad k := 1..\text{cols}(Z)$

- **Total sum of squares & cross products: ($SSCP_T$) for mean centered Y:**

$$Y_{mc}^{\langle j \rangle} := Y^{\langle j \rangle} - \text{mean}\left(Y^{\langle j \rangle}\right)$$

$$SSCP_T := Y_{mc}^T \cdot Y_{mc} \qquad SSCP_T = \begin{pmatrix} 7.7059 \times 10^6 & 7.4748 \times 10^6 \\ 7.4748 \times 10^6 & 7.6104 \times 10^6 \end{pmatrix}$$

- **Predicted (hypothesis) sum of squares & cross produxts ($SSCP_R$) for mean centered Y_{hat}:**

$$Yhat_{mc}^{\langle j \rangle} := Y_{hat}^{\langle j \rangle} - \text{mean}\left(Y_{hat}^{\langle j \rangle}\right)$$

$$SSCP_R := Yhat_{mc}^T \cdot Yhat_{mc} \qquad SSCP_R = \begin{pmatrix} 6.8359 \times 10^6 & 6.7091 \times 10^6 \\ 6.7091 \times 10^6 & 6.6697 \times 10^6 \end{pmatrix}$$

- **Residual (error) sum of squares:**

$$SSCP_E := \varepsilon_{hat}^T \cdot \varepsilon_{hat} \qquad SSCP_E = \begin{pmatrix} 8.7001 \times 10^5 & 7.6568 \times 10^5 \\ 7.6568 \times 10^5 & 9.4071 \times 10^5 \end{pmatrix}$$

- **Verify that $SSCP_T = SSCP_R + SSCP_E$:**

$$SSCP_T = \begin{pmatrix} 7.7059 \times 10^6 & 7.4748 \times 10^6 \\ 7.4748 \times 10^6 & 7.6104 \times 10^6 \end{pmatrix} \qquad SSCP_R + SSCP_E = \begin{pmatrix} 7.7059 \times 10^6 & 7.4748 \times 10^6 \\ 7.4748 \times 10^6 & 7.6104 \times 10^6 \end{pmatrix}$$

Calculating coefficient of multiple determination [= R^2]
and Multiple Correlation Coeffient [R] jw Eq. 7-9, p. 361

Univariate coefficient of multiple determination (R^2) and multiple correlation (R):

$$R_{sq_j} := \frac{SSCP_{R_{j,j}}}{SSCP_{T_{j,j}}} \qquad R_{sq} = \begin{pmatrix} 0.8871 \\ 0.87639 \end{pmatrix}$$

< R^2 and R are calculated for each dependent variable Y taking into account only the main diagonal sums of squares in $SSCP_R$ & $SSCP_T$

$$R := \sqrt{R_{sq}} \qquad R = \begin{pmatrix} 0.94186 \\ 0.93616 \end{pmatrix}$$

Multivariate coefficient of multiple determination (R^2) - Rencher, p. 380:

$$R_{sq} := \frac{|SSCP_R|}{|SSCP_T|} \qquad R_{sq} = 0.21 \qquad \sqrt{R_{sq}} = 0.458$$

< Multivariate analog to Univariate R^2 and R

2005 jw383

Test for effects: variable GEN

Error SSCP:

$E := SSCP_E$

$$E = \begin{pmatrix} 8.70008 \times 10^5 & 7.65676 \times 10^5 \\ 7.65676 \times 10^5 & 9.40709 \times 10^5 \end{pmatrix}$$

Hypothesis SSCP (Subtracting E from $SSCP_E$ of reduced model) jw p 395:

$i := 1..n \quad l_i := 1$

$Z_r := \text{augment}(l, \text{submatrix}(M,1,17,4,7))$ **< Reduced Z matrix (Z_r)**

$r := \text{cols}(Z) - 1 \quad r = 5$

$\beta hat_r := \left(Z_r^T \cdot Z_r\right)^{-1} \cdot Z_r^T \cdot Y$ **< Fitting reduced model regression coefficients (βhat_r)**

$Yhat_r := Z_r \cdot \beta hat_r$ **< Fitted Y values ($Yhat_r$)**

$\varepsilon hat_r := Y - Yhat_r$ **< Residuals/Error (εhat_r)**

$SSCP_{Er} := \varepsilon hat_r^T \cdot \varepsilon hat_r$ **< $SSCP_E$ for reduced model**

$H := SSCP_{Er} - SSCP_E$

$$H = \begin{pmatrix} 1.37482 \times 10^6 & 1.55262 \times 10^6 \\ 1.55262 \times 10^6 & 1.75342 \times 10^6 \end{pmatrix}$$

Null Hypothesis:

H_o: Regression coefficients β =0 for the independent variables REMOVED from the Full regression model to make the Reduced model.

Note: REJECTION of H_o means that the β's are NEEDED in the Regression equation whereas failure to reject H_o means one should proceed with the Reduced Model as the viable regression because the extra variables are NOT NEEDED.

Multivariate Test Statistics:

Wilks' Lambda:

$\Lambda := \dfrac{|E|}{|E+H|} \quad \Lambda = 0.3447916$

Pillai's Trace:

$P_{tr} := tr\left[H \cdot (H+E)^{-1}\right] \quad P_{tr} = 0.6552084$

Hotelling-Lawley Trace:

$HT_{tr} := tr\left(H \cdot E^{-1}\right) \quad HT_{tr} = 1.9003029$

$r := \text{rows}(\beta_{hat}) - 1 \quad r = 5 \quad q := \text{rows}(\beta hat_r) - 1 \quad q = 4 \quad m := \text{cols}(Y) \quad m = 2$

$K := -\left[n - r - 1 - \frac{1}{2}\cdot(m - r + q + 1)\right]\cdot \ln(\Lambda) \quad K = 10.648$

Stringency of the test: $\alpha := 0.05$ **< set as desired**

$C := \text{qchisq}\left[1 - \alpha, m\cdot(r-q)\right] \quad C = 5.991$ **< jw Result 7.11 p. 393**

Decision Rule: Reject H_0 if K > C:

Probability: $1 - \text{pchisq}\left[K, m\cdot(r-q)\right] = 0.0048729$

$\text{Decision} := \text{if}(K > C, 1, 0) \quad \text{Decision} = 1$ **< 0 = Do not reject H_0**
1 = Reject H_0

Univariate F Tests: (see Rencher, p. 358)

$k := 1..2 \quad h := r - q$ **< h = number of parameters deleted from the reduced model**

F-ratio statistic (F_κ):

$$F_{\kappa_k} := \frac{\frac{H_{k,k}}{h}}{\frac{E_{k,k}}{(n-r-1)}} \quad F_\kappa = \begin{pmatrix} 17.383 \\ 20.503 \end{pmatrix}$$

< Separate F-values for each Y variable

Probability:

$\text{Prob} := 1 - \text{pF}\left[F_\kappa, h, (n-r-1)\right]$

$$\text{Prob} = \begin{pmatrix} 0.0015647 \\ 0.0008606 \end{pmatrix}$$

2005 jw383

Test for effects: variable AMT

Error SSCP:

$E := SSCP_E$

$$E = \begin{pmatrix} 8.70008 \times 10^5 & 7.65676 \times 10^5 \\ 7.65676 \times 10^5 & 9.40709 \times 10^5 \end{pmatrix}$$

Hypothesis SSCP (Subtracting E from $SSCP_E$ of reduced model) jw p 395:

$i := 1 .. n \quad l_i := 1$

$Z_r := \text{augment}(l, \text{submatrix}(M,1,17,3,3), \text{submatrix}(M,1,17,5,7))$ **< Reduced Z matrix (Z_r)**

$r := \text{cols}(Z) - 1 \quad r = 5$

$\beta hat_r := \left(Z_r^T \cdot Z_r\right)^{-1} \cdot Z_r^T \cdot Y$ **< Fitting reduced model regression coefficients (βhat_r)**

$Yhat_r := Z_r \cdot \beta hat_r$ **< Fitted Y values ($Yhat_r$)**

$\varepsilon hat_r := Y - Yhat_r$ **< Residuals/Error (εhat_r)**

$SSCP_{Er} := \varepsilon hat_r^T \cdot \varepsilon hat_r$ **< $SSCP_E$ for reduced model**

$H := SSCP_{Er} - SSCP_E$

$$H = \begin{pmatrix} 1.72976 \times 10^6 & 1.86046 \times 10^6 \\ 1.86046 \times 10^6 & 2.00103 \times 10^6 \end{pmatrix}$$

Null Hypothesis:

H_o: Regression coefficients β =0 for the independent variables REMOVED from the Full regression model to make the Reduced model.

Note: REJECTION of H_o means that the β's are NEEDED in the Regression equation whereas failure to reject H_o means one should proceed with the Reduced Model as the viable regression because the extra variables are NOT NEEDED.

Multivariate Test Statistics:

Wilks' Lambda:

$\Lambda := \frac{|E|}{|E + H|} \quad \Lambda = 0.3090326$

Pillai's Trace:

$P_{tr} := tr\left[H \cdot (H + E)^{-1}\right] \quad P_{tr} = 0.6909674$

Hotelling-Lawley Trace:

$HT_{tr} := tr\left(H \cdot E^{-1}\right) \quad HT_{tr} = 2.2359042$

$r := \text{rows}(\beta_{hat}) - 1 \quad r = 5 \qquad q := \text{rows}(\beta hat_r) - 1 \quad q = 4 \qquad m := \text{cols}(Y) \quad m = 2$

$K := -\left[n - r - 1 - \frac{1}{2} \cdot (m - r + q + 1)\right] \cdot \ln(\Lambda) \qquad K = 11.743$

Stringency of the test: $\alpha := 0.05$ **< set as desired**

$C := \text{qchisq}[1 - \alpha, m \cdot (r - q)] \qquad C = 5.991$ **< jw Result 7.11 p. 393**

Decision Rule: Reject H_0 if K > C:

Probability: $1 - \text{pchisq}[K, m \cdot (r - q)] = 0.0028185$

$\text{Decision} := \text{if}(K > C, 1, 0) \qquad \text{Decision} = 1$

< 0 = Do not reject H_0
1 = Reject H_0

Univariate F Tests: (see Rencher, p. 358)

$k := 1 .. 2 \qquad h := r - q$ **< h = number of parameters deleted from the reduced model**

F-ratio statistic (F_κ):

$$F_{\kappa_k} := \frac{\frac{H_{k,k}}{h}}{\frac{E_{k,k}}{(n-r-1)}} \qquad F_\kappa = \begin{pmatrix} 21.87 \\ 23.399 \end{pmatrix}$$

< Separate F-values for each Y variable

Probability:

$Prob := 1 - pF[F_\kappa, h, (n - r - 1)]$

$$Prob = \begin{pmatrix} 0.0006753 \\ 0.0005213 \end{pmatrix}$$

2005 jw383

Test for effects: ALL VARIABLES (Constant β_o left in model only)

Error SSCP:

$E := SSCP_E$ $\qquad$ $E = \begin{pmatrix} 8.70008 \times 10^5 & 7.65676 \times 10^5 \\ 7.65676 \times 10^5 & 9.40709 \times 10^5 \end{pmatrix}$

Hypothesis SSCP (Subtracting E from $SSCP_E$ of reduced model) jw p 395:

$i := 1 .. n \quad l_i := 1$

$Z_r := l$ $\qquad$ **< Reduced Z matrix (Z_r)**

$r := cols(Z) - 1 \quad r = 5$

$\beta hat_r := \left(Z_r^T \cdot Z_r\right)^{-1} \cdot Z_r^T \cdot Y$ $\qquad$ **< Fitting reduced model regression coefficients (βhat_r)**

$Yhat_r := Z_r \cdot \beta hat_r$ $\qquad$ **< Fitted Y values ($Yhat_r$)**

$\varepsilon hat_r := Y - Yhat_r$ $\qquad$ **< Residuals/Error (εhat_r)**

$SSCP_{Er} := \varepsilon hat_r^T \cdot \varepsilon hat_r$ $\qquad$ **< $SSCP_E$ for reduced model**

$H := SSCP_{Er} - SSCP_E$ $\qquad$ $H = \begin{pmatrix} 6.83593 \times 10^6 & 6.70909 \times 10^6 \\ 6.70909 \times 10^6 & 6.66967 \times 10^6 \end{pmatrix}$

Null Hypothesis:

H_o: Regression coefficients β =0 for the independent variables REMOVED from the Full regression model to make the Reduced model.

Note: REJECTION of H_o means that the β's are NEEDED in the Regression equation whereas failure to reject H_o means one should proceed with the Reduced Model as the viable regression because the extra variables are NOT NEEDED.

Multivariate Test Statistics:

Wilks' Lambda:

$\Lambda := \frac{|E|}{|E + H|} \quad \Lambda = 0.0837239$

Pillai's Trace:

$P_{tr} := tr\left[H \cdot (H + E)^{-1}\right] \quad P_{tr} = 1.1259807$

Hotelling-Lawley Trace:

$HT_{tr} := tr\left(H \cdot E^{-1}\right) \quad HT_{tr} = 8.439305$

$r := rows\left(\beta_{hat}\right) - 1 \quad r = 5 \quad q := rows\left(\beta hat_r\right) - 1 \quad q = 0 \quad m := cols(Y) \quad m = 2$

$K := -\left[n - r - 1 - \frac{1}{2} \cdot (m - r + q + 1)\right] \cdot \ln(\Lambda) \qquad K = 29.763$

Stringency of the test: $\alpha := 0.05$ **< set as desired**

$C := qchisq\left[1 - \alpha, m \cdot (r - q)\right] \qquad C = 18.307$ $\qquad$ **< jw Result 7.11 p. 393**

Decision Rule: Reject H_0 if K > C:

Probability: $1 - pchisq\left[K, m \cdot (r - q)\right] = 0.0009366$

$Decision := if(K > C, 1, 0) \qquad Decision = 1$ $\qquad$ **< 0 = Do not reject H_0**
1 = Reject H_0

Univariate F Tests: (see Rencher, p. 358)

$k := 1 .. 2 \qquad h := r - q$ $\qquad$ **< h = number of parameters deleted from the reduced model**

F-ratio statistic (F_κ):

$F_{\kappa_k} := \frac{\frac{H_{k,k}}{h}}{\frac{E_{k,k}}{(n-r-1)}} \qquad F_\kappa = \begin{pmatrix} 17.286 \\ 15.598 \end{pmatrix}$ $\qquad$ **< Separate F-values for each Y variable**

Probability:

$Prob := 1 - pF\left[F_\kappa, h, (n - r - 1)\right]$

$Prob = \begin{pmatrix} 0.0000698 \\ 0.0001132 \end{pmatrix}$

Test for effects: CONSTANT

Error SSCP:

$E := SSCP_E$ $$E = \begin{pmatrix} 8.70008\times10^5 & 7.65676\times10^5 \\ 7.65676\times10^5 & 9.40709\times10^5 \end{pmatrix}$$

Hypothesis SSCP (Subtracting E from $SSCP_E$ of reduced model) jw p 395:

$i := 1..n \quad l_i := 1$

$Z_r := \text{submatrix}(M,1,17,3,7)$ **< Reduced Z matrix (Z_r)**

$r := \text{cols}(Z) - 1 \quad r = 5$

$\beta hat_r := \left(Z_r^T \cdot Z_r\right)^{-1} \cdot Z_r^T \cdot Y$ **< Fitting reduced model regression coefficients (βhat_r)**

$Yhat_r := Z_r \cdot \beta hat_r$ **< Fitted Y values ($Yhat_r$)**

$\varepsilon hat_r := Y - Yhat_r$ **< Residuals/Error (εhat_r)**

$SSCP_{Er} := \varepsilon hat_r^T \cdot \varepsilon hat_r$ **< $SSCP_E$ for reduced model**

$H := SSCP_{Er} - SSCP_E$ $$H = \begin{pmatrix} 8.21868\times10^5 & 7.78835\times10^5 \\ 7.78835\times10^5 & 7.38055\times10^5 \end{pmatrix}$$

Null Hypothesis:

H_o: Regression coefficients $\beta = 0$ for the independent variables REMOVED from the Full regression model to make the Reduced model.

Note: REJECTION of H_o means that the β's are NEEDED in the Regression equation whereas failure to reject H_o means one should proceed with the Reduced Model as the viable regression because the extra variables are NOT NEEDED.

Multivariate Test Statistics:

Wilks' Lambda: $\Lambda := \frac{|E|}{|E+H|} \quad \Lambda = 0.5105362$

Pillai's Trace: $P_{tr} := tr\left[H\cdot(H+E)^{-1}\right] \quad P_{tr} = 0.4894638$

Hotelling-Lawley Trace: $HT_{tr} := tr\left(H\cdot E^{-1}\right) \quad HT_{tr} = 0.9587249$

$r := \text{rows}(\beta_{hat}) - 1 \quad r = 5 \qquad q := \text{rows}(\beta hat_r) - 1 \quad q = 4 \qquad m := \text{cols}(Y) \quad m = 2$

$K := -\left[n - r - 1 - \frac{1}{2}\cdot(m - r + q + 1)\right]\cdot\ln(\Lambda) \qquad K = 6.723$

Stringency of the test: $\alpha := 0.05$ **< set as desired**

$C := \text{qchisq}\left[1-\alpha, m\cdot(r-q)\right] \qquad C = 5.991$ **< jw Result 7.11 p. 393**

Decision Rule: Reject H_0 if K > C:

Probability: $1 - \text{pchisq}\left[K, m\cdot(r-q)\right] = 0.0346843$

$\text{Decision} := \text{if}(K > C, 1, 0) \qquad \text{Decision} = 1$ **< 0 = Do not reject H_0**
1 = Reject H_0

Univariate F Tests: (see Rencher, p. 358)

$k := 1..2 \qquad h := r - q$ **< h = number of parameters deleted from the reduced model**

F-ratio statistic (F_κ):

$$F_{\kappa_k} := \frac{\frac{H_{k,k}}{h}}{\frac{E_{k,k}}{(n-r-1)}} \qquad F_\kappa = \begin{pmatrix} 10.391 \\ 8.63 \end{pmatrix}$$

< Separate F-values for each Y variable

Probability:

$Prob := 1 - pF\left[F_\kappa, h, (n-r-1)\right]$

$$Prob = \begin{pmatrix} 0.0081076 \\ 0.0135022 \end{pmatrix}$$

ORIGIN ≡ 1

PRINCIPAL COMPONENTS ANALYSIS
jw439.mcd

Prepared by:
Wm Stein

The 1970 Census provided tract information on 5 socioeconomic variables for the Madison, Wisc. area. The data for 14 tracts are listed in J.W. Table 8.5, p. 470.

Read in Data:

M := READPRN("DATA\T8-5.DAT")

$X := M^T$

M =

	1	2	3	4	5
1	5.935	14.2	2.265	2.27	2.91
2	1.523	13.1	0.597	0.75	2.62
3	2.599	12.7	1.237	1.11	1.72
4	4.009	15.2	1.649	0.81	3.02
5	4.687	14.7	2.312	2.5	2.22
6	8.044	15.6	3.641	4.51	2.36
7	2.766	13.3	1.244	1.03	1.97
8	6.538	17	2.618	2.39	1.85
9	6.451	12.9	3.147	5.52	2.01
10	3.314	12.2	1.606	2.18	1.82
11	3.777	13	2.119	2.83	1.8
12	1.53	13.8	0.798	0.84	4.25
13	2.768	13.6	1.336	1.75	2.64
14	6.585	14.9	2.763	1.91	3.17

Calculating summary statistics:

$n := \text{cols}(X) \qquad n = 14 \qquad p := \text{rows}(X) \qquad p = 5$

$i := 1..n \qquad j := 1..p$

$l_i := 1$

Variance/covariance (S):

$I := \text{identity}(n)$

$$S := \frac{1}{n-1}\cdot X\cdot\left(I - \frac{1}{n}\cdot l\cdot l^T\right)\cdot X^T$$

Mean vector (X_{bar}):

$$X_{bar} := \frac{1}{n}\cdot X\cdot l$$

$$X_{bar} = \begin{pmatrix} 4.323 \\ 14.014 \\ 1.952 \\ 2.171 \\ 2.454 \end{pmatrix}$$

$$S = \begin{pmatrix} 4.308 & 1.684 & 1.803 & 2.155 & -0.253 \\ 1.684 & 1.767 & 0.588 & 0.178 & 0.176 \\ 1.803 & 0.588 & 0.801 & 1.065 & -0.158 \\ 2.155 & 0.178 & 1.065 & 1.969 & -0.357 \\ -0.253 & 0.176 & -0.158 & -0.357 & 0.504 \end{pmatrix}$$

Sum of Variance in S:

$$\sum_j S_{j,j} = 9.35$$

Calculate eigenvalues & eigenvectors of variance/covariance matrix (S):

$\Lambda := \text{reverse}(\text{sort}(\text{eigenvals}(S)))$

$E^{\langle j \rangle} := \text{eigenvec}(S, \Lambda_j)$

$$\Lambda = \begin{pmatrix} 6.931 \\ 1.785 \\ 0.39 \\ 0.23 \\ 0.014 \end{pmatrix}$$

$$E = \begin{pmatrix} 0.78121 & 0.07087 & 0.00366 & -0.54171 & -0.30204 \\ 0.30565 & 0.76387 & -0.16182 & 0.5448 & -0.00928 \\ 0.33445 & -0.08291 & 0.01484 & -0.05102 & 0.93726 \\ 0.42601 & -0.57946 & 0.22045 & 0.63601 & -0.17215 \\ -0.05435 & 0.26236 & 0.96176 & -0.05128 & 0.02458 \end{pmatrix}$$

Testing the Identity SE = ED:

$I := \text{identity}(p)$

$D_{j,j} := \Lambda_j$

Diagonal Matrix of Eigenvalues (D) :

$$D = \begin{pmatrix} 6.931 & 0 & 0 & 0 & 0 \\ 0 & 1.785 & 0 & 0 & 0 \\ 0 & 0 & 0.39 & 0 & 0 \\ 0 & 0 & 0 & 0.23 & 0 \\ 0 & 0 & 0 & 0 & 0.014 \end{pmatrix}$$

Sum of Variance in Λ:

$$\sum_j \Lambda_j = 9.35$$

$$S\cdot E = \begin{pmatrix} 5.41461 & 0.12652 & 0.00142 & -0.12434 & -0.00428 \\ 2.11847 & 1.36362 & -0.06305 & 0.12505 & -0.00013 \\ 2.31809 & -0.148 & 0.00578 & -0.01171 & 0.01327 \\ 2.95269 & -1.03442 & 0.0859 & 0.14598 & -0.00244 \\ -0.37673 & 0.46834 & 0.37475 & -0.01177 & 0.00035 \end{pmatrix}$$

$$E\cdot D = \begin{pmatrix} 5.41461 & 0.12652 & 0.00142 & -0.12434 & -0.00428 \\ 2.11847 & 1.36362 & -0.06305 & 0.12505 & -0.00013 \\ 2.31809 & -0.148 & 0.00578 & -0.01171 & 0.01327 \\ 2.95269 & -1.03442 & 0.0859 & 0.14598 & -0.00244 \\ -0.37673 & 0.46834 & 0.37475 & -0.01177 & 0.00035 \end{pmatrix}$$

2005 jw439

Calculating Principal Component Scores (Y) as linear combinations of E and X:

$$Y := E^T \cdot X$$

Calculating loadings of original variables on principal components:

NOTE:
Be certain that Eigenvalues (Λ) are in rank order and that Eigenvectors (E) are in columns in the same rank order. Otherwise calculations below will take into account the wrong entry in the variance / covariance matrix S.

$i := 1 .. p \quad k := 1 .. p$

$$\rho_{i,k} := \frac{E_{k,i} \cdot \sqrt{\Lambda_i}}{\sqrt{S_{k,k}}}$$ **J.W Result 8.3, p. 429**

$Y^T =$

	1	2	3	4	5
1	10.54307	10.52792	1.05666	5.70009	-0.12073
2	5.57054	10.31795	0.57977	6.62406	-0.08673
3	6.70519	9.59088	-0.12829	6.06572	0.10773
4	8.51014	12.08123	0.66259	6.38543	0.12841
5	9.87215	10.50321	0.35897	6.8278	0.23908
6	14.0629	10.19044	0.8231	6.70301	0.11982
7	6.97371	10.1724	-0.00186	6.23808	0.0782
8	12.09675	12.3326	-0.382	7.01154	-0.0447
9	12.27726	7.37897	1.13289	6.78052	-0.01945
10	7.68473	8.63524	0.29277	6.06258	0.06053
11	8.74052	8.85472	0.29668	6.63587	0.28168
12	5.80693	11.21198	2.05702	6.96503	0.11763
13	7.36805	10.15264	0.75407	6.81931	0.05357
14	11.26387	11.34422	1.12385	5.46163	0.21157

Compare with results in chart J.W. p. 440

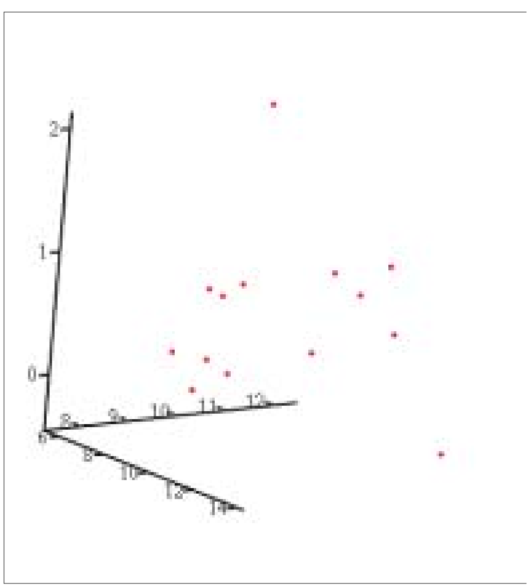
Eigenvectors (component values = "loadings")

$$E = \begin{pmatrix} 0.78121 & 0.07087 & 0.00366 & -0.54171 & -0.30204 \\ 0.30565 & 0.76387 & -0.16182 & 0.5448 & -0.00928 \\ 0.33445 & -0.08291 & 0.01484 & -0.05102 & 0.93726 \\ 0.42601 & -0.57946 & 0.22045 & 0.63601 & -0.17215 \\ -0.05435 & 0.26236 & 0.96176 & -0.05128 & 0.02458 \end{pmatrix}$$

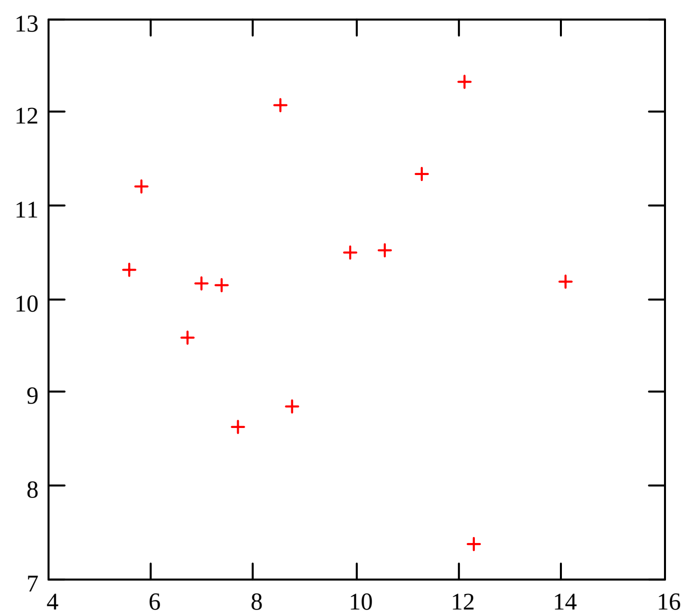
Correlations with original variables

$$\rho^T = \begin{pmatrix} 0.99095 & 0.04562 & 0.0011 & -0.12505 & -0.01731 \\ 0.60527 & 0.76768 & -0.07598 & 0.19633 & -0.00083 \\ 0.98402 & -0.1238 & 0.01035 & -0.02732 & 0.12462 \\ 0.79918 & -0.55168 & 0.09806 & 0.21712 & -0.01459 \\ -0.20149 & 0.49357 & 0.84533 & -0.03459 & 0.00412 \end{pmatrix}$$

Plot of component scores (Y):



$$\left[\left(Y^T\right)^{\langle 1\rangle}, \left(Y^T\right)^{\langle 2\rangle}, \left(Y^T\right)^{\langle 3\rangle}\right]$$



$\left(Y^T\right)^{\langle 2\rangle}$ +++

$\left(Y^T\right)^{\langle 1\rangle}$

Fraction Total Variance Explained by each Principal Component:

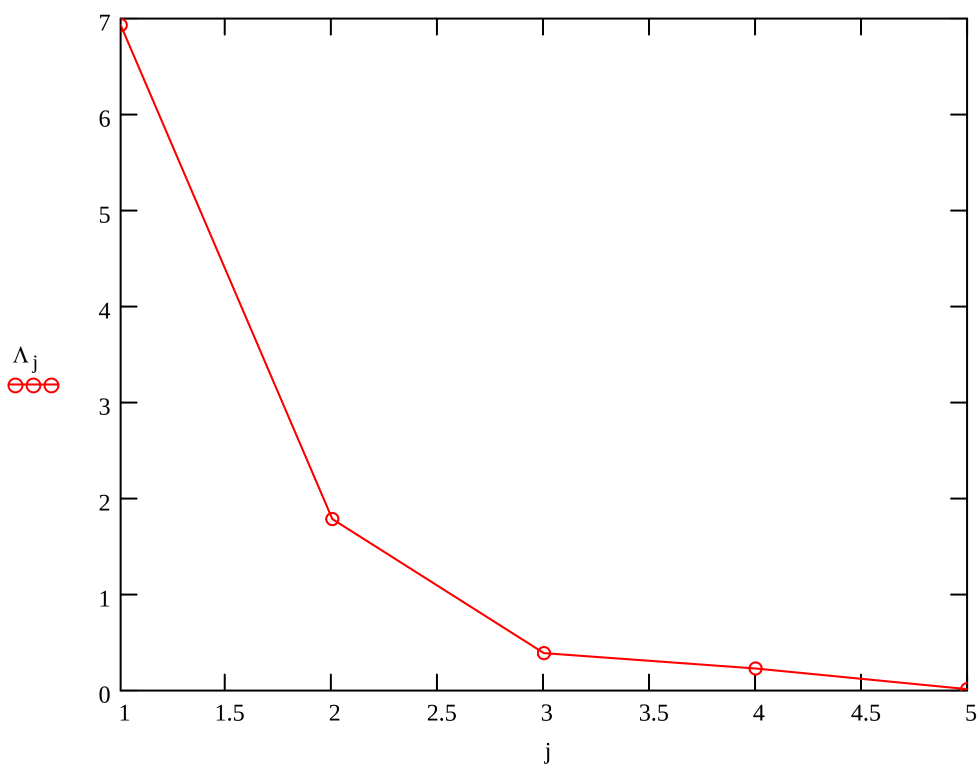
Eigenvalues and Total Variance:

$$\Lambda = \begin{pmatrix} 6.931 \\ 1.785 \\ 0.39 \\ 0.23 \\ 0.014 \end{pmatrix} \qquad \sum_j \Lambda_j = 9.35$$

Fraction of Variance Explained by each:

$$FE_j := \frac{\Lambda_j}{\sum_j \Lambda_j} \qquad FE = \begin{pmatrix} 0.7413 \\ 0.1909 \\ 0.0417 \\ 0.0245 \\ 0.0015 \end{pmatrix}$$

Scree Plot:



$$FE_1 + FE_2 = 0.932$$

Since the first two Principal Components "explain" ~ 93% of the Total Variance, they would seem to be reasonable descriptions of the data by themselves - thus reducing the data from five dimensions for the original variables to two linear combinations that are thePrincipal Components.

Note: Use of Principal Components Analysis (PCA) is as close as one can get to "guilt free" statistics!

Chosing the number of Principal Components to use in "data reduction" is a judgement call but doesn't depend on any underlying statistical assumption (such as Normality of the original data). Graphing data in reduced dimensions of the Principal Components is often a first step in interpretation or use of the data with our without further statistical analysis.

A similar analysis might have been conducted using the Correlation Matrix R instead of Variance-Covariance Matrix S. In general, the results will be different. Which to use and interpret is also a judgement call, with no hard-and-fast rules. Greatly different measurement scales among the original variables generally supports use of correlations instead of variance-covariance, but it much depends on whether scale per se is a major part of the problem or something the researcher intends to control.   I recommend trying both approaches!

Prepared by:
Wm Stein

PRINCIPAL COMPONENTS ANALYSIS AND SINGULAR VALUE DECOMPOSITION

ORIGIN ≡ 1

The 1970 Census provided tract information on 5 socioeconomic variables for the Madison, Wisc. area. The data for 14 tracts are listed in J.W. Table 8.5, p. 470.

Read in Data:

M := READPRN("DATA\T8-5.DAT")

M =

	1	2	3	4	5
1	5.935	14.2	2.265	2.27	2.91
2	1.523	13.1	0.597	0.75	2.62
3	2.599	12.7	1.237	1.11	1.72
4	4.009	15.2	1.649	0.81	3.02
5	4.687	14.7	2.312	2.5	2.22
6	8.044	15.6	3.641	4.51	2.36
7	2.766	13.3	1.244	1.03	1.97
8	6.538	17	2.618	2.39	1.85
9	6.451	12.9	3.147	5.52	2.01
10	3.314	12.2	1.606	2.18	1.82
11	3.777	13	2.119	2.83	1.8
12	1.53	13.8	0.798	0.84	4.25

Sizing data array:

$X := M^T$

$n := \text{cols}(X) \qquad n = 14 \qquad p := \text{rows}(X) \qquad p = 5$

$i := 1..n \qquad j := 1..p$

$l_i := 1$

Mean vector (X_{bar}):

$$X_{bar} := \frac{1}{n}\cdot X\cdot l \qquad X_{bar} = \begin{pmatrix} 4.323 \\ 14.014 \\ 1.952 \\ 2.171 \\ 2.454 \end{pmatrix}$$

Mean Centering the data:

$$Y^{\langle i\rangle} := X^{\langle i\rangle} - X_{bar}$$

$$M_c := Y^T$$

$$M_c = \begin{pmatrix} 1.612 & 0.186 & 0.313 & 0.099 & 0.456 \\ -2.8 & -0.914 & -1.355 & -1.421 & 0.166 \\ -1.724 & -1.314 & -0.715 & -1.061 & -0.734 \\ -0.314 & 1.186 & -0.303 & -1.361 & 0.566 \\ 0.364 & 0.686 & 0.36 & 0.329 & -0.234 \\ 3.721 & 1.586 & 1.689 & 2.339 & -0.094 \\ -1.557 & -0.714 & -0.708 & -1.141 & -0.484 \\ 2.215 & 2.986 & 0.666 & 0.219 & -0.604 \\ 2.128 & -1.114 & 1.195 & 3.349 & -0.444 \\ -1.009 & -1.814 & -0.346 & 0.009 & -0.634 \\ -0.546 & -1.014 & 0.167 & 0.659 & -0.654 \\ -2.793 & -0.214 & -1.154 & -1.331 & 1.796 \\ -1.555 & -0.414 & -0.616 & -0.421 & 0.186 \\ 2.262 & 0.886 & 0.811 & -0.261 & 0.716 \end{pmatrix}$$

Variance/covariance (S):

$I := \text{identity}(n)$

$$S := \frac{1}{n-1}\cdot X\cdot\left(I - \frac{1}{n}\cdot l\cdot l^T\right)\cdot X^T$$

$$S_Y := \frac{1}{n-1}\cdot Y\cdot\left(I - \frac{1}{n}\cdot l\cdot l^T\right)\cdot Y^T$$

It doesn't matter whether the > data matrix is mean-centered for calculating S

$$S = \begin{pmatrix} 4.308 & 1.684 & 1.803 & 2.155 & -0.253 \\ 1.684 & 1.767 & 0.588 & 0.178 & 0.176 \\ 1.803 & 0.588 & 0.801 & 1.065 & -0.158 \\ 2.155 & 0.178 & 1.065 & 1.969 & -0.357 \\ -0.253 & 0.176 & -0.158 & -0.357 & 0.504 \end{pmatrix} \qquad S_Y = \begin{pmatrix} 4.308 & 1.684 & 1.803 & 2.155 & -0.253 \\ 1.684 & 1.767 & 0.588 & 0.178 & 0.176 \\ 1.803 & 0.588 & 0.801 & 1.065 & -0.158 \\ 2.155 & 0.178 & 1.065 & 1.969 & -0.357 \\ -0.253 & 0.176 & -0.158 & -0.357 & 0.504 \end{pmatrix}$$

Calculate eigenvalues & eigenvectors of variance/covariance matrix (S):

$\Lambda := \text{reverse}(\text{sort}(\text{eigenvals}(S)))$

$E^{\langle j\rangle} := \text{eigenvec}(S, \Lambda_j)$

$$\Lambda = \begin{pmatrix} 6.931 \\ 1.785 \\ 0.39 \\ 0.23 \\ 0.014 \end{pmatrix} \qquad E = \begin{pmatrix} 0.78121 & 0.07087 & 0.00366 & -0.54171 & -0.30204 \\ 0.30565 & 0.76387 & -0.16182 & 0.5448 & -0.00928 \\ 0.33445 & -0.08291 & 0.01484 & -0.05102 & 0.93726 \\ 0.42601 & -0.57946 & 0.22045 & 0.63601 & -0.17215 \\ -0.05435 & 0.26236 & 0.96176 & -0.05128 & 0.02458 \end{pmatrix}$$

Testing the Identity SE = ED:

$I := \text{identity}(p)$

$D_{j,j} := \Lambda_j$

Diagonal Matrix of Eigenvalues (D) :

$$D = \begin{pmatrix} 6.931 & 0 & 0 & 0 & 0 \\ 0 & 1.785 & 0 & 0 & 0 \\ 0 & 0 & 0.39 & 0 & 0 \\ 0 & 0 & 0 & 0.23 & 0 \\ 0 & 0 & 0 & 0 & 0.014 \end{pmatrix}$$

$$S \cdot E = \begin{pmatrix} 5.41461 & 0.12652 & 0.00142 & -0.12434 & -0.00428 \\ 2.11847 & 1.36362 & -0.06305 & 0.12505 & -0.00013 \\ 2.31809 & -0.148 & 0.00578 & -0.01171 & 0.01327 \\ 2.95269 & -1.03442 & 0.0859 & 0.14598 & -0.00244 \\ -0.37673 & 0.46834 & 0.37475 & -0.01177 & 0.00035 \end{pmatrix}$$

$$E \cdot D = \begin{pmatrix} 5.41461 & 0.12652 & 0.00142 & -0.12434 & -0.00428 \\ 2.11847 & 1.36362 & -0.06305 & 0.12505 & -0.00013 \\ 2.31809 & -0.148 & 0.00578 & -0.01171 & 0.01327 \\ 2.95269 & -1.03442 & 0.0859 & 0.14598 & -0.00244 \\ -0.37673 & 0.46834 & 0.37475 & -0.01177 & 0.00035 \end{pmatrix}$$

Calculating Principal Component Scores (P) as linear combinations of M_c and E:

$P := M_c \cdot E$

$$P = \begin{pmatrix} 1.438 & 0.293 & 0.441 & -0.749 & -0.201 \\ -3.535 & 0.083 & -0.036 & 0.175 & -0.167 \\ -2.4 & -0.644 & -0.744 & -0.383 & 0.027 \\ -0.595 & 1.846 & 0.046 & -0.063 & 0.048 \\ 0.767 & 0.268 & -0.257 & 0.379 & 0.159 \\ 4.957 & -0.045 & 0.207 & 0.254 & 0.039 \\ -2.132 & -0.063 & -0.618 & -0.211 & -0.002 \\ 2.991 & 2.097 & -0.998 & 0.563 & -0.125 \\ 3.172 & -2.856 & 0.517 & 0.332 & -0.1 \\ -1.421 & -1.6 & -0.323 & -0.386 & -0.02 \\ -0.365 & -1.381 & -0.319 & 0.187 & 0.201 \\ -3.298 & 0.977 & 1.441 & 0.516 & 0.037 \\ -1.737 & -0.083 & 0.138 & 0.371 & -0.027 \\ 2.158 & 1.109 & 0.508 & -0.987 & 0.131 \end{pmatrix}$$

Calculating loadings of original variables on principal components:

NOTE:
Be certain that Eigenvalues (Λ) are in rank order and that Eigenvectors (E) are in columns in the same rank order. Otherwise calculations below will take into account the wrong entry in the variance / covariance matrix S.

$i := 1 .. p \quad k := 1 .. p$

$$\rho_{i,k} := \frac{E_{k,i} \cdot \sqrt{\Lambda_i}}{\sqrt{S_{k,k}}}$$ **jw Result 8.3, p. 429**

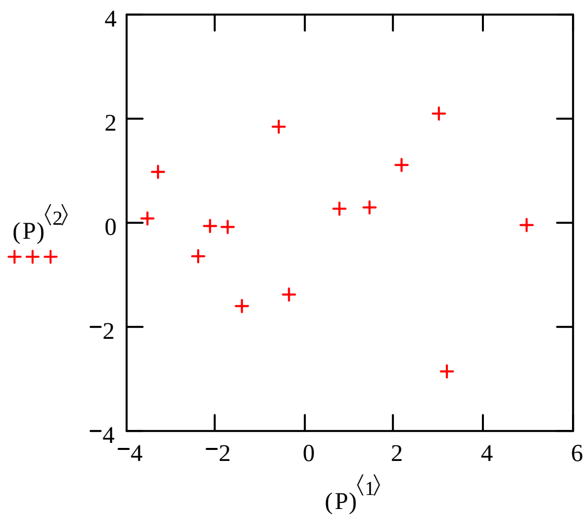
Eigenvectors (loadings):

$$E = \begin{pmatrix} 0.78121 & 0.07087 & 0.00366 & -0.54171 & -0.30204 \\ 0.30565 & 0.76387 & -0.16182 & 0.5448 & -0.00928 \\ 0.33445 & -0.08291 & 0.01484 & -0.05102 & 0.93726 \\ 0.42601 & -0.57946 & 0.22045 & 0.63601 & -0.17215 \\ -0.05435 & 0.26236 & 0.96176 & -0.05128 & 0.02458 \end{pmatrix}$$

Correlations with original variables:

$$\rho^T = \begin{pmatrix} 0.99095 & 0.04562 & 0.0011 & -0.12505 & -0.01731 \\ 0.60527 & 0.76768 & -0.07598 & 0.19633 & -0.00083 \\ 0.98402 & -0.1238 & 0.01035 & -0.02732 & 0.12462 \\ 0.79918 & -0.55168 & 0.09806 & 0.21712 & -0.01459 \\ -0.20149 & 0.49357 & 0.84533 & -0.03459 & 0.00412 \end{pmatrix}$$

Plot of component scores (P):



PC 2

PC 1

Singular Value Decomposition of M_c:

Display of squares matrices:

$$M_c \cdot M_c^T = \begin{pmatrix} 2.947 & -5.171 & -3.686 & -0.258 & 0.752 & 7.007 & -3.197 & 4.078 & 3.724 & -2.36 & -1.25 & -4.216 & -2.733 & 4.364 \\ -5.171 & 12.562 & 8.387 & 2.236 & -2.639 & -17.497 & 7.516 & -10.245 & -11.392 & 4.837 & 1.187 & 11.772 & 6.199 & -7.752 \\ -3.686 & 8.387 & 6.878 & 0.23 & -1.962 & -12.121 & 5.698 & -8.007 & -6.287 & 4.829 & 1.937 & 6.018 & 3.978 & -5.892 \\ -0.258 & 2.236 & 0.23 & 3.77 & 0.01 & -3.038 & 1.137 & 2.003 & -7.162 & -2.099 & -2.348 & 3.802 & 0.863 & 0.854 \\ 0.752 & -2.639 & -1.962 & 0.01 & 0.895 & 3.839 & -1.573 & 3.306 & 1.644 & -1.584 & -0.465 & -2.436 & -1.253 & 1.468 \\ 7.007 & -17.497 & -12.121 & -3.038 & 3.839 & 24.688 & -10.747 & 14.667 & 16.04 & -7.137 & -1.758 & -15.965 & -8.487 & 10.51 \\ -3.197 & 7.516 & 5.698 & 1.137 & -1.573 & -10.747 & 4.974 & -6.01 & -6.971 & 3.41 & 1.022 & 5.971 & 3.546 & -4.777 \\ 4.078 & -10.245 & -8.007 & 2.003 & 3.306 & 14.667 & -6.01 & 14.676 & 3.181 & -7.498 & -3.588 & -8.971 & -5.296 & 7.704 \\ 3.724 & -11.392 & -6.287 & -7.162 & 1.644 & 16.04 & -6.971 & 3.181 & 18.606 & -0.229 & 2.663 & -12.34 & -5.078 & 3.601 \\ -2.36 & 4.837 & 4.829 & -2.099 & -1.584 & -7.137 & 3.41 & -7.498 & -0.229 & 4.833 & 2.754 & 2.457 & 2.413 & -4.627 \\ -1.25 & 1.187 & 1.937 & -2.348 & -0.465 & -1.758 & 1.022 & -3.588 & 2.663 & 2.754 & 2.217 & -0.501 & 0.768 & -2.639 \\ -4.216 & 11.772 & 6.018 & 3.802 & -2.436 & -15.965 & 5.971 & -8.971 & -12.34 & 2.457 & -0.501 & 14.178 & 6.039 & -5.81 \\ -2.733 & 6.199 & 3.978 & 0.863 & -1.253 & -8.487 & 3.546 & -5.296 & -5.078 & 2.413 & 0.768 & 6.039 & 3.182 & -4.141 \\ 4.364 & -7.752 & -5.892 & 0.854 & 1.468 & 10.51 & -4.777 & 7.704 & 3.601 & -4.627 & -2.639 & -5.81 & -4.141 & 7.138 \end{pmatrix}$$

$$M_c^T \cdot M_c = \begin{pmatrix} 55.9982 & 21.8878 & 23.4361 & 28.0192 & -3.2952 \\ 21.8878 & 22.9771 & 7.6443 & 2.3137 & 2.2821 \\ 23.4361 & 7.6443 & 10.4087 & 13.8428 & -2.0584 \\ 28.0192 & 2.3137 & 13.8428 & 25.6032 & -4.6385 \\ -3.2952 & 2.2821 & -2.0584 & -4.6385 & 6.5569 \end{pmatrix} \qquad (n-1)\cdot S = \begin{pmatrix} 55.998 & 21.888 & 23.436 & 28.019 & -3.295 \\ 21.888 & 22.977 & 7.644 & 2.314 & 2.282 \\ 23.436 & 7.644 & 10.409 & 13.843 & -2.058 \\ 28.019 & 2.314 & 13.843 & 25.603 & -4.638 \\ -3.295 & 2.282 & -2.058 & -4.638 & 6.557 \end{pmatrix}$$

^ For mean-centered data there is a direct relationship between squares matrix $M_c^T M_c$ & variance/covariance matrix S.

Now for theSingular Value Decomposition:

$i := 1..n \qquad j := 1..p$

$SVD := svd(M_c)$

$U := submatrix(SVD, 1, n, 1, p)$

$V := submatrix(SVD, n+1, rows(SVD), 1, p)$

$\Lambda\Lambda := svds(M_c) \qquad \Lambda\Lambda d_{j,j} := \Lambda\Lambda_j$

Matrices displayed:

$$U = \begin{pmatrix} -0.151 & 0.061 & -0.196 & 0.433 & 0.469 \\ 0.372 & 0.017 & 0.016 & -0.102 & 0.39 \\ 0.253 & -0.134 & 0.331 & 0.222 & -0.064 \\ 0.063 & 0.383 & -0.021 & 0.037 & -0.112 \\ -0.081 & 0.056 & 0.114 & -0.22 & -0.37 \\ -0.522 & -0.009 & -0.092 & -0.147 & -0.092 \\ 0.225 & -0.013 & 0.275 & 0.122 & 0.005 \\ -0.315 & 0.435 & 0.443 & -0.326 & 0.292 \\ -0.334 & -0.593 & -0.23 & -0.192 & 0.233 \\ 0.15 & -0.332 & 0.144 & 0.223 & 0.046 \\ 0.038 & -0.287 & 0.142 & -0.108 & -0.469 \\ 0.347 & 0.203 & -0.64 & -0.299 & -0.087 \\ 0.183 & -0.017 & -0.061 & -0.215 & 0.063 \\ -0.227 & 0.23 & -0.226 & 0.571 & -0.306 \end{pmatrix}$$

$$V = \begin{pmatrix} -0.781 & 0.071 & -0.004 & 0.542 & 0.302 \\ -0.306 & 0.764 & 0.162 & -0.545 & 0.009 \\ -0.334 & -0.083 & -0.015 & 0.051 & -0.937 \\ -0.426 & -0.579 & -0.22 & -0.636 & 0.172 \\ 0.054 & 0.262 & -0.962 & 0.051 & -0.025 \end{pmatrix}$$

$$\Lambda\Lambda d = \begin{pmatrix} 9.492 & 0 & 0 & 0 & 0 \\ 0 & 4.817 & 0 & 0 & 0 \\ 0 & 0 & 2.251 & 0 & 0 \\ 0 & 0 & 0 & 1.727 & 0 \\ 0 & 0 & 0 & 0 & 0.429 \end{pmatrix}$$

Note:
In theory matrix U has dimensions nXn, V matrix is pXp, and ΛΛd is nXp, but in SVD columns beyond p in matrix U are zeroed out by the ΛΛd matrix with zero p+1 to n columns. MathCad's svd() and svds() functions only report columns of U and singular values in ΛΛd actually used in SVD.

Note that:

$$mean\left(U^{\langle j\rangle}\right) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad \left|U^{\langle j\rangle}\right| = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

Singular Value Decomposition verified:

$$M_c = \begin{pmatrix} 1.612 & 0.186 & 0.313 & 0.099 & 0.456 \\ -2.8 & -0.914 & -1.355 & -1.421 & 0.166 \\ -1.724 & -1.314 & -0.715 & -1.061 & -0.734 \\ -0.314 & 1.186 & -0.303 & -1.361 & 0.566 \\ 0.364 & 0.686 & 0.36 & 0.329 & -0.234 \\ 3.721 & 1.586 & 1.689 & 2.339 & -0.094 \\ -1.557 & -0.714 & -0.708 & -1.141 & -0.484 \\ 2.215 & 2.986 & 0.666 & 0.219 & -0.604 \\ 2.128 & -1.114 & 1.195 & 3.349 & -0.444 \\ -1.009 & -1.814 & -0.346 & 0.009 & -0.634 \\ -0.546 & -1.014 & 0.167 & 0.659 & -0.654 \\ -2.793 & -0.214 & -1.154 & -1.331 & 1.796 \\ -1.555 & -0.414 & -0.616 & -0.421 & 0.186 \\ 2.262 & 0.886 & 0.811 & -0.261 & 0.716 \end{pmatrix}$$

$$U \cdot \Lambda\Lambda d \cdot V^T = \begin{pmatrix} 1.612 & 0.186 & 0.313 & 0.099 & 0.456 \\ -2.8 & -0.914 & -1.355 & -1.421 & 0.166 \\ -1.724 & -1.314 & -0.715 & -1.061 & -0.734 \\ -0.314 & 1.186 & -0.303 & -1.361 & 0.566 \\ 0.364 & 0.686 & 0.36 & 0.329 & -0.234 \\ 3.721 & 1.586 & 1.689 & 2.339 & -0.094 \\ -1.557 & -0.714 & -0.708 & -1.141 & -0.484 \\ 2.215 & 2.986 & 0.666 & 0.219 & -0.604 \\ 2.128 & -1.114 & 1.195 & 3.349 & -0.444 \\ -1.009 & -1.814 & -0.346 & 0.009 & -0.634 \\ -0.546 & -1.014 & 0.167 & 0.659 & -0.654 \\ -2.793 & -0.214 & -1.154 & -1.331 & 1.796 \\ -1.555 & -0.414 & -0.616 & -0.421 & 0.186 \\ 2.262 & 0.886 & 0.811 & -0.261 & 0.716 \end{pmatrix}$$

The two methods compared:

PCA:

$$E = \begin{pmatrix} 0.781 & 0.071 & 0.004 & -0.542 & -0.302 \\ 0.306 & 0.764 & -0.162 & 0.545 & -0.009 \\ 0.334 & -0.083 & 0.015 & -0.051 & 0.937 \\ 0.426 & -0.579 & 0.22 & 0.636 & -0.172 \\ -0.054 & 0.262 & 0.962 & -0.051 & 0.025 \end{pmatrix}$$

SVD:

$$V = \begin{pmatrix} -0.781 & 0.071 & -0.004 & 0.542 & 0.302 \\ -0.306 & 0.764 & 0.162 & -0.545 & 0.009 \\ -0.334 & -0.083 & -0.015 & 0.051 & -0.937 \\ -0.426 & -0.579 & -0.22 & -0.636 & 0.172 \\ 0.054 & 0.262 & -0.962 & 0.051 & -0.025 \end{pmatrix}$$

^ Eigenvectors in the p-dimensional space of the variables are the same
(although + vs - *directions* of vectors matter in $U \cdot \Lambda\Lambda d \cdot V$)

$$\Lambda = \begin{pmatrix} 6.931 \\ 1.785 \\ 0.39 \\ 0.23 \\ 0.014 \end{pmatrix} \qquad (n-1)\cdot\Lambda = \begin{pmatrix} 90.104 \\ 23.207 \\ 5.065 \\ 2.984 \\ 0.184 \end{pmatrix} \qquad \Lambda\Lambda = \begin{pmatrix} 9.492 \\ 4.817 \\ 2.251 \\ 1.727 \\ 0.429 \end{pmatrix} \qquad \left(\Lambda\Lambda_j\right)^2 = \begin{pmatrix} 90.104 \\ 23.207 \\ 5.065 \\ 2.984 \\ 0.184 \end{pmatrix}$$

^ Eigenvalues of S are squared and scaled by 1/(n-1) compared with singular values
Since, from above, $M_c^T M_c = (n-1)\cdot S$ then $(n-1)\cdot\Lambda = \Lambda\Lambda^2$

This makes sense since S in PCA is a matrix of scaled sum of squares whereas matrix M_c in SVD is the original data.

Projection of mean-centered data onto eigenvectors:

PCA:

$$M_c \cdot E = \begin{pmatrix} 1.438 & 0.293 & 0.441 & -0.749 & -0.201 \\ -3.535 & 0.083 & -0.036 & 0.175 & -0.167 \\ -2.4 & -0.644 & -0.744 & -0.383 & 0.027 \\ -0.595 & 1.846 & 0.046 & -0.063 & 0.048 \\ 0.767 & 0.268 & -0.257 & 0.379 & 0.159 \\ 4.957 & -0.045 & 0.207 & 0.254 & 0.039 \\ -2.132 & -0.063 & -0.618 & -0.211 & -0.002 \\ 2.991 & 2.097 & -0.998 & 0.563 & -0.125 \\ 3.172 & -2.856 & 0.517 & 0.332 & -0.1 \\ -1.421 & -1.6 & -0.323 & -0.386 & -0.02 \\ -0.365 & -1.381 & -0.319 & 0.187 & 0.201 \\ -3.298 & 0.977 & 1.441 & 0.516 & 0.037 \\ -1.737 & -0.083 & 0.138 & 0.371 & -0.027 \\ 2.158 & 1.109 & 0.508 & -0.987 & 0.131 \end{pmatrix}$$

SVD:

$$U \cdot \Lambda\Lambda d = \begin{pmatrix} -1.438 & 0.293 & -0.441 & 0.749 & 0.201 \\ 3.535 & 0.083 & 0.036 & -0.175 & 0.167 \\ 2.4 & -0.644 & 0.744 & 0.383 & -0.027 \\ 0.595 & 1.846 & -0.046 & 0.063 & -0.048 \\ -0.767 & 0.268 & 0.257 & -0.379 & -0.159 \\ -4.957 & -0.045 & -0.207 & -0.254 & -0.039 \\ 2.132 & -0.063 & 0.618 & 0.211 & 0.002 \\ -2.991 & 2.097 & 0.998 & -0.563 & 0.125 \\ -3.172 & -2.856 & -0.517 & -0.332 & 0.1 \\ 1.421 & -1.6 & 0.323 & 0.386 & 0.02 \\ 0.365 & -1.381 & 0.319 & -0.187 & -0.201 \\ 3.298 & 0.977 & -1.441 & -0.516 & -0.037 \\ 1.737 & -0.083 & -0.138 & -0.371 & 0.027 \\ -2.158 & 1.109 & -0.508 & 0.987 & -0.131 \end{pmatrix}$$

^ PCA projections ($P = M_c \cdot E = M_c \cdot V$) are the same as scaling SVD matrix U by singular values $\Lambda\Lambda d$.

Rows of U give directionality for objects in PCA space; when each column is multiplied by the appropriate singular value, points (objects) are correctly placed in a PCA plot.

^ In Singular Value Decomposition: $M_c = U \cdot \Lambda\Lambda d \cdot V$:

- **Rows of U may be interpreted as directions of objects loaded onto eigenvectors comprising PCA space.**
- **Columns of V represent eigenvector directions of PCA space in relation to the original variables.**
- **Singular values represent scaling factors.**

2005 PCA & SVD

PCA as a spectral decomposition of S:

$$\sum_j \Lambda_j \cdot E^{\langle j \rangle} \cdot E^{\langle j \rangle T} = \begin{pmatrix} 4.308 & 1.684 & 1.803 & 2.155 & -0.253 \\ 1.684 & 1.767 & 0.588 & 0.178 & 0.176 \\ 1.803 & 0.588 & 0.801 & 1.065 & -0.158 \\ 2.155 & 0.178 & 1.065 & 1.969 & -0.357 \\ -0.253 & 0.176 & -0.158 & -0.357 & 0.504 \end{pmatrix} \qquad S = \begin{pmatrix} 4.308 & 1.684 & 1.803 & 2.155 & -0.253 \\ 1.684 & 1.767 & 0.588 & 0.178 & 0.176 \\ 1.803 & 0.588 & 0.801 & 1.065 & -0.158 \\ 2.155 & 0.178 & 1.065 & 1.969 & -0.357 \\ -0.253 & 0.176 & -0.158 & -0.357 & 0.504 \end{pmatrix}$$

first component in spectral decomposition:

$$\Lambda_1 \cdot E^{\langle 1 \rangle} \cdot E^{\langle 1 \rangle T} = \begin{pmatrix} 4.23 & 1.655 & 1.811 & 2.307 & -0.294 \\ 1.655 & 0.648 & 0.709 & 0.902 & -0.115 \\ 1.811 & 0.709 & 0.775 & 0.988 & -0.126 \\ 2.307 & 0.902 & 0.988 & 1.258 & -0.16 \\ -0.294 & -0.115 & -0.126 & -0.16 & 0.02 \end{pmatrix}$$

< Components represent partials of S that together sum to S. Thus each component can be viewed as representing the variance/covariance of the data consulting only the appropriate eigenvector/eigenvalue pairs in canonical directions from greatest to least variability in the data.

Same problem as singlular value decomposition M_c:

$$\sum_j U^{\langle j \rangle} \cdot \Lambda\Lambda_j \cdot V^{\langle j \rangle T} = \begin{pmatrix} 1.612 & 0.186 & 0.313 & 0.099 & 0.456 \\ -2.8 & -0.914 & -1.355 & -1.421 & 0.166 \\ -1.724 & -1.314 & -0.715 & -1.061 & -0.734 \\ -0.314 & 1.186 & -0.303 & -1.361 & 0.566 \\ 0.364 & 0.686 & 0.36 & 0.329 & -0.234 \\ 3.721 & 1.586 & 1.689 & 2.339 & -0.094 \\ -1.557 & -0.714 & -0.708 & -1.141 & -0.484 \\ 2.215 & 2.986 & 0.666 & 0.219 & -0.604 \\ 2.128 & -1.114 & 1.195 & 3.349 & -0.444 \\ -1.009 & -1.814 & -0.346 & 0.009 & -0.634 \\ -0.546 & -1.014 & 0.167 & 0.659 & -0.654 \\ -2.793 & -0.214 & -1.154 & -1.331 & 1.796 \\ -1.555 & -0.414 & -0.616 & -0.421 & 0.186 \\ 2.262 & 0.886 & 0.811 & -0.261 & 0.716 \end{pmatrix} \qquad M_c = \begin{pmatrix} 1.612 & 0.186 & 0.313 & 0.099 & 0.456 \\ -2.8 & -0.914 & -1.355 & -1.421 & 0.166 \\ -1.724 & -1.314 & -0.715 & -1.061 & -0.734 \\ -0.314 & 1.186 & -0.303 & -1.361 & 0.566 \\ 0.364 & 0.686 & 0.36 & 0.329 & -0.234 \\ 3.721 & 1.586 & 1.689 & 2.339 & -0.094 \\ -1.557 & -0.714 & -0.708 & -1.141 & -0.484 \\ 2.215 & 2.986 & 0.666 & 0.219 & -0.604 \\ 2.128 & -1.114 & 1.195 & 3.349 & -0.444 \\ -1.009 & -1.814 & -0.346 & 0.009 & -0.634 \\ -0.546 & -1.014 & 0.167 & 0.659 & -0.654 \\ -2.793 & -0.214 & -1.154 & -1.331 & 1.796 \\ -1.555 & -0.414 & -0.616 & -0.421 & 0.186 \\ 2.262 & 0.886 & 0.811 & -0.261 & 0.716 \end{pmatrix}$$

first component of singular value decomposition:

$$U^{\langle 1 \rangle} \cdot \Lambda\Lambda_1 \cdot V^{\langle 1 \rangle T} = \begin{pmatrix} 1.123 & 0.439 & 0.481 & 0.612 & -0.078 \\ -2.761 & -1.08 & -1.182 & -1.506 & 0.192 \\ -1.875 & -0.734 & -0.803 & -1.023 & 0.13 \\ -0.465 & -0.182 & -0.199 & -0.254 & 0.032 \\ 0.599 & 0.234 & 0.256 & 0.327 & -0.042 \\ 3.873 & 1.515 & 1.658 & 2.112 & -0.269 \\ -1.665 & -0.652 & -0.713 & -0.908 & 0.116 \\ 2.337 & 0.914 & 1 & 1.274 & -0.163 \\ 2.478 & 0.969 & 1.061 & 1.351 & -0.172 \\ -1.11 & -0.434 & -0.475 & -0.605 & 0.077 \\ -0.285 & -0.112 & -0.122 & -0.155 & 0.02 \\ -2.577 & -1.008 & -1.103 & -1.405 & 0.179 \\ -1.357 & -0.531 & -0.581 & -0.74 & 0.094 \\ 1.686 & 0.66 & 0.722 & 0.92 & -0.117 \end{pmatrix}$$

< Components represent partials of M_c that together sum to M_c = mean-centered object locations in original variable space. Thus each component can be viewed as partial locations of objects (points) consulting only the appropriate eigenvector from greatest to least difference in U rotated to the space of the original variables in M_c by each eigenvector's direction in V.

Prepared by:
Wm Stein

ORIGIN ≡ 1

PRINCIPAL COORDINATES ANALYSIS

The 1970 Census provided tract information on 5 socioeconomic variables for the Madison, Wisc. area. The data for 14 tracts are listed in J.W. Table 8.5, p. 470.

Read in Data:

$M := READPRN("DATA\backslash T8\text{-}5.DAT")$

M =

	1	2	3	4	5
1	5.935	14.2	2.265	2.27	2.91
2	1.523	13.1	0.597	0.75	2.62
3	2.599	12.7	1.237	1.11	1.72
4	4.009	15.2	1.649	0.81	3.02
5	4.687	14.7	2.312	2.5	2.22
6	8.044	15.6	3.641	4.51	2.36
7	2.766	13.3	1.244	1.03	1.97
8	6.538	17	2.618	2.39	1.85
9	6.451	12.9	3.147	5.52	2.01
10	3.314	12.2	1.606	2.18	1.82
11	3.777	13	2.119	2.83	1.8
12	1.53	13.8	0.798	0.84	4.25

Sizing data array:

$X := M^T$

$n := cols(X) \quad n = 14 \quad p := rows(X) \quad p = 5$

$i := 1 .. n \quad j := 1 .. p$

$l_i := 1$

Mean vector (X_{bar}):

$$X_{bar} := \frac{1}{n} \cdot X \cdot l \qquad X_{bar} = \begin{pmatrix} 4.323 \\ 14.014 \\ 1.952 \\ 2.171 \\ 2.454 \end{pmatrix}$$

Variance/covariance (S):

$I := identity(n)$

$$S := \frac{1}{n-1} \cdot X \cdot \left(I - \frac{1}{n} \cdot l \cdot l^T \right) \cdot X^T$$

$$S = \begin{pmatrix} 4.308 & 1.684 & 1.803 & 2.155 & -0.253 \\ 1.684 & 1.767 & 0.588 & 0.178 & 0.176 \\ 1.803 & 0.588 & 0.801 & 1.065 & -0.158 \\ 2.155 & 0.178 & 1.065 & 1.969 & -0.357 \\ -0.253 & 0.176 & -0.158 & -0.357 & 0.504 \end{pmatrix}$$

Calculating Squared Euclidean Distances between the objects in M:

On many occasions, data is collected in distance form - as for instance, base differences in alligned nucleotide sequences. As long as "distances" utilized obey the triangle inequality, PCO allows ordination of the data and plotting in a way that is analogous to PCA.

$i := 1 .. cols(X) \quad k := 1 .. cols(X)$

$$D_{i,k} := \left| \left(X^{\langle i \rangle} - X^{\langle k \rangle} \right)^T \cdot \left(X^{\langle i \rangle} - X^{\langle k \rangle} \right) \right|$$

D =

0	25.852	17.197	7.233	2.339	13.621	14.316	9.466	14.107	12.5	7.664	25.557	11.596	1.358
25.852	0	2.667	11.861	18.735	72.245	2.505	47.727	53.953	7.721	12.406	3.195	3.347	35.204
17.197	2.667	0	10.188	11.697	55.807	0.457	37.568	38.058	2.052	5.22	9.019	2.104	25.799
7.233	11.861	10.188	0	4.645	34.535	6.47	14.44	36.702	12.802	10.684	10.343	5.226	9.199
2.339	18.735	11.697	4.645	0	17.905	9.014	8.959	16.213	8.896	4.041	19.945	6.584	5.096
13.621	72.245	55.807	34.535	17.905	0	51.155	10.029	11.214	43.795	30.42	70.796	44.845	10.806
14.316	2.505	0.457	6.47	9.014	51.155	0	31.67	37.522	2.986	5.147	7.211	1.066	21.667
9.466	47.727	37.568	14.44	8.959	10.029	31.67	0	26.92	34.503	24.068	46.795	28.45	6.406
14.107	53.953	38.058	36.702	16.213	11.214	37.522	26.92	0	23.897	15.497	57.464	31.944	18.543
12.5	7.721	2.052	12.802	8.896	43.795	2.986	34.503	23.897	0	1.54	14.096	3.188	21.223
7.664	12.406	5.22	10.684	4.041	30.42	5.147	24.068	15.497	1.54	0	17.397	3.863	14.633
25.557	3.195	9.019	10.343	19.945	70.796	7.211	46.795	57.464	14.096	17.397	0	5.282	32.936
11.596	3.347	2.104	5.226	6.584	44.845	1.066	28.45	31.944	3.188	3.863	5.282	0	18.602
1.358	35.204	25.799	9.199	5.096	10.806	21.667	6.406	18.543	21.223	14.633	32.936	18.602	0

2005 PCO

Transforming distances to produce matrix A:

$$A := -0.5D$$

$$A_{bar} := \frac{1}{n} \cdot A \cdot 1$$

$$\Delta_{i,k} := A_{i,k} - A_{bar_i} - A_{bar_k} + \text{mean}(A)$$

$$A_{bar} = \begin{pmatrix} -5.815 \\ -10.622 \\ -7.78 \\ -6.226 \\ -4.788 \\ -16.685 \\ -6.828 \\ -11.679 \\ -13.644 \\ -6.757 \\ -5.449 \\ -11.43 \\ -5.932 \\ -7.91 \end{pmatrix}$$

$$\Delta = \begin{pmatrix} 2.947 & -5.171 & -3.686 & -0.258 & 0.752 & 7.007 & -3.197 & 4.078 & 3.724 & -2.36 & -1.25 & -4.216 & -2.733 & 4.364 \\ -5.171 & 12.562 & 8.387 & 2.236 & -2.639 & -17.497 & 7.516 & -10.245 & -11.392 & 4.837 & 1.187 & 11.772 & 6.199 & -7.752 \\ -3.686 & 8.387 & 6.878 & 0.23 & -1.962 & -12.121 & 5.698 & -8.007 & -6.287 & 4.829 & 1.937 & 6.018 & 3.978 & -5.892 \\ -0.258 & 2.236 & 0.23 & 3.77 & 0.01 & -3.038 & 1.137 & 2.003 & -7.162 & -2.099 & -2.348 & 3.802 & 0.863 & 0.854 \\ 0.752 & -2.639 & -1.962 & 0.01 & 0.895 & 3.839 & -1.573 & 3.306 & 1.644 & -1.584 & -0.465 & -2.436 & -1.253 & 1.468 \\ 7.007 & -17.497 & -12.121 & -3.038 & 3.839 & 24.688 & -10.747 & 14.667 & 16.04 & -7.137 & -1.758 & -15.965 & -8.487 & 10.51 \\ -3.197 & 7.516 & 5.698 & 1.137 & -1.573 & -10.747 & 4.974 & -6.01 & -6.971 & 3.41 & 1.022 & 5.971 & 3.546 & -4.777 \\ 4.078 & -10.245 & -8.007 & 2.003 & 3.306 & 14.667 & -6.01 & 14.676 & 3.181 & -7.498 & -3.588 & -8.971 & -5.296 & 7.704 \\ 3.724 & -11.392 & -6.287 & -7.162 & 1.644 & 16.04 & -6.971 & 3.181 & 18.606 & -0.229 & 2.663 & -12.34 & -5.078 & 3.601 \\ -2.36 & 4.837 & 4.829 & -2.099 & -1.584 & -7.137 & 3.41 & -7.498 & -0.229 & 4.833 & 2.754 & 2.457 & 2.413 & -4.627 \\ -1.25 & 1.187 & 1.937 & -2.348 & -0.465 & -1.758 & 1.022 & -3.588 & 2.663 & 2.754 & 2.217 & -0.501 & 0.768 & -2.639 \\ -4.216 & 11.772 & 6.018 & 3.802 & -2.436 & -15.965 & 5.971 & -8.971 & -12.34 & 2.457 & -0.501 & 14.178 & 6.039 & -5.81 \\ -2.733 & 6.199 & 3.978 & 0.863 & -1.253 & -8.487 & 3.546 & -5.296 & -5.078 & 2.413 & 0.768 & 6.039 & 3.182 & -4.141 \\ 4.364 & -7.752 & -5.892 & 0.854 & 1.468 & 10.51 & -4.777 & 7.704 & 3.601 & -4.627 & -2.639 & -5.81 & -4.141 & 7.138 \end{pmatrix}$$

2005 PCO

Calculate eigenvalues & eigenvectors of matrix Δ:

$\Lambda := \text{reverse}(\text{sort}(\text{eigenvals}(\Delta)))$

$E^{\langle j \rangle} := \text{eigenvec}(\Delta, \Lambda_j)$

Remember:

Eigenvectors are automatically scaled to unit length!

$\left| E^{\langle 1 \rangle} \right| = 1$

$\Lambda =$

90.104
23.207
5.065
2.984
0.184
5.898×10^{-15}
3.242×10^{-15}
1.179×10^{-15}
1.179×10^{-15}
0
-1.176×10^{-15}
-1.982×10^{-15}
-4.352×10^{-15}
-4.352×10^{-15}

$E =$

−0.15145	−0.06074	−0.19572	0.43333	−0.46903
0.37239	−0.01715	0.01617	−0.10156	−0.38977
0.25286	0.13377	0.33077	0.22166	0.06355
0.06271	−0.38318	−0.02063	0.03658	0.11175
−0.08077	−0.05561	0.11427	−0.21951	0.36973
−0.52226	0.00931	−0.09195	−0.14727	0.09172
0.22457	0.01306	0.2746	0.12188	−0.00529
−0.31513	−0.43536	0.4435	−0.32588	−0.2918
−0.33415	0.59293	−0.22959	−0.19214	−0.23294
0.14967	0.33215	0.14369	0.22348	−0.0465
0.03844	0.28659	0.14195	−0.1084	0.46906
0.34749	−0.20274	−0.64019	−0.29896	0.08661
0.18303	0.01716	−0.06128	−0.2146	−0.06271
−0.22739	−0.23019	−0.22557	0.57137	0.30561

Scaling the Eigenvectors of Δ by the square root of non-zero eigenvalues in Λ:

$EE_{i,j} := E_{i,j} \cdot \sqrt{\Lambda_j}$

$\left| EE^{\langle 1 \rangle} \right| = 9.492$

$EE =$

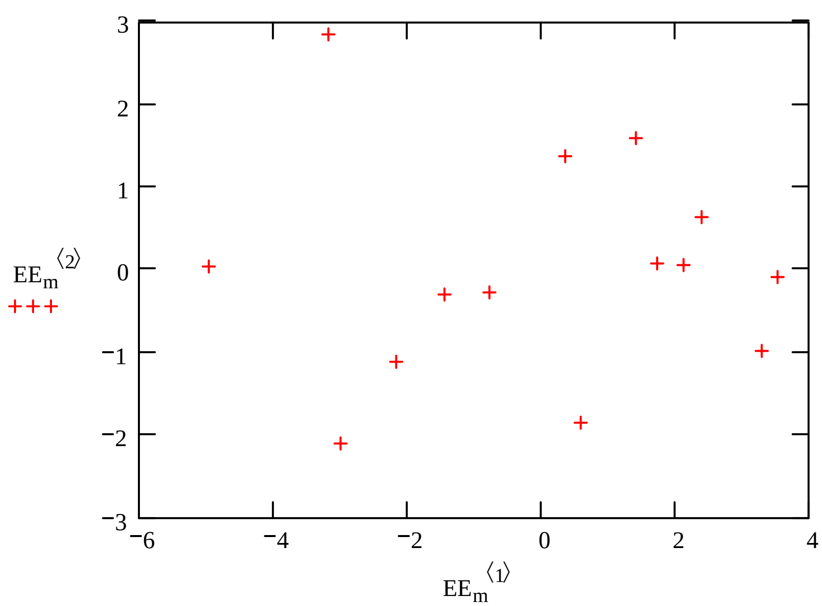
−1.438	−0.293	−0.441	0.749	−0.201
3.535	−0.083	0.036	−0.175	−0.167
2.4	0.644	0.744	0.383	0.027
0.595	−1.846	−0.046	0.063	0.048
−0.767	−0.268	0.257	−0.379	0.159
−4.957	0.045	−0.207	−0.254	0.039
2.132	0.063	0.618	0.211	−0.002
−2.991	−2.097	0.998	−0.563	−0.125
−3.172	2.856	−0.517	−0.332	−0.1
1.421	1.6	0.323	0.386	−0.02
0.365	1.381	0.319	−0.187	0.201
3.298	−0.977	−1.441	−0.516	0.037
1.737	0.083	−0.138	−0.371	−0.027
−2.158	−1.109	−0.508	0.987	0.131

Scaled eigenvectors (matrix EE) represent the object coordintates in PCO space.

Each row represents an object. As in PCA, choose only the first few PCO coordinates to represent the data. Here, the first two coordinates would seem to be enough:

$m := 1 .. 2$

$EEm_{i,m} := EE_{i,m}$



Compare this plot with the PCA plot of the same data in the PCA & SVD worksheet.

Note that PCA/PCO directions & therefore handedness of the coordinate system for the plots may differ, but are unimportant.

2005 PCO

Comparing PCO and Singular Value Decomposition (SVD):

(see the PCA SVD.mcd worksheet)

$p := 1 .. rows(X)$

$$\Lambda_j = \begin{pmatrix} 90.104 \\ 23.207 \\ 5.065 \\ 2.984 \\ 0.184 \end{pmatrix}$$

The first p (= original dimenion of variables in the data) Eigenvalues of matrix Δ - a transformation of original distance matrix D - represent the SQUARE of the singular values (ΛΛ) in SVD).

$$E = \begin{pmatrix} -0.151 & -0.061 & -0.196 & 0.433 & -0.469 \\ 0.372 & -0.017 & 0.016 & -0.102 & -0.39 \\ 0.253 & 0.134 & 0.331 & 0.222 & 0.064 \\ 0.063 & -0.383 & -0.021 & 0.037 & 0.112 \\ -0.081 & -0.056 & 0.114 & -0.22 & 0.37 \\ -0.522 & 0.009 & -0.092 & -0.147 & 0.092 \\ 0.225 & 0.013 & 0.275 & 0.122 & -0.005 \\ -0.315 & -0.435 & 0.443 & -0.326 & -0.292 \\ -0.334 & 0.593 & -0.23 & -0.192 & -0.233 \\ 0.15 & 0.332 & 0.144 & 0.223 & -0.046 \\ 0.038 & 0.287 & 0.142 & -0.108 & 0.469 \\ 0.347 & -0.203 & -0.64 & -0.299 & 0.087 \\ 0.183 & 0.017 & -0.061 & -0.215 & -0.063 \\ -0.227 & -0.23 & -0.226 & 0.571 & 0.306 \end{pmatrix}$$

The matrix of Eigenvectors in PCO corresponds to the matrix U in SVD.

$$EE = \begin{pmatrix} -1.438 & -0.293 & -0.441 & 0.749 & -0.201 \\ 3.535 & -0.083 & 0.036 & -0.175 & -0.167 \\ 2.4 & 0.644 & 0.744 & 0.383 & 0.027 \\ 0.595 & -1.846 & -0.046 & 0.063 & 0.048 \\ -0.767 & -0.268 & 0.257 & -0.379 & 0.159 \\ -4.957 & 0.045 & -0.207 & -0.254 & 0.039 \\ 2.132 & 0.063 & 0.618 & 0.211 & -0.002 \\ -2.991 & -2.097 & 0.998 & -0.563 & -0.125 \\ -3.172 & 2.856 & -0.517 & -0.332 & -0.1 \\ 1.421 & 1.6 & 0.323 & 0.386 & -0.02 \\ 0.365 & 1.381 & 0.319 & -0.187 & 0.201 \\ 3.298 & -0.977 & -1.441 & -0.516 & 0.037 \\ 1.737 & 0.083 & -0.138 & -0.371 & -0.027 \\ -2.158 & -1.109 & -0.508 & 0.987 & 0.131 \end{pmatrix}$$

Eigenvectors of matrix Δ SCALED by their associated SINGULAR VALUE lengths (= squareroot of Λ_j) represent a projection of mean centered data ($M_cE = M_cV = U \cdot \Lambda\Lambda d$) onto the eigenvectors in PCA.

2005 PCO Map Example

Prepared by:
Wm Stein

ORIGIN ≡ 1

PRINCIPAL COORDINATES ANALYSIS
MAP EXAMPLE

I measured linear distances (in mm) between cities on a map of New York State:

BUR - Burlington VT
ALB - Albany
BNG - Binghamton
WHP - White Plains
BUF - Buffalo
ITH - Ithaca
JAM - Jamestown

BUR	ALB	BNG	WHP	BUF	ITH	JAM
ALB	▪	▪	▪	▪	▪	▪
BNG	▪	▪	▪	▪	▪	▪
WHP	▪	▪	▪	▪	▪	▪
BUF	▪	▪	▪	▪	▪	▪
ITH	▪	▪	▪	▪	▪	▪
JAM	▪	▪	▪	▪	▪	▪

Read in Data:

D := READPRN("DATA\NYMap.DAT")

Sizing data array:

$n := cols(D) \quad n = 7$

$i := 1..n \quad k := 1..n$

$l_i := 1$

$$D = \begin{pmatrix} 0 & 92 & 154 & 170 & 214 & 153 & 244 \\ 92 & 0 & 81 & 80 & 183 & 98.5 & 200 \\ 154 & 81 & 0 & 93 & 112.5 & 26.5 & 122 \\ 170 & 80 & 93 & 0 & 204 & 120 & 207 \\ 214 & 183 & 112.5 & 204 & 0 & 87 & 40 \\ 153 & 98.5 & 26.5 & 120 & 87 & 0 & 101 \\ 244 & 200 & 122 & 207 & 40 & 101 & 0 \end{pmatrix}$$

Squaring Distances and transforming to produce matrices A & Δ:

$D_{sq_{i,k}} := \left(D_{i,k}\right)^2$

$A := -0.5 D_{sq}$

$A_{bar} := \frac{1}{n} \cdot A \cdot l$

$\Delta_{i,k} := A_{i,k} - A_{bar_i} - A_{bar_k} + mean(A)$

$$A_{bar} = \begin{pmatrix} -1.356 \times 10^4 \\ -7472.589 \\ -4797.75 \\ -1.02 \times 10^4 \\ -1.019 \times 10^4 \\ -4713.107 \\ -1.208 \times 10^4 \end{pmatrix}$$

$$\Delta = \begin{pmatrix} 18115.25 & 7797.196 & -2503.643 & 307.607 & -8146.661 & -2434.786 & -13134.964 \\ 7797.196 & 5943.143 & -12.196 & 5471.554 & -8079.214 & -1667.464 & -9453.018 \\ -2503.643 & -12.196 & 593.464 & 1672.214 & -337.679 & 157.696 & 430.143 \\ 307.607 & 5471.554 & 1672.214 & 11399.964 & -9414.304 & -1287.929 & -8149.107 \\ -8146.661 & -8079.214 & -337.679 & -9414.304 & 11387.429 & 2121.304 & 12469.125 \\ -2434.786 & -1667.464 & 157.696 & -1287.929 & 2121.304 & 424.179 & 2687 \\ -13134.964 & -9453.018 & 430.143 & -8149.107 & 12469.125 & 2687 & 15150.821 \end{pmatrix}$$

2005 PCO Map Example

Calculate eigenvalues & eigenvectors of matrix Δ:

$$\Lambda := \mathrm{reverse}(\mathrm{sort}(\mathrm{eigenvals}(\Delta)))$$

$$E^{\langle i \rangle} := \mathrm{eigenvec}(\Delta, \Lambda_i)$$

$$\Lambda = \begin{pmatrix} 47659.872 \\ 15306.353 \\ 265.653 \\ 87.587 \\ 5.802 \times 10^{-13} \\ -48.653 \\ -256.562 \end{pmatrix}$$

Remember:

Eigenvectors are automatically scaled to unit length!

$$\left|E^{\langle 1 \rangle}\right| = 1$$

$$E = \begin{pmatrix} -0.48511 & -0.67134 & -0.28583 & 0.07421 & 0.37796 & 0.03901 & 0.28703 \\ -0.35305 & 0.03599 & 0.51771 & 0.1939 & 0.37796 & 0.29023 & -0.58426 \\ 0.01709 & 0.1963 & 0.56258 & -0.31599 & 0.37796 & -0.01212 & 0.63389 \\ -0.30789 & 0.66668 & -0.54369 & -0.10069 & 0.37796 & 0.10965 & 0.01092 \\ 0.46676 & -0.25426 & -0.16649 & -0.64022 & 0.37796 & 0.17293 & -0.32729 \\ 0.09892 & 0.01588 & 0.01662 & 0.14346 & 0.37796 & -0.88822 & -0.19317 \\ 0.56328 & 0.01075 & -0.1009 & 0.64532 & 0.37796 & 0.28851 & 0.17288 \end{pmatrix}$$

Scaling the Eigenvectors of Δ by the square root of non-zero eigenvalues in Λ:

$$EE_{i,k} := E_{i,k} \cdot \sqrt{\Lambda_k} \qquad \left|EE^{\langle 1 \rangle}\right| = 218.311$$

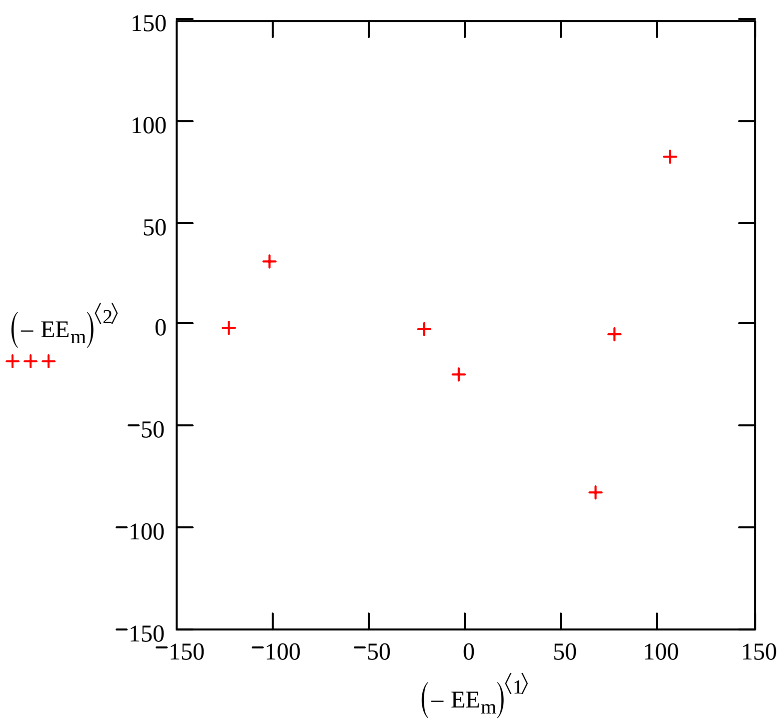
$$EE = \begin{pmatrix} -105.905 & -83.057 & -4.659 & 0.695 & 2.879 \times 10^{-7} & 0.272i & 4.598i \\ -77.075 & 4.452 & 8.438 & 1.815 & 2.879 \times 10^{-7} & 2.024i & -9.358i \\ 3.73 & 24.286 & 9.169 & -2.957 & 2.879 \times 10^{-7} & -0.085i & 10.153i \\ -67.217 & 82.481 & -8.862 & -0.942 & 2.879 \times 10^{-7} & 0.765i & 0.175i \\ 101.9 & -31.457 & -2.714 & -5.992 & 2.879 \times 10^{-7} & 1.206i & -5.242i \\ 21.596 & 1.965 & 0.271 & 1.343 & 2.879 \times 10^{-7} & -6.195i & -3.094i \\ 122.97 & 1.33 & -1.645 & 6.039 & 2.879 \times 10^{-7} & 2.012i & 2.769i \end{pmatrix}$$

Scaled eigenvectors (matrix EE) represent the object coordintates in PCO space.

Remember, the original dimensionality of the problem was a two-dimensional map...

$$m := 1..2$$

$$EEm_{i,m} := EE_{i,m}$$



So, from a distance matrix (D), one can use PCO to reconstruct a reasonable approximation for map coordinates...

PCO 3 (small in comparison with PCO 1 & 2) might be used to graphically represent an error-in-plot term.

I turned around both PCO axes to allow representation of the map in its usual orientation and I fixed the graph axes to plot equal scales. Note, however, that I didn't correct for map rotation! North does not correspond to any PCO axis.

2005 jw606

ORIGIN ≡ 1

CLASSIFICATION AND DISCRIMANT FUNCTIONS USING LINEAR AND QUADRATIC ECM RULES
jw606.mcd

prepared by:
Wm Stein

Table 11.2, p. 607 Salmon Data for Example 11.7

M := READPRN("\DATA\T11-2.dat")

$i := 1..50 \quad j := 1..2$

$U_{i,j} := M_{i,j+2}$

$C_{i,j} := M_{i+50,j+2}$

Summary Statistics:

$n_1 := 50 \quad n_2 := 50$

$l_{n_i} := 1$

Mean vectors:

$U_{bar} := \frac{1}{n_1} \cdot U^T \cdot l_n$

$C_{bar} := \frac{1}{n_2} \cdot C^T \cdot l_n$

Variance-covariance matrices:

$I := \text{identity}(n_1)$

US Salmon

U =

	1	2
1	108	368
2	131	355
3	105	469
4	86	506
5	99	402
6	87	423
7	94	440
8	117	489
9	79	432
10	99	403
11	114	428
12	123	372
13	123	372
14	109	420
15	112	394
16	104	407

Canadian Salmon

C =

	1	2
1	129	420
2	148	371
3	179	407
4	152	381
5	166	377
6	124	389
7	156	419
8	131	345
9	140	362
10	144	345
11	149	393
12	108	330
13	135	355
14	170	386
15	152	301
16	153	397

Total Data Set

M =

	1	2	3	4
1	1	2	108	368
2	1	1	131	355
3	1	1	105	469
4	1	2	86	506
5	1	1	99	402
6	1	2	87	423
7	1	1	94	440
8	1	2	117	489
9	1	2	79	432
10	1	1	99	403
11	1	1	114	428
12	1	2	123	372
13	1	1	123	372
14	1	2	109	420
15	1	2	112	394
16	1	1	104	407

$S_U := \frac{1}{n_1 - 1} \cdot U^T \cdot \left(I - \frac{1}{n_1} \cdot l_n \cdot l_n^T\right) \cdot U$

$U_{bar} = \begin{pmatrix} 98.38 \\ 429.66 \end{pmatrix}$ **< verified p. 658 >** $C_{bar} = \begin{pmatrix} 137.46 \\ 366.62 \end{pmatrix}$

$S_C := \frac{1}{n_1 - 1} \cdot C^T \cdot \left(I - \frac{1}{n_2} \cdot l_n \cdot l_n^T\right) \cdot C$

$S_U = \begin{pmatrix} 260.60776 & -188.09265 \\ -188.09265 & 1399.08612 \end{pmatrix}$ $S_C = \begin{pmatrix} 326.0902 & 133.5049 \\ 133.5049 & 893.26082 \end{pmatrix}$

$S_p := \frac{n_1 - 1}{n_1 + n_2 - 2} \cdot S_U + \frac{n_2 - 1}{n_1 + n_2 - 2} \cdot S_C$

$S_p = \begin{pmatrix} 293.34898 & -27.29388 \\ -27.29388 & 1146.17347 \end{pmatrix}$ **< pooled variance-covariance matrix Eq. 6-21 p. 284 & Eq. 11-17 p. 592**

Estimated LINEAR Minimum Expected cost of Misclassification (ECM) Rule - Eq 11-18
Assuming equal costs of misclassification, equal prior probabilities, and common covariance matrix Σ:

discriminant coefficients a_{hat} - Eq. 11-19:

$a_{hat} := (U_{bar} - C_{bar})^T \cdot S_p^{-1}$ $a_{hat} = (-0.12839 \quad 0.05194)$ **< values check with jw p. 608.**

linear transformation of means:

$y_{barU} := (U_{bar} - C_{bar})^T \cdot S_p^{-1} \cdot U_{bar}$ $y_{barU} = (9.68714)$

$y_{barC} := (U_{bar} - C_{bar})^T \cdot S_p^{-1} \cdot C_{bar}$ $y_{barC} = (1.39527)$

midpoint:

$m := \frac{1}{2} \cdot (U_{bar} - C_{bar})^T \cdot S_p^{-1} \cdot (U_{bar} + C_{bar})$ $m = (5.5412)$ $\frac{1}{2} \cdot (y_{barU} + y_{barC}) = (5.5412)$

^ value for m checks with jw p. 608

2005 jw606

Assigning training set to classes:

$i := 1 .. 100$

$MM_{i,j} := M_{i,j+2}$

$class_i := \left[a_{hat} \cdot \left(MM^T\right)^{\langle i \rangle} \right]_1$

$zero_i := class_i - m_1$

Solving for discrimination boundary:

$a_{hat}'x_0 = m$ < from Eq. 11-18
Lets suggest some values for the second coordinate for x_0

$$x_2 := \begin{pmatrix} 350 \\ 400 \\ 450 \\ 500 \end{pmatrix}$$

$$x_1 := \frac{m_1 - \left(a_{hat}^T\right)_2 \cdot x_2}{\left(a_{hat}^T\right)_1}$$

$X := \text{augment}(x_1, x_2)$

MM =

	1	2
1	108	368
2	131	355
3	105	469
4	86	506
5	99	402
6	87	423
7	94	440
8	117	489
9	79	432
10	99	403
11	114	428
12	123	372
13	123	372
14	109	420
15	112	394
16	104	407

class =

	1
1	5.24924
2	1.62107
3	10.88066
4	15.24191
5	8.17079
6	10.80224
7	10.78657
8	10.37887
9	12.29683
10	8.22273
11	7.5955
12	3.5312
13	3.5312
14	7.82189
15	6.08621
16	7.78857

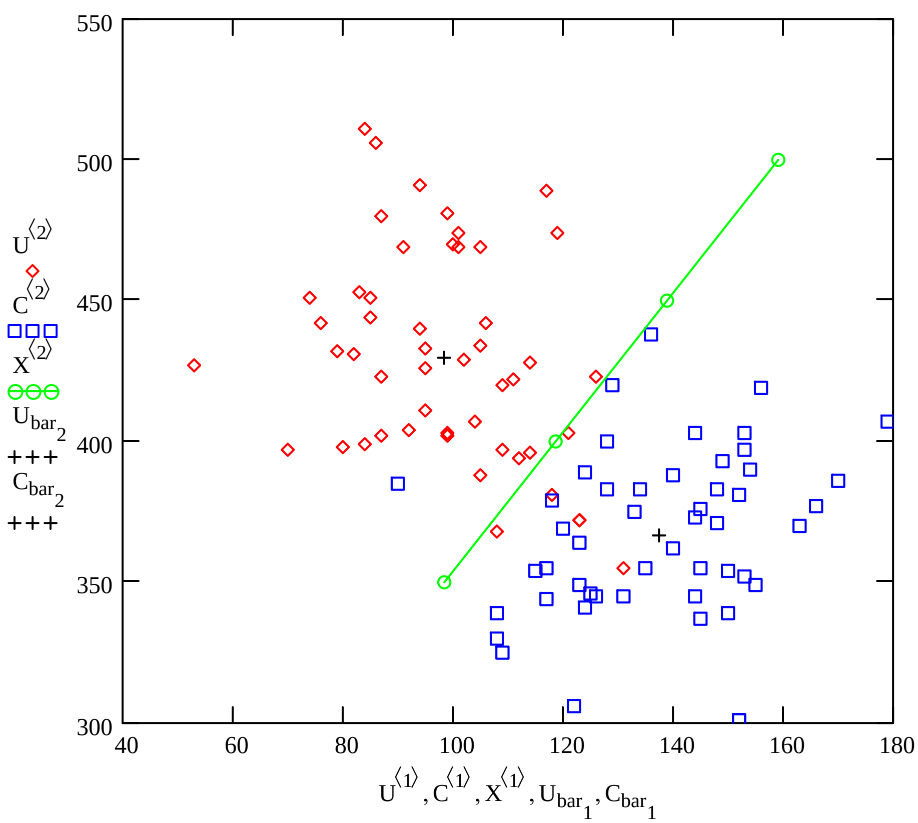
zero =

	1
1	-0.29196
2	-3.92013
3	5.33945
4	9.70071
5	2.62959
6	5.26104
7	5.24536
8	4.83767
9	6.75563
10	2.68153
11	2.0543
12	-2.01
13	-2.01
14	2.28069
15	0.54501
16	2.24737

$$x_1 = \begin{pmatrix} 98.44345 \\ 118.67252 \\ 138.9016 \\ 159.13067 \end{pmatrix}$$

< Since we know m and a_{hat}, and we supply the second coordinate for x_0, we can rearange Eq. 11-18 and solve for the other coordinate. The line provide by these values for x_0 constitute the descriminant dividing line between the populations. (I have only been able to solve for coordinates like this in the linear bivariate case).

$$U_{bar} = \begin{pmatrix} 98.38 \\ 429.66 \end{pmatrix}$$



Estimated Quadratic ECM Rule - Eq. 11-25 for different Σ_1 & Σ_2
Assuming equal costs of misclassification, equal prior probabilities:

$$k := \frac{1}{2}\cdot\ln\left(\frac{|S_U|}{|S_C|}\right) + \frac{1}{2}\left(U_{bar}{}^T\cdot S_U{}^{-1}\cdot U_{bar} - C_{bar}{}^T\cdot S_C{}^{-1}\cdot C_{bar}\right) \qquad k = (31.47367)$$

< jw Eq. 11-24

$$A := \left(S_U{}^{-1} - S_C{}^{-1}\right) \qquad B := \left(U_{bar}{}^T\cdot S_U{}^{-1} - C_{bar}{}^T\cdot S_C{}^{-1}\right)$$

$$X := MM^T$$

$$class_i := \left(\frac{-1}{2}\cdot\left(\left(X^{\langle i\rangle}\right)\right)^T\cdot A\cdot X^{\langle i\rangle} + B\cdot X^{\langle i\rangle} - k\right)_1$$

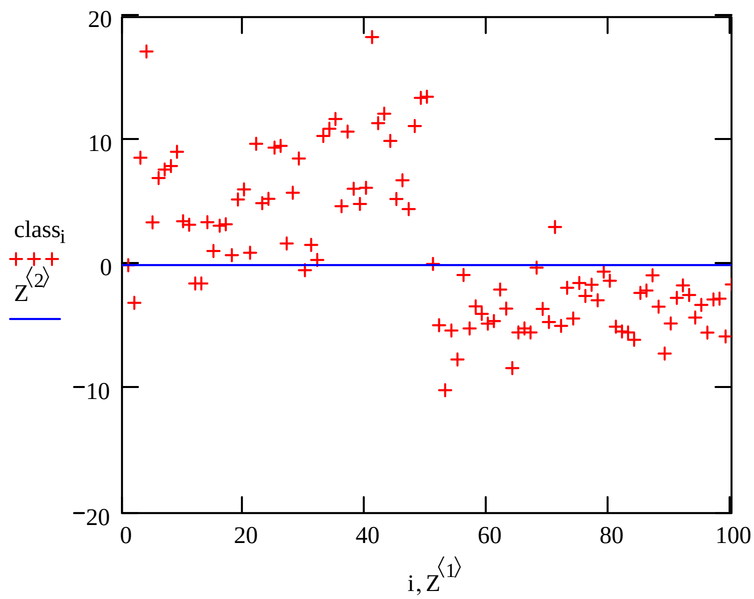
< Eq. 11-25 equal prior probabilities equal cost boundary in class is 0 below:

$$Z := \begin{pmatrix} 0 & 0 \\ 100 & 0 \end{pmatrix}$$

< draws zero line on graph below

class =

	1
1	-0.01682
2	-3.05643
3	8.64557
4	17.22453
5	3.43988
6	7.01394
7	7.70225
8	7.98143
9	9.12927
10	3.52258
11	3.25071
12	-1.4888
13	-1.4888
14	3.45272
15	1.13322
16	3.16997



Statistical distance between population means (D²):

$$D_{sq} := \left(U_{bar} - C_{bar}\right)^T\cdot S_p{}^{-1}\cdot\left(U_{bar} - C_{bar}\right) \qquad D_{sq} = (8.29187)$$

< Multivariate Statistical distance between population means - p. 610

Fisher's linear discriminant function is the same as the linear ECM rule above

$$y_{barU} - y_{barC} = (8.29187)$$

< location of population means along the discriminant axis

2005 jw611

CLASSIFICATION AND DISCRIMANT FUNCTIONS
USING LINEAR FISHER'S DISCRIMINANT FUNCTION FOR TWO GROUPS
jw611.mcd

ORIGIN ≡ 1

Table 8.1 p. 300 in A. C. Rencher

$M := READPRN("\backslash DATA\backslash STEEL.DAT")$

$X_1 := submatrix(M,1,5,2,3)$

$X_2 := submatrix(M,6,12,2,3)$

Total Data Set

M =

	1	2	3
1	1	33	60
2	1	36	61
3	1	35	64
4	1	38	63
5	1	40	65
6	2	35	57
7	2	36	59
8	2	38	59
9	2	39	61
10	2	41	63
11	2	43	65
12	2	41	59

Summary Statistics:

$n_1 := rows(X_1)$ $\quad n_1 = 5$ $\quad i := 1 .. n_1$

$n_2 := rows(X_2)$ $\quad n_2 = 7$ $\quad ii := 1 .. n_2$

$l_{n_i} := 1 \quad ll_{n_{ii}} := 1$

$$X_1 = \begin{pmatrix} 33 & 60 \\ 36 & 61 \\ 35 & 64 \\ 38 & 63 \\ 40 & 65 \end{pmatrix} \qquad X_2 = \begin{pmatrix} 35 & 57 \\ 36 & 59 \\ 38 & 59 \\ 39 & 61 \\ 41 & 63 \\ 43 & 65 \\ 41 & 59 \end{pmatrix}$$

Mean vectors:

$$X1_{bar} := \frac{1}{n_1} \cdot X_1{}^T \cdot l_n \qquad X1_{bar} = \begin{pmatrix} 36.4 \\ 62.6 \end{pmatrix}$$

$$X2_{bar} := \frac{1}{n_2} \cdot X_2{}^T \cdot ll_n \qquad X2_{bar} = \begin{pmatrix} 39 \\ 60.42857 \end{pmatrix}$$

Variance-covariance matrices:

$I := identity(n_1) \qquad II := identity(n_2)$

$$S_1 := \frac{1}{n_1 - 1} \cdot X_1{}^T \cdot \left(I - \frac{1}{n_1} \cdot l_n \cdot l_n{}^T \right) \cdot X_1 \qquad S_1 = \begin{pmatrix} 7.3 & 4.2 \\ 4.2 & 4.3 \end{pmatrix}$$

$$S_2 := \frac{1}{n_2 - 1} \cdot X_2{}^T \cdot \left(II - \frac{1}{n_2} \cdot ll_n \cdot ll_n{}^T \right) \cdot X_2 \qquad S_2 = \begin{pmatrix} 8.33333 & 6.66667 \\ 6.66667 & 7.61905 \end{pmatrix}$$

$$S_p := \frac{n_1 - 1}{n_1 + n_2 - 2} \cdot S_1 + \frac{n_2 - 1}{n_1 + n_2 - 2} \cdot S_2 \qquad S_p = \begin{pmatrix} 7.92 & 5.68 \\ 5.68 & 6.29143 \end{pmatrix}$$

< pooled variance-covariance matrix
Eq. 6-21 p. 284 & Eq. 11-17 p. 592

Estimated LINEAR Minimum Expected cost of Misclassification (ECM) Rule - Eq 11-18
Assuming equal costs of misclassification, equal prior probabilities, and common covariance matrix Σ:

discriminant coefficients a_{hat} - Eq. 11-19:

$$a_{hat} := (X1_{bar} - X2_{bar})^T \cdot S_p{}^{-1} \qquad a_{hat} = (-1.63338 \quad 1.81978)$$

linear transformation of means:

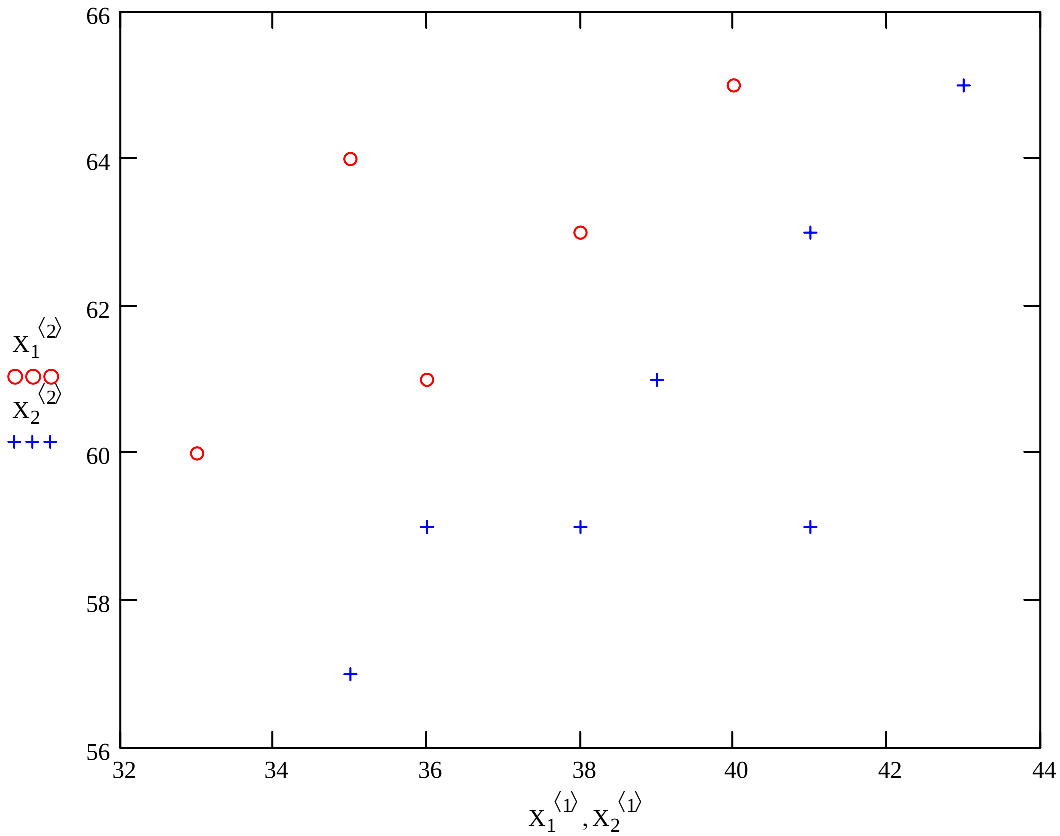
$$y_{bar_1} := (X1_{bar} - X2_{bar})^T \cdot S_p{}^{-1} \cdot X1_{bar} \qquad y_{bar_1} = (54.46324)$$

$$y_{bar_2} := (X1_{bar} - X2_{bar})^T \cdot S_p{}^{-1} \cdot X2_{bar} \qquad y_{bar_2} = (46.26494)$$

midpoint:

$$m := \frac{1}{2} \cdot (X1_{bar} - X2_{bar})^T \cdot S_p{}^{-1} \cdot (X1_{bar} + X2_{bar}) \qquad m = (50.36409) \qquad \frac{1}{2} \cdot (y_{bar_1} + y_{bar_2}) = (50.36409)$$

2005 jw611



66
64
62
60
58
56
32
34
36
38
40
42
44
$X_1^{\langle 2 \rangle}$
$X_2^{\langle 2 \rangle}$
$X_1^{\langle 1 \rangle}, X_2^{\langle 1 \rangle}$

ORIGIN ≡ 1

FISHER'S SAMPLE LINEAR DISCRIMINANTS FOR SEVERAL POPULATIONS
jw 631.mcd

prepared by:
Wm Stein

Using jw Table 11.7 Crude-Oil Data.

Five variables (columns) from three sandstones π

Reading the Data:

M := READPRN("\DATA\T11-7trans.txt")

$\pi_1 := \text{submatrix}(M,1,7,2,6)$

$\pi_2 := \text{submatrix}(M,8,18,2,6)$ p := 5 **< variables**

$\pi_3 := \text{submatrix}(M,19,56,2,6)$ g := 3 **<groups**

$k := 1..g \quad n_k := \text{rows}(\pi_k) \qquad n = \begin{pmatrix} 7 \\ 11 \\ 38 \end{pmatrix}$

M =

	1	2	3	4	5	6
1	1	3.9	7.1414	0.4472	0.1416	12.19
2	1	2.7	7	0.2646	0.1401	12.23
3	1	2.8	6	0.5477	0.1429	11.3
4	1	3.1	6.7082	0.2828	0.1389	13.01
5	1	3.5	6.7823	0.3162	0.128	12.63
6	1	3.9	6.5574	0.2646	0.16	10.42
7	1	2.7	5.9161	0	0.1957	9
8	2	5	6.8557	0.2646	0.1416	6.1
9	2	3.4	5.6569	0.4472	0.1718	4.69
10	2	1.2	3.4641	0	0.1805	3.15
11	2	8.4	4.1231	0.2646	0.1585	4.55
12	2	4.2	6	0.7071	0.1081	4.95
13	2	4.2	5.9161	0.7071	0.1757	2.22
14	2	3.9	6.4031	0.3162	0.1776	2.94
15	2	3.9	6	0.2646	0.1616	2.27
16	2	7.3	5.6569	0.5477	0.1247	12.92
17	2	4.4	6.7823	0.2646	0.1326	5.76
18	2	3	5.4772	0	0.1953	10.77
19	3	6.3	3.6056	0.7071	0.2358	8.27
20	3	1.7	2.3664	1	0.1757	4.64
21	3	7.3	4.899	0	0.2304	2.99
22	3	7.8	4.2426	0.7071	0.2551	6.09
23	3	7.8	5	0.8367	0.1855	6.2
24	3	7.8	5.099	1	0.1992	2.5
25	3	9.5	4.1231	0.2236	0.2841	5.71
26	3	7.7	3.7417	0.5477	0.2151	8.63
27	3	11	4.4721	0.7071	0.2342	8.4
28	3	8	3.7417	0.5477	0.2315	7.87

Summary Statistics:

Sample means:

$i_1 := 1..n_1 \quad ln1_{i_1} := 1 \quad x_{bar_1} := \frac{1}{n_1}\cdot{\pi_1}^T\cdot l_{n1} \quad x_{bar_1} = \begin{pmatrix} 3.22857 \\ 6.58649 \\ 0.3033 \\ 0.1496 \\ 11.54 \end{pmatrix}$

$i_2 := 1..n_2 \quad ln2_{i_2} := 1 \quad x_{bar_2} := \frac{1}{n_2}\cdot{\pi_2}^T\cdot l_{n2} \quad x_{bar_2} = \begin{pmatrix} 4.44545 \\ 5.66685 \\ 0.34397 \\ 0.15709 \\ 5.48364 \end{pmatrix}$

$i_3 := 1..n_3 \quad ln3_{i_3} := 1 \quad x_{bar_3} := \frac{1}{n_3}\cdot{\pi_3}^T\cdot l_{n3} \quad x_{bar_3} = \begin{pmatrix} 7.22632 \\ 4.63367 \\ 0.59812 \\ 0.22317 \\ 5.76789 \end{pmatrix}$

Grand mean:

$\Pi := \text{submatrix}(M,1,56,2,6) \quad N := \text{rows}(\Pi) \qquad N = 56$

$i := 1..N \quad l_{n_i} := 1 \qquad x_{barGM} := \frac{1}{N}\cdot\Pi^T\cdot l_n \qquad x_{barGM} = \begin{pmatrix} 6.18036 \\ 5.08072 \\ 0.51135 \\ 0.201 \\ 6.43357 \end{pmatrix}$

Between matrix - Eq. 11-60 p. 630:

$$B := \sum_{k=1}^{g} n_k\cdot\left(x_{bar_k} - x_{barGM}\right)\cdot\left(x_{bar_k} - x_{barGM}\right)^T \qquad B = \begin{pmatrix} 135.67315 & -60.06734 & 10.94189 & 2.78134 & -113.84136 \\ -60.06734 & 27.2449 & -4.74615 & -1.20156 & 59.0074 \\ 10.94189 & -4.74615 & 0.89727 & 0.22881 & -7.88271 \\ 2.78134 & -1.20156 & 0.22881 & 0.05839 & -1.93937 \\ -113.84136 & 59.0074 & -7.88271 & -1.93937 & 209.2942 \end{pmatrix}$$

2005 jw631

Variance-covariance matrices:

$$I_1 := \text{identity}(n_1)$$

$$S_1 := \frac{1}{n_1 - 1}\cdot \pi_1^T \cdot \left(I_1 - \frac{1}{n_1}\cdot l_{n1}\cdot l_{n1}^T\right)\cdot \pi_1$$

$$S_1 = \begin{pmatrix} 0.28905 & 0.12961 & 0.02547 & -0.00322 & 0.14533 \\ 0.12961 & 0.22105 & 0.02083 & -0.00686 & 0.48575 \\ 0.02547 & 0.02083 & 0.02934 & -0.00281 & 0.1343 \\ -0.00322 & -0.00686 & -0.00281 & 0.0005 & -0.02969 \\ 0.14533 & 0.48575 & 0.1343 & -0.02969 & 2.00187 \end{pmatrix}$$

$$I_2 := \text{identity}(n_2)$$

$$S_2 := \frac{1}{n_2 - 1}\cdot \pi_2^T \cdot \left(I_2 - \frac{1}{n_2}\cdot l_{n2}\cdot l_{n2}^T\right)\cdot \pi_2$$

$$S_2 = \begin{pmatrix} 3.85273 & 0.11825 & 0.16131 & -0.02471 & 2.26992 \\ 0.11825 & 1.07552 & 0.09241 & -0.01016 & 0.26195 \\ 0.16131 & 0.09241 & 0.05785 & -0.00352 & -0.06325 \\ -0.02471 & -0.01016 & -0.00352 & 0.00073 & -0.02241 \\ 2.26992 & 0.26195 & -0.06325 & -0.02241 & 11.83825 \end{pmatrix}$$

$$I_3 := \text{identity}(n_3)$$

$$S_3 := \frac{1}{n_3 - 1}\cdot \pi_3^T \cdot \left(I_3 - \frac{1}{n_3}\cdot l_{n3}\cdot l_{n3}^T\right)\cdot \pi_3$$

$$S_3 = \begin{pmatrix} 3.98145 & -0.00011 & -0.15669 & 0.03673 & 1.51625 \\ -0.00011 & 0.80292 & 0.02917 & 1.67077 \times 10^{-6} & -0.91287 \\ -0.15669 & 0.02917 & 0.07636 & -0.00626 & 0.06447 \\ 0.03673 & 1.67077 \times 10^{-6} & -0.00626 & 0.00181 & -0.01606 \\ 1.51625 & -0.91287 & 0.06447 & -0.01606 & 5.6116 \end{pmatrix}$$

Pooled covariance matrix:

$$S_{pooled} := \frac{1}{\sum_{i=1}^{g} (n_i) - g}\cdot \sum_{i=1}^{g} (n_i - 1)\cdot S_i$$

^jw Eq 11-50 p. 618

$$S_{pooled} = \begin{pmatrix} 3.53916 & 0.03691 & -0.07607 & 0.02061 & 1.50325 \\ 0.03691 & 0.78848 & 0.04016 & -0.00269 & -0.53287 \\ -0.07607 & 0.04016 & 0.06755 & -0.00535 & 0.04828 \\ 0.02061 & -0.00269 & -0.00535 & 0.00146 & -0.0188 \\ 1.50325 & -0.53287 & 0.04828 & -0.0188 & 6.37779 \end{pmatrix}$$

Within matrix - Eq. 11-61 p. 630:

$$W := \sum_{k=1}^{g} (n_k - 1)\cdot S_k$$

$$W = \begin{pmatrix} 187.57524 & 1.95623 & -4.03151 & 1.09245 & 79.67229 \\ 1.95623 & 41.78941 & 2.12844 & -0.14274 & -28.24236 \\ -4.03151 & 2.12844 & 3.57995 & -0.28354 & 2.55869 \\ 1.09245 & -0.14274 & -0.28354 & 0.07712 & -0.99629 \\ 79.67229 & -28.24236 & 2.55869 & -0.99629 & 338.02289 \end{pmatrix}$$

Discriminant matrix ($W^{-1}B$):

$$W^{-1}\cdot B = \begin{pmatrix} 0.9993 & -0.4673 & 0.07684 & 0.01935 & -1.16736 \\ -2.298 & 1.06251 & -0.17853 & -0.04504 & 2.52449 \\ 9.4922 & -4.13217 & 0.77615 & 0.19782 & -7.03453 \\ 43.40659 & -17.6293 & 3.74035 & 0.96264 & -15.42337 \\ -0.70824 & 0.3528 & -0.0512 & -0.01272 & 1.11303 \end{pmatrix}$$

$$\lambda := \text{reverse}\left(\text{sort}\left(\text{eigenvals}\left(W^{-1}\cdot B\right)\right)\right)$$

$$i := 1 .. \text{rank}\left(W^{-1}\cdot B\right)$$

$$\lambda = \begin{pmatrix} 4.35423 \\ 0.5594 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Descriminant function coefficients:

$$\varepsilon^{\langle i \rangle} := \text{eigenvec}\left(W^{-1}\cdot B, \lambda_i\right)$$

$$\varepsilon = \begin{pmatrix} -0.02569 & -0.00691 \\ 0.05843 & 0.01 \\ -0.22745 & 0.08335 \\ -0.9715 & 0.99633 \\ 0.01938 & 0.01539 \end{pmatrix}$$

Canonical Correlations:

$$r_i := \sqrt{\frac{\lambda_i}{1 + \lambda_i}} \qquad r = \begin{pmatrix} 0.90179 \\ 0.59894 \end{pmatrix}$$

Scaling eigenvectors into linear discriminant coefficients:

$$\varepsilon^T \cdot S_{pooled} \cdot \varepsilon = \begin{pmatrix} 0.00677 & 0 \\ 0 & 0.00166 \end{pmatrix}$$

<Note that $\varepsilon^T S_{pooled} \varepsilon$ are not unity

$$L^{\langle i \rangle} := \left(\varepsilon^{\langle i \rangle T} \cdot S_{pooled} \cdot \varepsilon^{\langle i \rangle} \right) \qquad L = (0.00677 \quad 0.00166)$$

$$a^{\langle i \rangle} := \frac{1}{\sqrt{\left(L^{\langle i \rangle}\right)_1}} \cdot \varepsilon^{\langle i \rangle}$$

< discriminant coefficients

$$a = \begin{pmatrix} -0.3122 & -0.16954 \\ 0.70997 & 0.24557 \\ -2.76375 & 2.04605 \\ -11.80485 & 24.45877 \\ 0.23549 & 0.37782 \end{pmatrix}$$

<Same as scaled eigenvectors (e) in jw313 citing Rencher

$$a^T \cdot S_{pooled} \cdot a = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

< coefficients a are now scaled so that $a^T S_{pooled} \alpha$ are unity

$$j := 1 .. p$$

$$\phi_j := S_{pooled_{j,j}}$$

$$\text{diag}\left(\sqrt{\phi}\right) \cdot a = \begin{pmatrix} -0.58733 & -0.31895 \\ 0.63043 & 0.21806 \\ -0.71829 & 0.53176 \\ -0.4503 & 0.93299 \\ 0.5947 & 0.95417 \end{pmatrix}$$

< Standardized discriminant coefficients as described by Rencher and in jw313

These may be used to determine the "importance" of each original variable with regard to each discriminant function axis within the context of the other variables.

Discriminant scores:

object scores:

$$yhat\pi_1 := \pi_1 \cdot a$$

$$yhat\pi_2 := \pi_2 \cdot a$$

$$yhat\pi_3 := \pi_3 \cdot a$$

< Each ROW represents a point transformed from the original data matrix π_1, π_2, or π_3.

First COLUMNS are first discriminant scores
Second COLUMNS are second discriminant scores

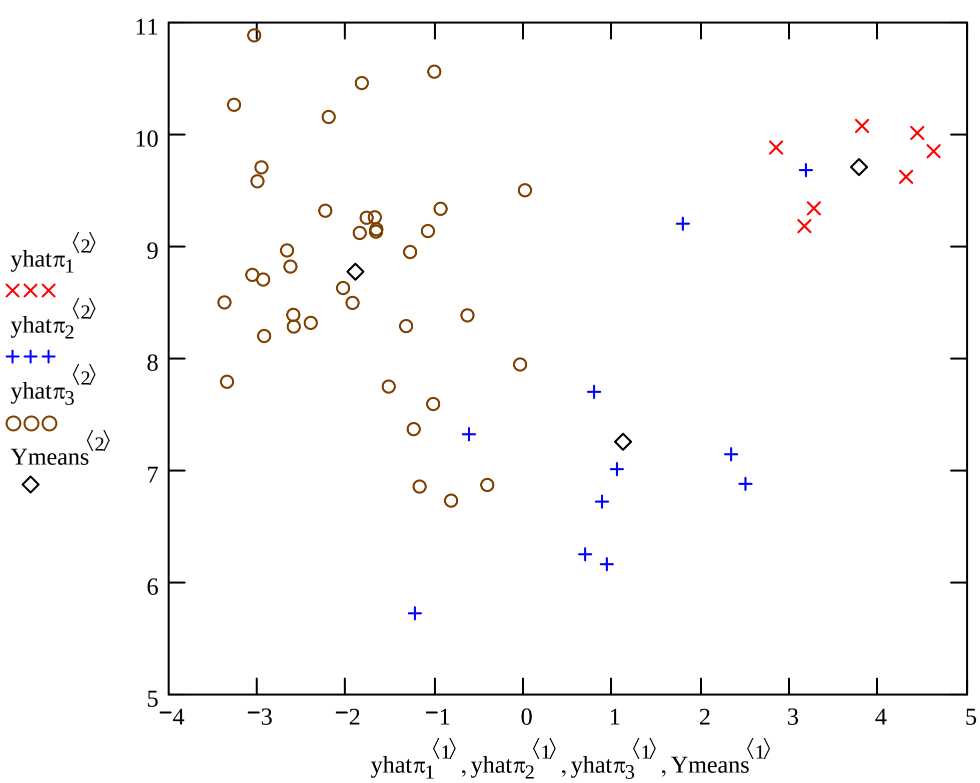
Discriminant scores may be plotted as a summary of the data.

mean scores:

$$Yhat_{bar_k} := x_{bar_k}{}^T \cdot a$$

$$Ymeans := \text{stack}\left(Yhat_{bar_1}, Yhat_{bar_2}, Yhat_{bar_3}\right)$$

$$Ymeans = \begin{pmatrix} 3.78154 & 9.70976 \\ 1.12168 & 7.25581 \\ -1.8956 & 8.77433 \end{pmatrix}$$



2005 jw631

Mean Centered Discriminant scores (ζ):

$j := 1 .. p$

object scores:

$i := 1 .. n_1$

$\zeta1_{i,j} := x_{barGM_j} - (\pi1)_{i,j}$

$yhat\zeta_1 := \zeta_1 \cdot a$

$$yhat\zeta_1 = \begin{pmatrix} -4.40897 & -1.48356 \\ -5.215 & -1.2571 \\ -3.43934 & -1.29092 \\ -5.0305 & -1.42021 \\ -4.90511 & -1.02876 \\ -3.86498 & -0.74783 \\ -3.75978 & -0.58907 \end{pmatrix}$$

$i := 1 .. n_2$

$\zeta2_{i,j} := x_{barGM_j} - (\pi2)_{i,j}$

$yhat\zeta_2 := \zeta_2 \cdot a$

$$yhat\zeta_2 = \begin{pmatrix} -2.93326 & 1.44765 \\ -1.38846 & 0.89125 \\ -1.28906 & 2.3408 \\ 0.6328 & 2.8674 \\ -1.47719 & 1.87064 \\ 0.02326 & 1.2693 \\ -1.64362 & 1.58014 \\ -1.53115 & 2.42919 \\ -2.38718 & -0.61066 \\ -3.09465 & 1.71254 \\ -3.77605 & -1.0894 \end{pmatrix}$$

$i := 1 .. n_3$

$\zeta3_{i,j} := x_{barGM_j} - (\pi3)_{i,j}$

$yhat\zeta_3 := \zeta_3 \cdot a$

$$yhat\zeta_3 = \begin{pmatrix} 1.60406 & -1.56309 \\ 2.00259 & 0.20352 \\ 0.22335 & 1.86258 \\ 2.3613 & -1.11361 \\ 1.33423 & 0.096 \\ 2.74828 & 0.80043 \\ 2.07243 & -0.37251 \\ 1.17483 & -0.66274 \\ 2.4067 & -0.98903 \\ 1.64106 & -0.72586 \\ 1.06888 & -0.53871 \\ 2.44213 & -2.29148 \\ 1.08339 & -0.66832 \\ 2.34068 & -0.112 \\ 0.48383 & -0.54474 \\ 1.06535 & -0.56001 \\ 2.03431 & -0.22718 \\ 2.77848 & 0.09236 \\ 0.72911 & 0.30312 \\ 0.03841 & 0.20751 \\ 2.46307 & -0.15296 \\ 1.80581 & 0.27519 \\ 1.23028 & -1.86589 \\ -0.5552 & 0.64606 \\ 0.68437 & -0.35726 \\ 1.25272 & -0.52765 \\ 1.99686 & 0.30727 \\ 2.66977 & -1.67265 \\ 0.34263 & -0.74328 \\ 0.92677 & 0.84299 \\ 2.3299 & 0.39155 \\ 0.64521 & 1.22403 \\ -0.18543 & 1.7222 \\ 0.57786 & 1.73721 \\ -0.60931 & -0.90956 \\ 0.4251 & 0.99896 \\ 0.41329 & -1.96798 \\ 1.44115 & -0.03583 \end{pmatrix}$$

mean scores:

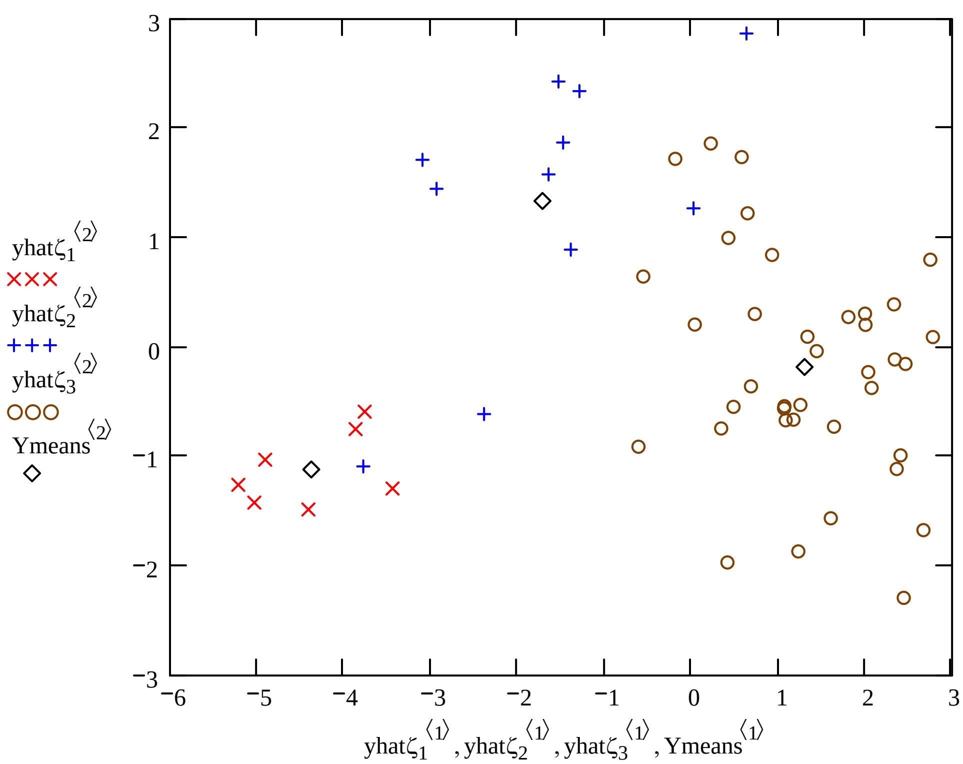
$Yhat_{bar_k} := \left(\left(x_{barGM} - x_{bar_k}\right)^T\right) \cdot a$

$Ymeans := stack\left(Yhat_{bar_1}, Yhat_{bar_2}, Yhat_{bar_3}\right)$

< Each ROW represents a point transformed from the original data matrix π_1, π_2, or π_3.

First COLUMNS are first discriminant scores
Second COLUMNS are second discriminant scores

Discriminant scores may be plotted as a summary of the data.



$$Ymeans = \begin{pmatrix} -4.37481 & -1.11678 \\ -1.71496 & 1.33717 \\ 1.30232 & -0.18135 \end{pmatrix}$$

MANOVA Table jw p. 299

Source	SSP Matrix	Degrees of Freedom
Treatment Matrix B	$B = \begin{pmatrix} 135.67315 & -60.06734 & 10.94189 & 2.78134 & -113.84136 \\ -60.06734 & 27.2449 & -4.74615 & -1.20156 & 59.0074 \\ 10.94189 & -4.74615 & 0.89727 & 0.2288 & \text{df}_B := g - 1 \\ 2.78134 & -1.20156 & 0.22881 & 0.05839 & -1.93937 \\ -113.84136 & 59.0074 & -7.88271 & -1.93937 & 209.2942 \end{pmatrix}$	$df_B = 2$
Residual/Error Matrix W	$W = \begin{pmatrix} 187.57524 & 1.95623 & -4.03151 & 1.09245 & 79.67229 \\ 1.95623 & 41.78941 & 2.12844 & -0.14274 & -28.24236 \\ -4.03151 & 2.12844 & 3.57995 & -0.28354 & 2.55869 \\ 1.09245 & -0.14274 & -0.28354 & 0.07712 & -0.99629 \\ 79.67229 & -28.24236 & 2.55869 & -0.99629 & 338.02289 \end{pmatrix}$	$df_W = 53$
TOTAL (mean corrected)	$B + W = \begin{pmatrix} 323.24839 & -58.11111 & 6.91038 & 3.87379 & -34.16907 \\ -58.11111 & 69.03431 & -2.61771 & -1.3443 & 30.76504 \\ 6.91038 & -2.61771 & 4.47722 & -0.0547 & \text{df}_T := N - 1 \\ 3.87379 & -1.3443 & -0.05473 & 0.1355 & -2.93566 \\ -34.16907 & 30.76504 & -5.32402 & -2.93566 & 547.31709 \end{pmatrix}$	$df_T = 55$

MANOVA tests jw p. 299-300.

Decomposition Model:

$X_{i,j} = \mu + \tau_i + \varepsilon_{i,j}$ **where:** $j = 1$ to n_m, $m = 1$ to g

Restriction:

$\Sigma n_m \tau_m = 0$

Assumptions:

All populations P_1-P_g rs

$\varepsilon_{i,j}$ **rs** $N(0, \sigma^2)$

Population variances equal

Hypothesis testing:

$H_0 : \tau_1 = \tau_2 = \ldots \tau_g = 0$

H_1 : at least one $\tau <> 0$

Wilk's Lambda Test Statistic:

$$\Lambda_s := \frac{|W|}{|W + B|} \qquad \Lambda_s = 0.11977$$

Lawley-Hotelling Trace:

$$LHtr := tr\left(B \cdot W^{-1}\right) \qquad LHtr = 4.91364$$

Pillai trace:

$$Ptr := tr\left[B \cdot (B + W)^{-1}\right] \qquad Ptr = 1.17196$$

Stringency of the test: $\alpha := 0.01$ **< set as desired**

If assumptions hold and H_0 is true then jw Table 6.3 & p. 300:

p > 0 **g = 3**

$$C := qF[1 - \alpha, 2 \cdot p, 2 \cdot (N - p - 2)] \qquad C = 2.50716$$

Decision Rule: Reject H_0 if $K > C$:

$$K := \left(\frac{N - p - 2}{p}\right) \cdot \left(\frac{1 - \sqrt{\Lambda_s}}{\sqrt{\Lambda_s}}\right) \qquad K = 18.51743$$

Probability: $1 - pF[K, 2 \cdot p, 2 \cdot (N - p - 2)] = 0$

$$Decision := if(K > C, 1, 0) \qquad Decision = 1$$

< 0 = Do not reject H_0
1 = Reject H_0

Simultaneous Bonferroni confidence intervals for treatment effects jw p. 305:

$\alpha := 0.05$ **< Set probability of Type 1 error**

$c := qt\left[1 - \frac{\alpha}{p \cdot g \cdot (g-1)}, N - g\right]$ $\quad c = 3.07408$

Difference in means between first & second populations:

$i := 1 .. p$

$\kappa_i := \left| \sqrt{\frac{W_{i,i}}{N-g} \cdot \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} \right|$ $\quad \kappa = \begin{pmatrix} 0.90958 \\ 0.42932 \\ 0.12566 \\ 0.01844 \\ 1.22103 \end{pmatrix}$

$ci_{lower} := x_{bar_1} - x_{bar_2} - c \cdot \kappa$ $\quad ci_{upper} := x_{bar_1} - x_{bar_2} + c \cdot \kappa$

$ci := augment(ci_{lower}, ci_{upper})$ $\quad ci = \begin{pmatrix} -4.01301 & 1.57924 \\ -0.40015 & 2.23941 \\ -0.42696 & 0.34561 \\ -0.06419 & 0.0492 \\ 2.30282 & 9.80991 \end{pmatrix}$

Difference in means between first & third populations:

Note: According to jw Result 6.5, Bonferroni simultaneous confidence intervals apply to ALL variables (=components of mean vectors) simultaneously for ALL pairwise comparisons across groups.

$i := 1 .. p$

$\kappa_i := \left| \sqrt{\frac{W_{i,i}}{N-g} \cdot \left(\frac{1}{n_1} + \frac{1}{n_3} \right)} \right|$ $\quad \kappa = \begin{pmatrix} 0.77378 \\ 0.36523 \\ 0.1069 \\ 0.01569 \\ 1.03872 \end{pmatrix}$

$ci_{lower} := x_{bar_1} - x_{bar_3} - c \cdot \kappa$ $\quad ci_{upper} := x_{bar_1} - x_{bar_3} + c \cdot \kappa$

$ci := augment(ci_{lower}, ci_{upper})$ $\quad ci = \begin{pmatrix} -6.3764 & -1.61909 \\ 0.83009 & 3.07555 \\ -0.62343 & 0.03379 \\ -0.1218 & -0.02534 \\ 2.57898 & 8.96523 \end{pmatrix}$

Difference in means between second & third populations:

$i := 1 .. p$

$\kappa_i := \left| \sqrt{\frac{W_{i,i}}{N-g} \cdot \left(\frac{1}{n_2} + \frac{1}{n_3} \right)} \right|$ $\quad \kappa = \begin{pmatrix} 0.64411 \\ 0.30402 \\ 0.08898 \\ 0.01306 \\ 0.86466 \end{pmatrix}$

$ci_{lower} := x_{bar_2} - x_{bar_3} - c \cdot \kappa$ $\quad ci_{upper} := x_{bar_2} - x_{bar_3} + c \cdot \kappa$

$ci := augment(ci_{lower}, ci_{upper})$ $\quad ci = \begin{pmatrix} -4.76091 & -0.80082 \\ 0.0986 & 1.96777 \\ -0.52769 & 0.01939 \\ -0.10623 & -0.02593 \\ -2.94229 & 2.37378 \end{pmatrix}$

2005 jw631

Applying Fisher's Classification Procedure: **See jw Eq. 11-67 p. 638**

$m := 1 .. g - 1 \quad l_{z_m} := 1 \quad o := 1 .. N \quad N = 56$

$yhat := stack(yhat\zeta_1, yhat\zeta_2, yhat\zeta_3)^T$

<Discriminant scores are stacked into a single list of objects (N=56), and are transposed here for computational convenience

$Obj := augment(M^{\langle 1 \rangle}, yhat^T)$

<Matrix Obj is created to show ORIGINAL membership of each object (ROWS) in a group (first COLUMN) along with discriminant scores:
COLUMN 2 = first discriminant function
COLUMN 3 = second discriminant function

$$G1_o := \left[l_z^T \cdot \left[(yhat)^{\langle o \rangle} - \left(Ymeans^T \right)^{\langle 1 \rangle} \right]^2 \right]_1$$

$$G2_o := \left[l_z^T \cdot \left[(yhat)^{\langle o \rangle} - \left(Ymeans^T \right)^{\langle 2 \rangle} \right]^2 \right]_1$$

$$G3_o := \left[l_z^T \cdot \left[(yhat)^{\langle o \rangle} - \left(Ymeans^T \right)^{\langle 3 \rangle} \right]^2 \right]_1$$

$G := augment(G1, G2, G3)$

<Squarred discriminant distances between each object and the means for each group are calculated and placed in COLUMNS of matrix G

Fisher's criterion says to place each object (ROW) in the group (COLUMNS) showing the smallest distance (numeric value in G)

e.g., Object 1 (originally in Group 1 = π_1) shows smallest distance with Group 1 in matrix G - so if it was a "new" object (not part of the training set) it would be placed (correctly) in Group 1.
Note, however, that Object 8 is not so lucky! Originally in Group 2, as a "new" object, it would be mis-classified as belonging to Group 1

Obj =

	1	2	3
1	1	-4.40897	-1.48356
2	1	-5.215	-1.2571
3	1	-3.43934	-1.29092
4	1	-5.0305	-1.42021
5	1	-4.90511	-1.02876
6	1	-3.86498	-0.74783
7	1	-3.75978	-0.58907
8	2	-2.93326	1.44765
9	2	-1.38846	0.89125
10	2	-1.28906	2.3408
11	2	0.6328	2.8674
12	2	-1.47719	1.87064
13	2	0.02326	1.2693
14	2	-1.64362	1.58014
15	2	-1.53115	2.42919
16	2	-2.38718	-0.61066
17	2	-3.09465	1.71254
18	2	-3.77605	-1.0894
19	3	1.60406	-1.56309
20	3	2.00259	0.20352
21	3	0.22335	1.86258
22	3	2.3613	-1.11361
23	3	1.33423	0.096
24	3	2.74828	0.80043
25	3	2.07243	-0.37251
26	3	1.17483	-0.66274
27	3	2.4067	-0.98903
28	3	1.64106	-0.72586

G =

	1	2	3
1	0.13569	15.21417	34.31457
2	0.72561	18.98053	43.63277
3	0.90544	9.88032	23.71446
4	0.522	18.59596	41.6394
5	0.28896	15.77466	39.25027
6	0.39605	8.96982	27.02195
7	0.65673	7.89171	25.79114
8	8.65438	1.49646	20.59381
9	12.95047	0.30545	8.39077
10	21.47666	1.18866	13.07652
11	40.94988	7.85359	9.74317
12	17.32092	0.34113	11.93636
13	25.03642	3.02603	3.74037
14	14.73274	0.06412	11.78145
15	20.6603	1.22629	14.84344
16	4.20682	4.24593	13.79675
17	9.64389	2.04444	22.92017
18	0.35927	10.13631	26.61438
19	35.94611	19.42741	2.00026
20	42.41444	15.10533	0.6385
21	30.01965	4.0331	5.34183
22	45.3752	22.6222	1.99054
23	34.06394	10.83803	0.07794
24	54.41419	20.20863	3.0547
25	42.12087	17.26733	0.62961
26	31.00473	12.35054	0.24799
27	46.00528	22.3993	1.872
28	36.34357	15.51896	0.41123

ORIGIN ≡ 1

FISHER'S SAMPLE LINEAR DISCRIMINANTS FOR SEVERAL POPULATIONS
jw 631A.mcd

prepared by: Wm Stein

Verifying Calculations in Example 11.13 - p. 631 with three populations:

$$X_1 := \begin{pmatrix} -2 & 5 \\ 0 & 3 \\ -1 & 1 \end{pmatrix} \qquad X_2 := \begin{pmatrix} 0 & 6 \\ 2 & 4 \\ 1 & 2 \end{pmatrix} \qquad X_3 := \begin{pmatrix} 1 & -2 \\ 0 & 0 \\ -1 & -4 \end{pmatrix}$$

Sample means:

$n_1 := \text{rows}(X_1)$ $\quad n_2 := \text{rows}(X_2)$ $\quad n_3 := \text{rows}(X_3)$

$i_1 := 1..n_1$ $\quad i_2 := 1..n_2$ $\quad i_3 := 1..n_3$

$l_{n1_{i_1}} := 1$ $\quad l_{n2_{i_2}} := 1$ $\quad l_{n3_{i_3}} := 1$

$$x_{bar_1} := \frac{1}{n_1}\cdot X_1^T\cdot l_{n1} \qquad x_{bar_2} := \frac{1}{n_2}\cdot X_2^T\cdot l_{n2} \qquad x_{bar_3} := \frac{1}{n_3}\cdot X_3^T\cdot l_{n3}$$

$$x_{bar_1} = \begin{pmatrix} -1 \\ 3 \end{pmatrix} \qquad x_{bar_2} = \begin{pmatrix} 1 \\ 4 \end{pmatrix} \qquad x_{bar_3} = \begin{pmatrix} 0 \\ -2 \end{pmatrix}$$

Grand mean:

$g := 3$

$$x_g := \frac{1}{g}\cdot\sum_{k=1}^{g} x_{bar_k} \qquad x_g = \begin{pmatrix} 0 \\ 1.66667 \end{pmatrix}$$

Note:

- **All values are verified jw pp. 631-632.**
- **To use formulae for B & W matricies below, n_i, x.bar$_i$ & S_i must be set up as VECTORS of numbers, vectors, & matrices respectively using MathCad's "[" function.**

Between matrix - Eq. 11-60 p. 630:

$$B := \sum_{k=1}^{g} n_k\cdot(x_{bar_k} - x_g)\cdot(x_{bar_k} - x_g)^T \qquad B = \begin{pmatrix} 6 & 3 \\ 3 & 62 \end{pmatrix}$$

Variance-covariance matrices:

$I_1 := \text{identity}(n_1)$

$$S_1 := \frac{1}{n_1 - 1}\cdot X_1^T\cdot\left(I_1 - \frac{1}{n_1}\cdot l_{n1}\cdot l_{n1}^T\right)\cdot X_1 \qquad S_1 = \begin{pmatrix} 1 & -1 \\ -1 & 4 \end{pmatrix}$$

$I_2 := \text{identity}(n_2)$

$$S_2 := \frac{1}{n_2 - 1}\cdot X_2^T\cdot\left(I_2 - \frac{1}{n_2}\cdot l_{n2}\cdot l_{n2}^T\right)\cdot X_2 \qquad S_2 = \begin{pmatrix} 1 & -1 \\ -1 & 4 \end{pmatrix}$$

$I_3 := \text{identity}(n_3)$

$$S_3 := \frac{1}{n_3 - 1}\cdot X_3^T\cdot\left(I_3 - \frac{1}{n_3}\cdot l_{n3}\cdot l_{n3}^T\right)\cdot X_3 \qquad S_3 = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix}$$

Within matrix - Eq. 11-61 p. 630:

$$W := \sum_{k=1}^{g} (n_k - 1)\cdot S_k \qquad W = \begin{pmatrix} 6 & -2 \\ -2 & 24 \end{pmatrix}$$

Discriminant matrix ($W^{-1}B$):

$$W^{-1}\cdot B = \begin{pmatrix} 1.07143 & 1.4 \\ 0.21429 & 2.7 \end{pmatrix}$$

Descriminant function coefficients:

$$\lambda := \text{reverse}\left(\text{sort}\left(\text{eigenvals}\left(W^{-1}\cdot B\right)\right)\right) \qquad \lambda = \begin{pmatrix} 2.86707 \\ 0.90436 \end{pmatrix}$$

$$i := 1 .. \text{rank}\left(W^{-1}\cdot B\right)$$

$$\varepsilon^{\langle i\rangle} := \text{eigenvec}\left(W^{-1}\cdot B, \lambda_i\right) \qquad \varepsilon = \begin{pmatrix} 0.61487 & -0.99295 \\ 0.78863 & 0.1185 \end{pmatrix}$$

Scaling eigenvectors into linear discriminant coefficients:

$$S_{pooled} := \frac{W}{\sum_{k=1}^{g} n_k - g}$$

< Eq. 11-62 p. 630

$$\varepsilon^T\cdot S_{pooled}\cdot\varepsilon = \begin{pmatrix} 2.54254 & 0 \\ 0 & 1.12056 \end{pmatrix}$$

^Note that $\varepsilon^T S_{pooled}\varepsilon$ are not unity

$$L^{\langle i\rangle} := \left(\varepsilon^{\langle i\rangle T}\cdot S_{pooled}\cdot\varepsilon^{\langle i\rangle}\right) \qquad L = \begin{pmatrix} 2.54254 & 1.12056 \end{pmatrix}$$

< Squared length of $\varepsilon^T S_{pooled}\varepsilon$ prior to scaling

$$a^{\langle i\rangle} := \frac{1}{\sqrt{\left(L^{\langle i\rangle}\right)_1}}\cdot\varepsilon^{\langle i\rangle}$$

< discriminant coefficients

$$a = \begin{pmatrix} 0.38561 & -0.93802 \\ 0.49458 & 0.11194 \end{pmatrix}$$

< values verified p. 687

$$a^T\cdot S_{pooled}\cdot a = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

< coefficients a are now scaled so that $a^T S_{pooled}\alpha$ are unity

Discriminant scores:

$$a^T\cdot X_1^T = \begin{pmatrix} 1.7017 & 1.48375 & 0.10897 \\ 2.43573 & 0.33582 & 1.04996 \end{pmatrix}$$

$$a^T\cdot X_2^T = \begin{pmatrix} 2.9675 & 2.74955 & 1.37478 \\ 0.67164 & -1.42828 & -0.71414 \end{pmatrix}$$

$$a^T\cdot X_3^T = \begin{pmatrix} -0.60356 & 0 & -2.36394 \\ -1.1619 & 0 & 0.49026 \end{pmatrix}$$

< Each column represents a point transformed from the original data matrix X_1, X_2, or X_3.

First rows are first discriminant scores
Second rows are second discriminant scores

Discriminant scores may be plotted as a summary of the data.

2005 jw631B

ORIGIN ≡ 1

FISHER'S SAMPLE LINEAR DISCRIMINANTS FOR SEVERAL POPULATIONS
jw 631B.mcd

prepared by:
Wm Stein

Using jw Table 11.7 Salmon Data.

Two variables (columns 3 & 4) for US and Canadian populations (column 1)

Reading the Data:

$M := \text{READPRN}("\backslash DATA\backslash T11\text{-}2.DAT")$

$\pi_1 := \text{submatrix}(M, 1, 50, 3, 4)$

$\pi_2 := \text{submatrix}(M, 51, 100, 3, 4)$ $p := 2$ **< variables**

$g := 2$ **<groups**

$k := 1..g \quad n_k := \text{rows}(\pi_k) \qquad n = \begin{pmatrix} 50 \\ 50 \end{pmatrix}$

M =

	1	2	3	4	5	6
1	1	2	108	368		
2	1	1	131	355		
3	1	1	105	469		
4	1	2	86	506		
5	1	1	99	402		
6	1	2	87	423		
7	1	1	94	440		
8	1	2	117	489		
9	1	2	79	432		
10	1	1	99	403		
11	1	1	114	428		
12	1	2	123	372		
13	1	1	123	372		
14	1	2	109	420		
15	1	2	112	394		
16	1	1	104	407		
17	1	2	111	422		
18	1	2	126	423		
19	1	2	105	434		
20	1	1	119	474		
21	1	1	114	396		
22	1	2	100	470		
23	1	2	84	399		
24	1	2	102	429		
25	1	2	101	469		
26	1	2	85	444		
27	1	1	109	397		
28	1	2	106	442		

Summary Statistics:

Sample means:

$i_1 := 1..n_1 \quad l_{n1_{i_1}} := 1 \quad x_{bar_1} := \frac{1}{n_1} \cdot \pi_1^T \cdot l_{n1} \quad x_{bar_1} = \begin{pmatrix} 98.38 \\ 429.66 \end{pmatrix}$

$i_2 := 1..n_2 \quad l_{n2_{i_2}} := 1 \quad x_{bar_2} := \frac{1}{n_2} \cdot \pi_2^T \cdot l_{n2} \quad x_{bar_2} = \begin{pmatrix} 137.46 \\ 366.62 \end{pmatrix}$

Grand mean:

$\Pi := \text{submatrix}(M, 1, 100, 3, 4) \quad N := \text{rows}(\Pi) \qquad N = 100$

$i := 1..N \quad l_{n_i} := 1 \qquad x_{barGM} := \frac{1}{N} \cdot \Pi^T \cdot l_n \qquad x_{barGM} = \begin{pmatrix} 117.92 \\ 398.14 \end{pmatrix}$

Between matrix - Eq. 11-60 p. 630:

$$B := \sum_{k=1}^{g} n_k \cdot (x_{bar_k} - x_{barGM}) \cdot (x_{bar_k} - x_{barGM})^T \qquad B = \begin{pmatrix} 38181.16 & -61590.08 \\ -61590.08 & 99351.04 \end{pmatrix}$$

2005 jw631B

Variance-covariance matrices:

$$I_1 := \text{identity}(n_1)$$

$$S_1 := \frac{1}{n_1 - 1}\cdot\pi_1^T\cdot\left(I_1 - \frac{1}{n_1}\cdot l_{n1}\cdot l_{n1}^T\right)\cdot\pi_1 \qquad S_1 = \begin{pmatrix} 260.60776 & -188.09265 \\ -188.09265 & 1399.08612 \end{pmatrix}$$

$$I_2 := \text{identity}(n_2)$$

$$S_2 := \frac{1}{n_2 - 1}\cdot\pi_2^T\cdot\left(I_2 - \frac{1}{n_2}\cdot l_{n2}\cdot l_{n2}^T\right)\cdot\pi_2 \qquad S_2 = \begin{pmatrix} 326.0902 & 133.5049 \\ 133.5049 & 893.26082 \end{pmatrix}$$

Pooled covariance matrix:

$$S_{pooled} := \frac{1}{\sum_{i=1}^{g}(n_i) - g}\cdot\sum_{i=1}^{g}(n_i - 1)\cdot S_i \qquad S_{pooled} = \begin{pmatrix} 293.34898 & -27.29388 \\ -27.29388 & 1146.17347 \end{pmatrix}$$

^jw Eq 11-50 p. 618

Within matrix - Eq. 11-61 p. 630:

$$W := \sum_{k=1}^{g}(n_k - 1)\cdot S_k \qquad W = \begin{pmatrix} 28748.2 & -2674.8 \\ -2674.8 & 1.12325\times 10^5 \end{pmatrix}$$

Discriminant matrix ($W^{-1}B$):

$$W^{-1}\cdot B = \begin{pmatrix} 1.27994 & -2.06468 \\ -0.51784 & 0.83533 \end{pmatrix}$$

$$\lambda := \text{reverse}\left(\text{sort}\left(\text{eigenvals}\left(W^{-1}\cdot B\right)\right)\right)$$

$$i := 1 .. \text{rank}\left(W^{-1}\cdot B\right) \qquad \lambda = \begin{pmatrix} 2.11527 \\ 0 \end{pmatrix}$$

Descriminant function coefficients:

$$\varepsilon^{\langle i\rangle} := \text{eigenvec}\left(W^{-1}\cdot B, \lambda_i\right) \qquad \varepsilon = \begin{pmatrix} -0.927 \\ 0.37505 \end{pmatrix}$$

Canonical Correlations:

$$r_i := \sqrt{\frac{\lambda_i}{1 + \lambda_i}} \qquad r = (0.82402)$$

Scaling eigenvectors into linear discriminant coefficients:

$\varepsilon^T \cdot S_{pooled} \cdot \varepsilon = (432.28746)$

<Note that $\varepsilon^T S_{pooled} \varepsilon$ are not unity

$L^{\langle i \rangle} := \left(\varepsilon^{\langle i \rangle T} \cdot S_{pooled} \cdot \varepsilon^{\langle i \rangle}\right)$ $L = (432.28746)$

<Same as scaled eigenvectors (e) in jw313 citing Rencher

$a^{\langle i \rangle} := \frac{1}{\sqrt{\left(L^{\langle i \rangle}\right)_1}} \cdot \varepsilon^{\langle i \rangle}$ **< discriminant coefficients** $a = \begin{pmatrix} -0.04459 \\ 0.01804 \end{pmatrix}$

$a^T \cdot S_{pooled} \cdot a = (1)$ **< coefficients a are now scaled so that $a^T S_{pooled} \alpha$ are unity**

$j := 1 .. p$

$\phi_j := S_{pooled_{j,j}}$

< Standardized discriminant coefficients as described by Rencher and in jw313

$\text{diag}\left(\sqrt{\phi}\right) \cdot a = \begin{pmatrix} -0.76364 \\ 0.6107 \end{pmatrix}$

These may be used to determine the "importance" of each original variable with regard to each discriminant function axis within the context of the other variables.

Discriminant scores:

object scores:

$yhat\pi_1 := \pi_1 \cdot a$

$yhat\pi_2 := \pi_2 \cdot a$

< Each ROW represents a point transformed from the original data matrix π_1, π_2, or π_3.

First COLUMNS are first discriminant scores
Second COLUMNS are second discriminant scores

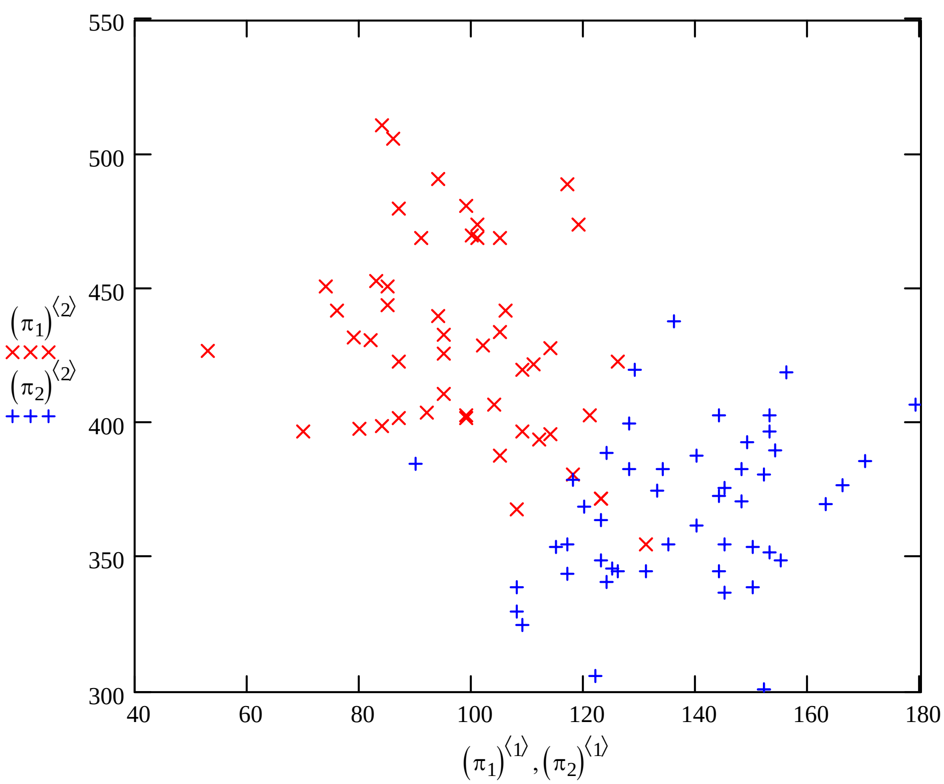
Discriminant scores may be plotted as a summary of the data.

mean scores:

$Yhat_{bar_k} := x_{bar_k}{}^T \cdot a$

$Ymeans := \text{stack}\left(Yhat_{bar_1}, Yhat_{bar_2}\right)$

$Ymeans = \begin{pmatrix} 3.3641 \\ 0.48454 \end{pmatrix}$



2005 jw631B

Mean Centered Discriminant scores (ζ):

$j := 1..p$

object scores:

$i := 1..n_1$

$\zeta1_{i,j} := x_{barGM_j} - (\pi1)_{i,j}$

$yhat\zeta_1 := \zeta_1 \cdot a$

$i := 1..n_2$

$\zeta2_{i,j} := x_{barGM_j} - (\pi2)_{i,j}$

$yhat\zeta_2 := \zeta_2 \cdot a$

$yhat\zeta_1 =$

	1
1	0.10139
2	1.36136
3	-1.85426
4	-3.36881
5	-0.91319
6	-1.82703
7	-1.82158
8	-1.68
9	-2.34606
10	-0.93123
11	-0.71341
12	0.69802
13	0.69802
14	-0.79203
15	-0.18927
16	-0.78045

$yhat\zeta_2 =$

	1
1	0.09969
2	1.8307
3	2.56347
4	1.82866
5	2.52502
6	0.43595
7	1.32154
8	1.54175
9	1.63637
10	2.12136
11	1.47844
12	0.78686
13	1.53971
14	2.54101
15	3.27175
16	1.58463

mean scores:

$Yhat_{bar_k} := \left(\left(x_{barGM} - x_{bar_k}\right)^T\right) \cdot a$

$Ymeans := stack\left(Yhat_{bar_1}, Yhat_{bar_2}\right)$

$Ymeans = \begin{pmatrix} -1.43978 \\ 1.43978 \end{pmatrix}$

< Each ROW represents a point transformed from the original data matrix π_1, π_2, or π_3.

First COLUMNS are first discriminant scores
Second COLUMNS are second discriminant scores

Discriminant scores may be plotted as a summary of the data.

2005 jw631B

MANOVA Table jw p. 299

Source	SSP Matrix	Degrees of Freedom	
Treatment Matrix B	$B = \begin{pmatrix} 38181.16 & -61590.08 \\ -61590.08 & 99351.04 \end{pmatrix}$	$df_B := g - 1$	$df_B = 1$
Residual/Error Matrix W	$W = \begin{pmatrix} 28748.2 & -2674.8 \\ -2674.8 & 1.12325 \times 10^5 \end{pmatrix}$	$df_W := N - g$	$df_W = 98$
TOTAL (mean corrected)	$B + W = \begin{pmatrix} 66929.36 & -64264.88 \\ -64264.88 & 2.11676 \times 10^5 \end{pmatrix}$	$df_T := N - 1$	$df_T = 99$

MANOVA tests jw p. 299-300.

Decomposition Model:

$X_{i,j} = \mu + \tau_i + \varepsilon_{i,j}$ **where: j = 1 to** n_m

m = 1 to g

Restriction:

$\Sigma n_m \tau_m = 0$

Assumptions:

All populations P_1**-**P_g **rs**

$\varepsilon_{i,j}$ **rs** $N(0, \sigma^2)$

Population variances equal

Hypothesis testing:

$H_0 : \tau_1 = \tau_2 = \dots \tau_g = 0$

H_1 **: at least one** τ **<> 0**

Wilk's Lambda Test Statistic:

$\Lambda_s := \frac{|W|}{|W + B|}$ $\Lambda_s = 0.321$

Lawley-Hotelling Trace:

$LHtr := tr(B \cdot W^{-1})$ $LHtr = 2.11527$

Pillai trace:

$Ptr := tr\left[B \cdot (B + W)^{-1}\right]$ $Ptr = 0.679$

Stringency of the test: $\alpha := 0.01$ **< set as desired**

If assumptions hold and H_0 **is true then jw Table 6.3 & p. 300:**

p > 0 **g = 3**

$C := qF\left[1 - \alpha, 2 \cdot p, 2 \cdot (N - p - 2)\right]$ $C = 3.41835$

Decision Rule: Reject H_0 **if** $K > C$**:**

$K := \left(\frac{N - p - 2}{p}\right) \cdot \left(\frac{1 - \sqrt{\Lambda_s}}{\sqrt{\Lambda_s}}\right)$ $K = 36.72064$

Probability: $1 - pF\left[K, 2 \cdot p, 2 \cdot (N - p - 2)\right] = 0$

$Decision := if(K > C, 1, 0)$ $Decision = 1$

< 0 = Do not reject H_0

1 = Reject H_0

2005 jw631B

Simultaneous Bonferroni confidence intervals for treatment effects jw p. 305:

$\alpha := 0.05$ **< Set probability of Type 1 error**

$$c := qt\left[1 - \frac{\alpha}{p \cdot g \cdot (g-1)}, N - g\right] \qquad c = 2.27636$$

Difference in means between first & second populations:

$i := 1 .. p$

$$\kappa_i := \left| \sqrt{\frac{W_{i,i}}{N-g} \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \right| \qquad \kappa = \begin{pmatrix} 3.42549 \\ 6.77104 \end{pmatrix}$$

$$ci_{lower} := x_{bar_1} - x_{bar_2} - c \cdot \kappa \qquad ci_{upper} := x_{bar_1} - x_{bar_2} + c \cdot \kappa$$

$$ci := \text{augment}(ci_{lower}, ci_{upper}) \qquad ci = \begin{pmatrix} -46.87765 & -31.28235 \\ 47.62667 & 78.45333 \end{pmatrix}$$

Note: According to jw Result 6.5, Bonferroni simultaneous confidence intervals apply to ALL variables (=components of mean vectors) simultaneously for ALL pairwise comparisons across groups.

2005 jw631B

Applying Fisher's Classification Procedure: **See jw Eq. 11-67 p. 638**

$m := 1..g-1 \quad l_{z_m} := 1 \quad o := 1..N \quad N = 100$

$yhat := \text{stack}(yhat\zeta_1, yhat\zeta_2)^T$

<Discriminant scores are stacked into a single list of objects (N=56), and are transposed here for computational convenience

$Obj := \text{augment}\left(M^{\langle 1 \rangle}, yhat^T\right)$

<Matrix Obj is created to show ORIGINAL membership of each object (ROWS) in a group (first COLUMN) along with discriminant scores:
COLUMN 2 = first discriminant function
COLUMN 3 = second discriminant function

$$G1_o := \left[l_z^T \cdot \left[(yhat)^{\langle o \rangle} - \left(Ymeans^T\right)^{\langle 1 \rangle} \right]^2 \right]_1$$

$$G2_o := \left[l_z^T \cdot \left[(yhat)^{\langle o \rangle} - \left(Ymeans^T\right)^{\langle 2 \rangle} \right]^2 \right]_1$$

<Squarred discriminant distances between each object and the means for each group are calculated and placed in COLUMNS of matrix G

Fisher's criterion says to place each object (ROW) in the group (COLUMNS) showing the smallest distance (numeric value in G)

e.g., Object 1 (originally in Group 1 = π_1) shows smallest distance with Group 1 in matrix G - so if it was a "new" object (not part of the training set) it would be placed (correctly) in Group 1.
Note, however, that Object 8 is not so lucky! Originally in Group 2, as a "new" object, it would be mis-classified as belonging to Group 1

$G := \text{augment}(G1, G2)$

Obj =

	1	2	3
1	1	0.10139	
2	1	1.36136	
3	1	-1.85426	
4	1	-3.36881	
5	1	-0.91319	
6	1	-1.82703	
7	1	-1.82158	
8	1	-1.68	
9	1	-2.34606	
10	1	-0.93123	
11	1	-0.71341	
12	1	0.69802	
13	1	0.69802	
14	1	-0.79203	
15	1	-0.18927	
16	1	-0.78045	
17	1	-0.73893	
18	1	-0.08819	
19	1	-1.22291	
20	1	-1.32025	
21	1	-0.13617	
22	1	-2.09523	
23	1	-1.52786	
24	1	-1.26647	
25	1	-2.0326	
26	1	-2.29501	
27	1	-0.37714	
28	1	-1.32263	

G =

	1	2	3
1	2.37521	1.79128	
2	7.84641	0.00615	
3	0.17179	10.8507	
4	3.72117	23.12258	
5	0.2773	5.53647	
6	0.14996	10.67204	
7	0.14577	10.6365	
8	0.05771	9.73304	
9	0.82135	14.3326	
10	0.25862	5.62169	
11	0.52762	4.63622	
12	4.5702	0.5502	
13	4.5702	0.5502	
14	0.41958	4.98097	
15	1.56378	2.6538	
16	0.43471	4.92944	
17	0.49119	4.74679	
18	1.82681	2.33468	
19	0.04703	7.08992	
20	0.01429	7.61778	
21	1.69939	2.48363	
22	0.42961	12.49627	
23	0.00776	8.80689	
24	0.03003	7.32381	
25	0.35144	12.05744	
26	0.73142	13.94866	
27	1.1292	3.3012	
28	0.01372	7.63093	

2005 jw631C

ORIGIN ≡ 1

FISHER'S SAMPLE LINEAR DISCRIMINANTS FOR SEVERAL POPULATIONS
jw 631C.mcd

prepared by:
Wm Stein

Using jw Table 11.7 Crude-Oil Data - Graphing the Discriminant Functions.

Five variables (columns) from three sandstones π
For this sheet we will only use variables in columns 2 & 3 so we can graph them easily.

Reading the Data:

$M := \text{READPRN}(\text{"\textbackslash DATA\textbackslash T11-7trans.txt"})$

$\pi_1 := \text{submatrix}(M,1,7,2,3)$

$\pi_2 := \text{submatrix}(M,8,18,2,3)$ $\quad p := 2$ **< variables**

$\pi_3 := \text{submatrix}(M,19,56,2,3)$ $\quad g := 3$ **<groups**

$k := 1..g \quad n_k := \text{rows}(\pi_k) \qquad n = \begin{pmatrix} 7 \\ 11 \\ 38 \end{pmatrix}$

M =

	1	2	3	4	5	6
1	1	3.9	7.1414	0.4472	0.1416	12.19
2	1	2.7	7	0.2646	0.1401	12.23
3	1	2.8	6	0.5477	0.1429	11.3
4	1	3.1	6.7082	0.2828	0.1389	13.01
5	1	3.5	6.7823	0.3162	0.128	12.63
6	1	3.9	6.5574	0.2646	0.16	10.42
7	1	2.7	5.9161	0	0.1957	9
8	2	5	6.8557	0.2646	0.1416	6.1
9	2	3.4	5.6569	0.4472	0.1718	4.69
10	2	1.2	3.4641	0	0.1805	3.15
11	2	8.4	4.1231	0.2646	0.1585	4.55
12	2	4.2	6	0.7071	0.1081	4.95
13	2	4.2	5.9161	0.7071	0.1757	2.22
14	2	3.9	6.4031	0.3162	0.1776	2.94
15	2	3.9	6	0.2646	0.1616	2.27
16	2	7.3	5.6569	0.5477	0.1247	12.92
17	2	4.4	6.7823	0.2646	0.1326	5.76
18	2	3	5.4772	0	0.1953	10.77
19	3	6.3	3.6056	0.7071	0.2358	8.27
20	3	1.7	2.3664	1	0.1757	4.64
21	3	7.3	4.899	0	0.2304	2.99
22	3	7.8	4.2426	0.7071	0.2551	6.09
23	3	7.8	5	0.8367	0.1855	6.2
24	3	7.8	5.099	1	0.1992	2.5
25	3	9.5	4.1231	0.2236	0.2841	5.71
26	3	7.7	3.7417	0.5477	0.2151	8.63
27	3	11	4.4721	0.7071	0.2342	8.4
28	3	8	3.7417	0.5477	0.2315	7.87

Summary Statistics:

Sample means:

$i_1 := 1..n_1 \quad ln1_{i_1} := 1 \quad x_{bar_1} := \frac{1}{n_1}\cdot\pi_1{}^T\cdot l_{n1} \quad x_{bar_1} = \begin{pmatrix} 3.22857 \\ 6.58649 \end{pmatrix}$

$i_2 := 1..n_2 \quad ln2_{i_2} := 1 \quad x_{bar_2} := \frac{1}{n_2}\cdot\pi_2{}^T\cdot l_{n2} \quad x_{bar_2} = \begin{pmatrix} 4.44545 \\ 5.66685 \end{pmatrix}$

$i_3 := 1..n_3 \quad ln3_{i_3} := 1 \quad x_{bar_3} := \frac{1}{n_3}\cdot\pi_3{}^T\cdot l_{n3} \quad x_{bar_3} = \begin{pmatrix} 7.22632 \\ 4.63367 \end{pmatrix}$

Grand mean:

$\Pi := \text{submatrix}(M,1,56,2,3) \quad N := \text{rows}(\Pi) \qquad N = 56$

$i := 1..N \quad l_{n_i} := 1 \qquad x_{barGM} := \frac{1}{N}\cdot\Pi^T\cdot l_n \qquad x_{barGM} = \begin{pmatrix} 6.18036 \\ 5.08072 \end{pmatrix}$

Between matrix - Eq. 11-60 p. 630:

$$B := \sum_{k=1}^{g} n_k\cdot\left(x_{bar_k} - x_{barGM}\right)\cdot\left(x_{bar_k} - x_{barGM}\right)^T \qquad B = \begin{pmatrix} 135.67315 & -60.06734 \\ -60.06734 & 27.2449 \end{pmatrix}$$

Variance-covariance matrices:

$I_1 := \text{identity}(n_1)$

$$S_1 := \frac{1}{n_1 - 1}\cdot\pi_1^T\cdot\left(I_1 - \frac{1}{n_1}\cdot l_{n1}\cdot l_{n1}^T\right)\cdot\pi_1 \qquad S_1 = \begin{pmatrix} 0.28905 & 0.12961 \\ 0.12961 & 0.22105 \end{pmatrix}$$

$I_2 := \text{identity}(n_2)$

$$S_2 := \frac{1}{n_2 - 1}\cdot\pi_2^T\cdot\left(I_2 - \frac{1}{n_2}\cdot l_{n2}\cdot l_{n2}^T\right)\cdot\pi_2 \qquad S_2 = \begin{pmatrix} 3.85273 & 0.11825 \\ 0.11825 & 1.07552 \end{pmatrix}$$

$I_3 := \text{identity}(n_3)$

$$S_3 := \frac{1}{n_3 - 1}\cdot\pi_3^T\cdot\left(I_3 - \frac{1}{n_3}\cdot l_{n3}\cdot l_{n3}^T\right)\cdot\pi_3 \qquad S_3 = \begin{pmatrix} 3.98145 & -0.00011 \\ -0.00011 & 0.80292 \end{pmatrix}$$

Pooled covariance matrix:

$$S_{pooled} := \frac{1}{\sum_{i=1}^{g}(n_i) - g}\cdot\sum_{i=1}^{g}(n_i - 1)\cdot S_i \qquad S_{pooled} = \begin{pmatrix} 3.53916 & 0.03691 \\ 0.03691 & 0.78848 \end{pmatrix}$$

^jw Eq 11-50 p. 618

Within matrix - Eq. 11-61 p. 630:

$$W := \sum_{k=1}^{g}(n_k - 1)\cdot S_k \qquad W = \begin{pmatrix} 187.57524 & 1.95623 \\ 1.95623 & 41.78941 \end{pmatrix}$$

Discriminant matrix ($W^{-1}B$):

$$W^{-1}\cdot B = \begin{pmatrix} 0.73865 & -0.32719 \\ -1.47196 & 0.66727 \end{pmatrix}$$

$\lambda := \text{reverse}(\text{sort}(\text{eigenvals}(W^{-1}\cdot B)))$

$i := 1 .. \text{rank}(W^{-1}\cdot B)$

$$\lambda = \begin{pmatrix} 1.39786 \\ 0.00806 \end{pmatrix}$$

Descriminant function coefficients:

$$\varepsilon^{\langle i\rangle} := \text{eigenvec}(W^{-1}\cdot B, \lambda_i) \qquad \varepsilon = \begin{pmatrix} -0.44459 & 0.40873 \\ 0.89574 & 0.91266 \end{pmatrix}$$

Canonical Correlations:

$$r_i := \sqrt{\frac{\lambda_i}{1 + \lambda_i}} \qquad r = \begin{pmatrix} 0.76352 \\ 0.08944 \end{pmatrix}$$

2005 jw631C

Scaling eigenvectors into linear discriminant coefficients:

$$\varepsilon^T \cdot S_{pooled} \cdot \varepsilon = \begin{pmatrix} 1.30277 & 0 \\ 0 & 1.27554 \end{pmatrix}$$

<Note that $\varepsilon^T S_{pooled} \varepsilon$ are not unity

$$L^{\langle i \rangle} := \left(\varepsilon^{\langle i \rangle T} \cdot S_{pooled} \cdot \varepsilon^{\langle i \rangle}\right) \qquad L = (1.30277 \quad 1.27554)$$

<Same as scaled eigenvectors (e) in jw313 citing Rencher

$$a^{\langle i \rangle} := \frac{1}{\sqrt{\left(L^{\langle i \rangle}\right)_1}} \cdot \varepsilon^{\langle i \rangle}$$

< discriminant coefficients

$$a = \begin{pmatrix} -0.38951 & 0.3619 \\ 0.78478 & 0.80809 \end{pmatrix}$$

$$a^T \cdot S_{pooled} \cdot a = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

< coefficients a are now scaled so that $a^T S_{pooled} \alpha$ are unity

$$j := 1 .. p$$

$$\phi_j := S_{pooled_{j,j}}$$

< Standardized discriminant coefficients as described by Rencher and in jw313

$$\mathrm{diag}\left(\sqrt{\phi}\right) \cdot a = \begin{pmatrix} -0.73278 & 0.68083 \\ 0.69685 & 0.71755 \end{pmatrix}$$

These may be used to determine the "importance" of each original variable with regard to each discriminant function axis within the context of the other variables.

Discriminant scores:

object scores:

$$yhat\pi_1 := \pi_1 \cdot a$$

$$yhat\pi_2 := \pi_2 \cdot a$$

$$yhat\pi_3 := \pi_3 \cdot a$$

< Each ROW represents a point transformed from the original data matrix π_1, π_2, or π_3.

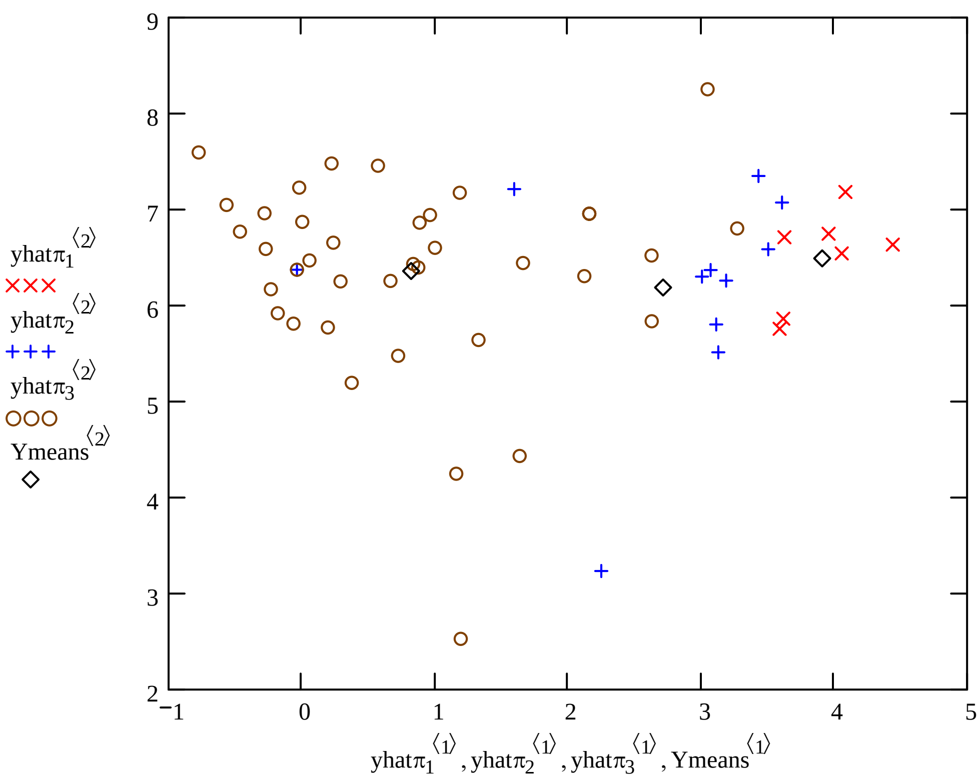
First COLUMNS are first discriminant scores
Second COLUMNS are second discriminant scores

Discriminant scores may be plotted as a summary of the data.

mean scores:

$$Yhat_{bar_k} := x_{bar_k}^{\ T} \cdot a$$

$$Ymeans := \mathrm{stack}\left(Yhat_{bar_1}, Yhat_{bar_2}, Yhat_{bar_3}\right)$$



$$Ymeans = \begin{pmatrix} 3.91135 & 6.49089 \\ 2.71565 & 6.18814 \\ 0.82165 & 6.35962 \end{pmatrix}$$

2005 jw631C

Mean Centered Discriminant scores (ζ):

$j := 1 .. p$

object scores:

$i := 1 .. n_1$

$\zeta1_{i,j} := xbarGM_j - (\pi1)_{i,j}$

$yhat\zeta_1 := \zeta_1 \cdot a$

$$yhat\zeta_1 = \begin{pmatrix} -2.5054 & -0.83996 \\ -2.86185 & -0.29142 \\ -2.03812 & 0.48049 \\ -2.47705 & -0.20037 \\ -2.37939 & -0.40501 \\ -2.04709 & -0.36803 \\ -2.01123 & 0.58447 \end{pmatrix}$$

$i := 1 .. n_2$

$\zeta2_{i,j} := xbarGM_j - (\pi2)_{i,j}$

$yhat\zeta_2 := \zeta_2 \cdot a$

$$yhat\zeta_2 = \begin{pmatrix} -1.85273 & -1.00718 \\ -1.53516 & 0.5406 \\ -0.67123 & 3.10876 \\ 1.6161 & -0.02944 \\ -1.49281 & -0.02617 \\ -1.42696 & 0.04163 \\ -1.926 & -0.24335 \\ -1.60966 & 0.0824 \\ -0.01606 & -0.8708 \\ -2.02883 & -0.73072 \\ -1.54994 & 0.83057 \end{pmatrix}$$

$i := 1 .. n_3$

$\zeta3_{i,j} := xbarGM_j - (\pi3)_{i,j}$

$yhat\zeta_3 := \zeta_3 \cdot a$

$$yhat\zeta_3 = \begin{pmatrix} 1.20424 & 1.14873 \\ 0.38497 & 3.81485 \\ 0.57872 & -0.25835 \\ 1.28861 & 0.09113 \\ 0.69422 & -0.52092 \\ 0.61652 & -0.60092 \\ 2.04456 & -0.42753 \\ 1.64275 & 0.53209 \\ 2.35494 & -1.25241 \\ 1.7596 & 0.42352 \\ 1.52231 & -0.12601 \\ 2.14554 & -0.70505 \\ 1.3839 & 0.57075 \\ 1.59936 & -0.88596 \\ 1.81237 & 0.17266 \\ 1.6161 & -0.02944 \\ 1.35639 & -1.13615 \\ 0.70348 & -0.0536 \\ 0.41941 & 2.09544 \\ -1.46947 & -1.9096 \\ 1.00766 & -1.11375 \\ 0.39192 & -0.8322 \\ -1.69229 & -0.46048 \\ -0.58115 & -0.6134 \\ 0.25157 & 0.70185 \\ 1.8498 & -0.24658 \\ 1.34221 & -0.31147 \\ 1.85951 & -0.61808 \\ 1.57592 & -0.52861 \\ 0.85577 & 0.86621 \\ 0.91374 & 0.08662 \\ -0.05729 & 1.90977 \\ -0.58115 & -0.6134 \\ 0.74243 & -0.08979 \\ -1.04928 & 0.50683 \\ -0.54291 & 0.03636 \\ -1.04857 & -0.17913 \\ -0.08298 & -0.10043 \end{pmatrix}$$

mean scores:

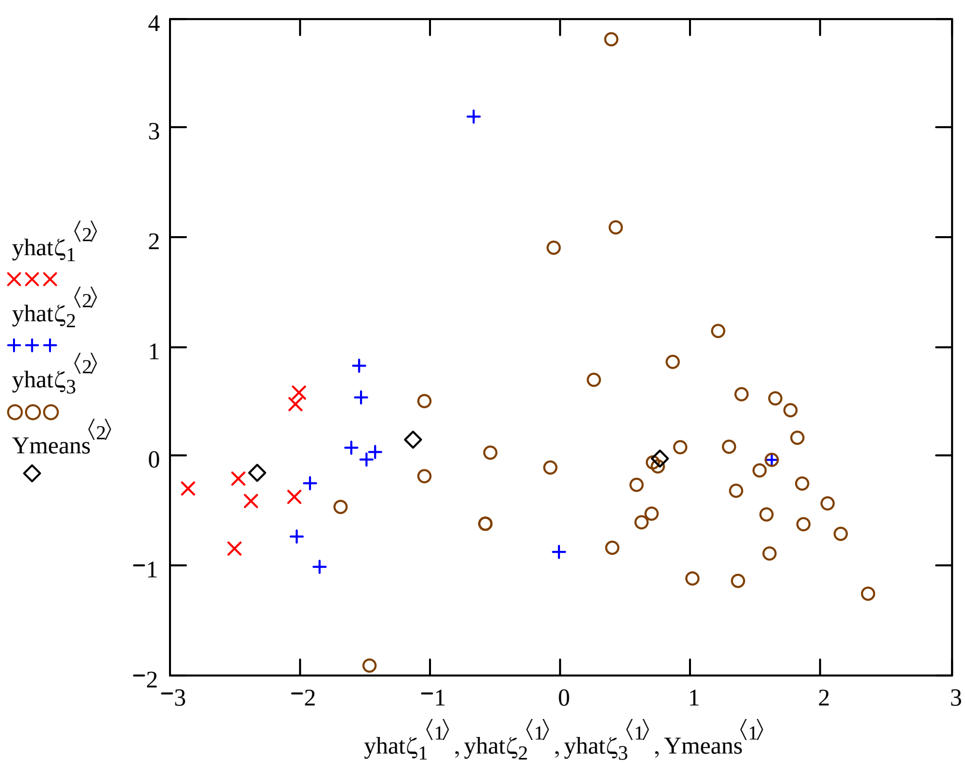
$Yhatbar_k := \left(\left(xbarGM - xbar_k\right)^T\right) \cdot a$

$Ymeans := stack\left(Yhatbar_1, Yhatbar_2, Yhatbar_3\right)$

< Each ROW represents a point transformed from the original data matrix π_1, π_2, or π_3.

First COLUMNS are first discriminant scores
Second COLUMNS are second discriminant scores

Discriminant scores may be plotted as a summary of the data.

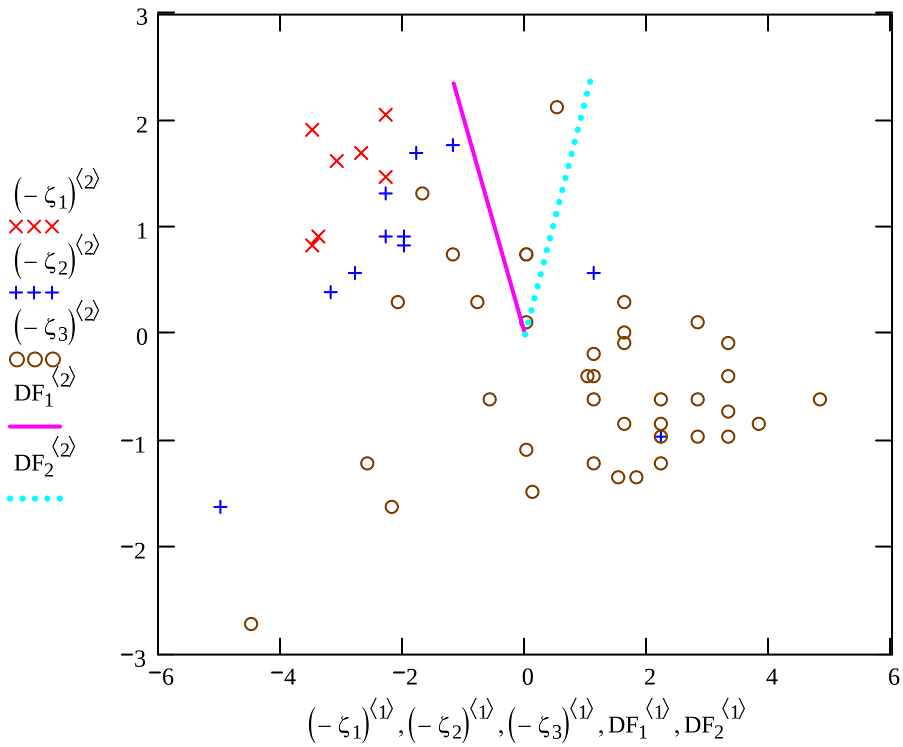


$$Ymeans = \begin{pmatrix} -2.33145 & -0.14855 \\ -1.13575 & 0.15421 \\ 0.75825 & -0.01728 \end{pmatrix}$$

2005 jw631C

Plotting the original variables for each group and discriminant function directions:

$$Mzero := \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad DF_1 := 3 \cdot augment\left(Mzero, a^{\langle 1 \rangle}\right)^T \qquad DF_2 := 3 \cdot \left(augment\left(Mzero, a^{\langle 2 \rangle}\right)\right)^T$$



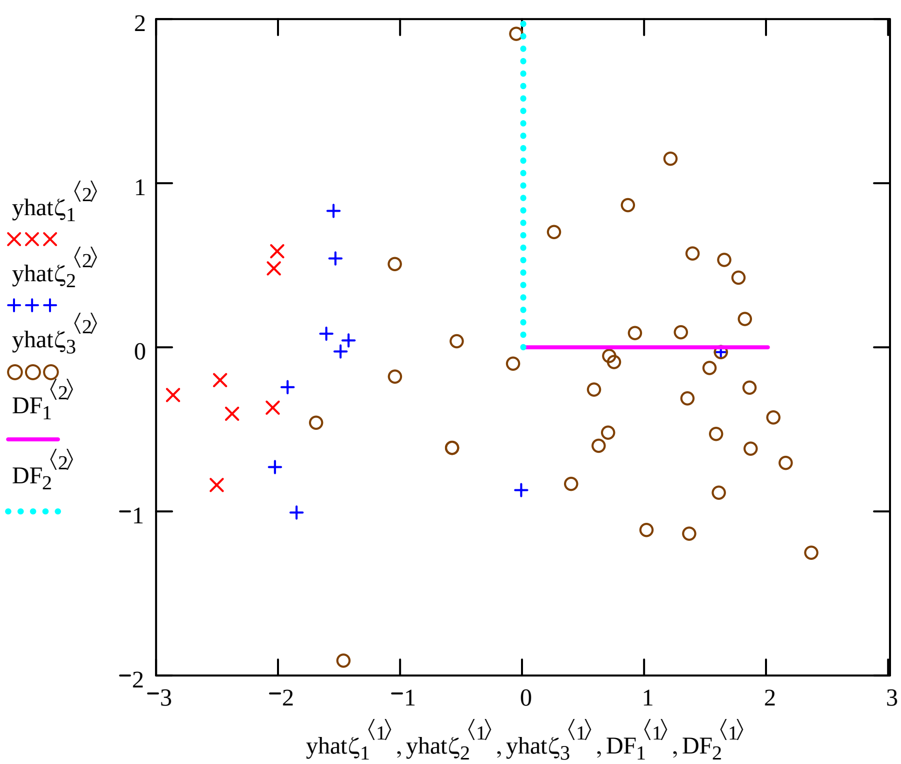
$$W^{-1} \cdot B = \begin{pmatrix} 0.73865 & -0.32719 \\ -1.47196 & 0.66727 \end{pmatrix}$$

<Plot of Original Datapoints (reversed in sign to allow simple comparisons)

Note: since $W^{-1}B$ - the Discriminant Matrix is typically not symmetric, the eigenvectors (discriminant coefficients) are not at right angles to each other. As with other linear transformations, the discriminant matrix may be viewed as a way to distort the multivariate space in a way that "maximally" displays the difference between means for the groups.

Here in both graphs:
first discriminant function = solid lines
second discriminant function = dotted lines

$$DF_1 := \begin{pmatrix} 0 & 0 \\ 2 & 0 \end{pmatrix} \qquad DF_2 := \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$$



< Plot of discriminant function "scores"

Discriminant function "scores" are the projections of the datapoints onto the discriminant function axes. Although the discriminant functions are typically not at right angles to each other, they are usually plotted as if they were. Since the objective of such a plot is "ordination" (data reduction), this does little harm.

2005 jw631C

MANOVA Table jw p. 299

Source	**SSP Matrix**	**Degrees of Freedom**	
Treatment Matrix B	$B = \begin{pmatrix} 135.67315 & -60.06734 \\ -60.06734 & 27.2449 \end{pmatrix}$	$df_B := g - 1$	$df_B = 2$
Residual/Error Matrix W	$W = \begin{pmatrix} 187.57524 & 1.95623 \\ 1.95623 & 41.78941 \end{pmatrix}$	$df_W := N - g$	$df_W = 53$
TOTAL (mean corrected)	$B + W = \begin{pmatrix} 323.24839 & -58.11111 \\ -58.11111 & 69.03431 \end{pmatrix}$	$df_T := N - 1$	$df_T = 55$

MANOVA tests jw p. 299-300.

Decomposition Model:

$X_{i,j} = \mu + \tau_i + \varepsilon_{i,j}$ **where:** $j = 1$ to n_m, $m = 1$ to g

Restriction:

$\Sigma n_m \tau_m = 0$

Assumptions:

All populations P_1-P_g **rs**

$\varepsilon_{i,j}$ **rs** $N(0, \sigma^2)$

Population variances equal

Hypothesis testing:

$H_0 : \tau_1 = \tau_2 = ... \tau_g = 0$

H_1 **: at least one** τ **<> 0**

Wilk's Lambda Test Statistic:

$$\Lambda_s := \frac{|W|}{|W + B|} \qquad \Lambda_s = 0.4137$$

Lawley-Hotelling Trace:

$$LHtr := tr\left(B \cdot W^{-1}\right) \qquad LHtr = 1.40592$$

Pillai trace:

$$Ptr := tr\left[B \cdot (B + W)^{-1}\right] \qquad Ptr = 0.59096$$

Stringency of the test: $\alpha := 0.01$ **< set as desired**

If assumptions hold and H_0 **is true then jw Table 6.3 & p. 300:**

p > 0 **g = 3**

$$C := qF\left[1 - \alpha, 2 \cdot p, 2 \cdot (N - p - 2)\right] \qquad C = 3.505$$

Decision Rule: Reject H_0 **if** $K > C$**:**

$$K := \left(\frac{N - p - 2}{p}\right) \cdot \left(\frac{1 - \sqrt{\Lambda_s}}{\sqrt{\Lambda_s}}\right) \qquad K = 14.42307$$

Probability: $1 - pF\left[K, 2 \cdot p, 2 \cdot (N - p - 2)\right] = 2.11407 \times 10^{-9}$

$Decision := if(K > C, 1, 0)$ $\qquad Decision = 1$

< 0 = Do not reject H_0
1 = Reject H_0

Simultaneous Bonferroni confidence intervals for treatment effects jw p. 305:

$\alpha := 0.05$ **< Set probability of Type 1 error**

$c := qt\left[1 - \frac{\alpha}{p \cdot g \cdot (g-1)}, N - g\right]$ $c = 2.74091$

Difference in means between first & second populations:

$i := 1 .. p$

$\kappa_i := \left|\sqrt{\frac{W_{i,i}}{N-g} \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}\right|$ $\kappa = \begin{pmatrix} 0.90958 \\ 0.42932 \end{pmatrix}$

$ci_{lower} := x_{bar_1} - x_{bar_2} - c \cdot \kappa$ $ci_{upper} := x_{bar_1} - x_{bar_2} + c \cdot \kappa$

$ci := \text{augment}(ci_{lower}, ci_{upper})$ $ci = \begin{pmatrix} -3.70996 & 1.27619 \\ -0.25711 & 2.09637 \end{pmatrix}$

Difference in means between first & third populations:

$i := 1 .. p$

$\kappa_i := \left|\sqrt{\frac{W_{i,i}}{N-g} \cdot \left(\frac{1}{n_1} + \frac{1}{n_3}\right)}\right|$ $\kappa = \begin{pmatrix} 0.77378 \\ 0.36523 \end{pmatrix}$

$ci_{lower} := x_{bar_1} - x_{bar_3} - c \cdot \kappa$ $ci_{upper} := x_{bar_1} - x_{bar_3} + c \cdot \kappa$

$ci := \text{augment}(ci_{lower}, ci_{upper})$ $ci = \begin{pmatrix} -6.11859 & -1.8769 \\ 0.95177 & 2.95386 \end{pmatrix}$

Note: According to jw Result 6.5, Bonferroni simultaneous confidence intervals apply to ALL variables (=components of mean vectors) simultaneously for ALL pairwise comparisons across groups.

Difference in means between second & third populations:

$i := 1 .. p$

$\kappa_i := \left|\sqrt{\frac{W_{i,i}}{N-g} \cdot \left(\frac{1}{n_2} + \frac{1}{n_3}\right)}\right|$ $\kappa = \begin{pmatrix} 0.64411 \\ 0.30402 \end{pmatrix}$

$ci_{lower} := x_{bar_2} - x_{bar_3} - c \cdot \kappa$ $ci_{upper} := x_{bar_2} - x_{bar_3} + c \cdot \kappa$

$ci := \text{augment}(ci_{lower}, ci_{upper})$ $ci = \begin{pmatrix} -4.5463 & -1.01542 \\ 0.19989 & 1.86648 \end{pmatrix}$

2005 jw631C

Applying Fisher's Classification Procedure: **See jw Eq. 11-67 p. 638**

$m := 1..g-1 \quad l_{z_m} := 1 \quad o := 1..N \quad N = 56$

<Discriminant scores are stacked into a single list of objects (N=56), and are transposed here for computational convenience

$yhat := stack(yhat\zeta_1, yhat\zeta_2, yhat\zeta_3)^T$

<Matrix Obj is created to show ORIGINAL membership of each object (ROWS) in a group (first COLUMN) along with discriminant scores:
COLUMN 2 = first discriminant function
COLUMN 3 = second discriminant function

$Obj := augment\left(M^{\langle 1 \rangle}, yhat^T\right)$

$$G1_o := \left[l_z^T \cdot \left[(yhat)^{\langle o \rangle} - \left(Ymeans^T\right)^{\langle 1 \rangle} \right]^2 \right]_1$$

$$G2_o := \left[l_z^T \cdot \left[(yhat)^{\langle o \rangle} - \left(Ymeans^T\right)^{\langle 2 \rangle} \right]^2 \right]_1$$

$$G3_o := \left[l_z^T \cdot \left[(yhat)^{\langle o \rangle} - \left(Ymeans^T\right)^{\langle 3 \rangle} \right]^2 \right]_1$$

$G := augment(G1, G2, G3)$

<Squarred discriminant distances between each object and the means for each group are calculated and placed in COLUMNS of matrix G

Fisher's criterion says to place each object (ROW) in the group (COLUMNS) showing the smallest distance (numeric value in G)

e.g., Object 1 (originally in Group 1 = π_1) shows smallest distance with Group 1 in matrix G - so if it was a "new" object (not part of the training set) it would be placed (correctly) in Group 1.
Note, however, that Object 8 is not so lucky! Originally in Group 2, as a "new" object, it would be mis-classified as belonging to Group 1

Obj =

	1	2	3
1	1	-2.5054	-0.83996
2	1	-2.86185	-0.29142
3	1	-2.03812	0.48049
4	1	-2.47705	-0.20037
5	1	-2.37939	-0.40501
6	1	-2.04709	-0.36803
7	1	-2.01123	0.58447
8	2	-1.85273	-1.00718
9	2	-1.53516	0.5406
10	2	-0.67123	3.10876
11	2	1.6161	-0.02944
12	2	-1.49281	-0.02617
13	2	-1.42696	0.04163
14	2	-1.926	-0.24335
15	2	-1.60966	0.0824
16	2	-0.01606	-0.8708
17	2	-2.02883	-0.73072
18	2	-1.54994	0.83057
19	3	1.20424	1.14873
20	3	0.38497	3.81485
21	3	0.57872	-0.25835
22	3	1.28861	0.09113
23	3	0.69422	-0.52092
24	3	0.61652	-0.60092
25	3	2.04456	-0.42753
26	3	1.64275	0.53209
27	3	2.35494	-1.25241
28	3	1.7596	0.42352

G =

	1	2	3
1	0.50831	2.86431	11.32822
2	0.30174	3.17799	13.18027
3	0.48172	0.92073	8.06746
4	0.02388	1.9248	10.50066
5	0.06807	1.85937	9.99514
6	0.12903	1.10328	7.99297
7	0.63986	0.95159	8.03212
8	0.96642	1.86287	7.7971
9	1.109	0.30882	5.57094
10	13.36638	8.94516	11.81551
11	15.59729	7.6064	0.73605
12	0.7183	0.16002	5.06732
13	0.85426	0.09748	4.77861
14	0.17337	0.78254	7.25631
15	0.57432	0.22974	5.61692
16	5.88269	2.30437	1.32806
17	0.4305	1.58069	8.27683
18	1.56944	0.62902	6.04658
19	14.18402	6.46464	1.55848
20	23.08749	15.71292	14.82453
21	8.48116	3.10964	0.09035
22	13.16225	5.8815	0.29303
23	9.29332	3.80459	0.25776
24	8.89519	3.6407	0.36073
25	19.22728	10.45281	1.82291
26	16.25752	7.86287	1.08414
27	23.18076	14.16352	4.07498
28	17.06397	8.45562	1.19701

2005 jw670

COMMON SIMULARITY/DISTANCE MEASURES
jw670.mcd

prepared by:
Wm Stein

ORIGIN ≡ 1

Table 12-5 Utility Data jw p. 687

Rows = utilities
Columns= measured variables

M := READPRN("\DATA\T12-5.DAT")

M =

	1	2	3
1	1.06	9.2	151
2	0.89	10.3	202
3	1.43	15.4	113
4	1.02	11.2	168
5	1.49	8.8	192

$n := rows(M) \qquad p := cols(M)$

$i := 1 .. n \qquad k := 1 .. n \qquad j := 1 .. p$

$x^{\langle i\rangle} := \left(M^T\right)^{\langle i\rangle} \qquad y^{\langle i\rangle} := \left(M^T\right)^{\langle i\rangle}$

DISTANCES between objects (rows):

Eucledean Distances:

$$E_{i,k} := \sqrt{\left|\left(x^{\langle i\rangle} - y^{\langle k\rangle}\right)^T \cdot \left(x^{\langle i\rangle} - y^{\langle k\rangle}\right)\right|}$$

E =

	1	2	3
1	0	3989.41	140.4
2	3989.41	0	4125.04
3	140.4	4125.04	0
4	2654.28	1335.47	2789.76
5	5777.17	1788.07	5912.55
6	2050.53	6039.69	1915.16
7	1435.27	2554.29	1571.3

Mahalanobis (Statistical) Distances (Squared):

$I := identity(n) \qquad l_{n_i} := 1$ **< I = identity matrix, l_n = identity vector**

$$S := \frac{1}{n-1} \cdot M^T \cdot \left(I - \frac{1}{n} \cdot l_n \cdot l_n^{\,T}\right) \cdot M$$ **< Variance-covariance matrix**

$$D_{i,k} := \left|\left(x^{\langle i\rangle} - y^{\langle k\rangle}\right)^T \cdot S^{-1} \cdot \left(x^{\langle i\rangle} - y^{\langle k\rangle}\right)\right|$$

D =

	1	2	3
1	0	17.09	15.99
2	17.09	0	19.91
3	15.99	19.91	0
4	7.23	6.21	17.96
5	20.37	27.47	23.27
6	23.2	22.39	19.18

Manhattan (City Block) Distances:

$$C_{i,k} := \sum_j \left|\left(x^{\langle i\rangle} - y^{\langle k\rangle}\right)_j\right|$$

C =

	1	2	3
1	0	4071.6	183.2
2	4071.6	0	4250.54
3	183.2	4250.54	0
4	2710.31	1383.68	2889.37
5	5839.65	1818.19	6018.45
6	2127.07	6143.24	1954.29
7	1476.98	2618.33	1651.8

Minkowski metrics:

$m := 1$ **< Set as desired (m=1 is city block, m=2 is Eucledean)**

$$MM_{i,k} := \left[\sum_j \left[\left|\left(x^{\langle i\rangle} - y^{\langle k\rangle}\right)_j\right|\right]^m\right]^{\frac{1}{m}}$$

MM =

	1	2	3
1	0	4071.6	183.2
2	4071.6	0	4250.54
3	183.2	4250.54	0
4	2710.31	1383.68	2889.37
5	5839.65	1818.19	6018.45
6	2127.07	6143.24	1954.29

ORIGIN ≡ 1

CLUSTERING - SINGLE LINK
jw683.mcd

prepared by:
Wm Stein

Example Agglomerative Hierarchical Clustering using the Single Link criterion - Example 12.5 jw p. 683

M := READPRN("\DATA\T12-4revised.txt")

Converting simularities to distances:

$D := 10 - M$

Single Link Criterion = MINIMUM distances between groups

D =

	1	2	3	4	5	6	7	8	9	10	11
1	0	2	2	7	6	6	6	6	7	9	9
2	2	0	1	5	4	6	6	6	7	8	9
3	2	1	0	6	5	6	5	5	6	8	9
4	7	5	6	0	5	9	9	9	10	8	9
5	6	4	5	5	0	7	7	7	8	9	9
6	6	6	6	9	7	0	2	1	3	10	9
7	6	6	5	9	7	2	0	1	3	10	9
8	6	6	5	9	7	1	1	0	4	10	9
9	7	7	6	10	8	5	3	4	0	10	9
10	9	8	8	8	9	10	10	10	10	0	8
11	9	9	9	9	9	9	9	9	9	8	0

2005 jw683

CLUSTERING - COMPLETE LINK

Example Agglomerative Hierarchical Clustering using the Complete Link criterion - Example 12.5 jw p. 683

M := READPRN("\DATA\T12-4revised.txt")

Converting simularities to distances:

$D := 10 - M$

Complete Link Criterion = MAXIMUM distances between groups

D =

	1	2	3	4	5	6	7	8	9	10	11
1	0	2	2	7	6	6	6	6	7	9	9
2	2	0	1	5	4	6	6	6	7	8	9
3	2	1	0	6	5	6	5	5	6	8	9
4	7	5	6	0	5	9	9	9	10	8	9
5	6	4	5	5	0	7	7	7	8	9	9
6	6	6	6	9	7	0	2	1	3	10	9
7	6	6	5	9	7	2	0	1	3	10	9
8	6	6	5	9	7	1	1	0	4	10	9
9	7	7	6	10	8	5	3	4	0	10	9
10	9	8	8	8	9	10	10	10	10	0	8
11	9	9	9	9	9	9	9	9	9	8	0

CLUSTERING - UNWEIGHTED PAIR-GROUP USING AVERAGES (UPGMA)

Example Agglomerative Hierarchical Clustering using the UPGMA criterion - Example 12.5 jw p. 683

$M := \text{READPRN}(\text{"\textbackslash DATA\textbackslash T12-4revised.txt"})$

Converting simularities to distances:

$D := 10 - M$

UPGMA Criterion = Take the UNWEIGHTED average distance between groupsMAXIMUM distances between groups

D =

	1	2	3	4	5	6	7	8	9	10	11
1	0	2	2	7	6	6	6	6	7	9	9
2	2	0	1	5	4	6	6	6	7	8	9
3	2	1	0	6	5	6	5	5	6	8	9
4	7	5	6	0	5	9	9	9	10	8	9
5	6	4	5	5	0	7	7	7	8	9	9
6	6	6	6	9	7	0	2	1	3	10	9
7	6	6	5	9	7	2	0	1	3	10	9
8	6	6	5	9	7	1	1	0	4	10	9
9	7	7	6	10	8	5	3	4	0	10	9
10	9	8	8	8	9	10	10	10	10	0	8
11	9	9	9	9	9	9	9	9	9	8	0

2005 jw683

CLUSTERING - WEIGHTED PAIR-GROUP USING AVERAGES (WPGMA)

Example Agglomerative Hierarchical Clustering using the WPGMA criterion - Example 12.5 jw p. 683

M := READPRN("\DATA\T12-4revised.txt")

Converting simularities to distances:

WPGMA Criterion = Take the WEIGHTED average distance between groupsMAXIMUM distances between groups. The last element linked to a group counts as much as the others in the group combined.

D =

	1	2	3	4	5	6	7	8	9	10	11
1	0	2	2	7	6	6	6	6	7	9	9
2	2	0	1	5	4	6	6	6	7	8	9
3	2	1	0	6	5	6	5	5	6	8	9
4	7	5	6	0	5	9	9	9	10	8	9
5	6	4	5	5	0	7	7	7	8	9	9
6	6	6	6	9	7	0	2	1	3	10	9
7	6	6	5	9	7	2	0	1	3	10	9
8	6	6	5	9	7	1	1	0	4	10	9
9	7	7	6	10	8	5	3	4	0	10	9
10	9	8	8	8	9	10	10	10	10	0	8
11	9	9	9	9	9	9	9	9	9	8	0

CORRESPONDENCE ANALYSIS
jw715.mcd

prepared by:
Wm Stein

ORIGIN ≡ 1

Example 12.18 jw p. 715 Contingency table with two variables: A (three levels) & B (two levels)

M := READPRN("\DATA\T12-10.DAT")

$r := \mathrm{rows}(M) \quad r = 8 \qquad c := \mathrm{cols}(M) \quad c = 10$

$i := 1..r \quad j := 1..c$

$l_{r_i} := 1 \quad l_{c_j} := 1$

M =

	1	2	3	4	5	6	7	8	9	10
1	9	8	3	5	6	0	5	0	0	0
2	8	9	8	7	0	0	0	0	0	0
3	5	4	9	9	7	7	4	6	0	2
4	3	4	0	6	9	8	7	6	4	3
5	2	2	4	5	6	0	5	0	2	5
6	0	0	0	0	2	7	6	6	7	6
7	0	0	0	0	0	0	7	4	6	5
8	0	0	0	0	0	5	4	8	8	9

Calculating Total number of observations in matrix M:

$$T := \sum_i \sum_j M_{i,j} \qquad T = 292$$

Matrix P of frequencies:

$$P := \frac{1}{T}M$$

$$P = \begin{pmatrix} 0.03082 & 0.0274 & 0.01027 & 0.01712 & 0.02055 & 0 & 0.01712 & 0 & 0 & 0 \\ 0.0274 & 0.03082 & 0.0274 & 0.02397 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.01712 & 0.0137 & 0.03082 & 0.03082 & 0.02397 & 0.02397 & 0.0137 & 0.02055 & 0 & 0.00685 \\ 0.01027 & 0.0137 & 0 & 0.02055 & 0.03082 & 0.0274 & 0.02397 & 0.02055 & 0.0137 & 0.01027 \\ 0.00685 & 0.00685 & 0.0137 & 0.01712 & 0.02055 & 0 & 0.01712 & 0 & 0.00685 & 0.01712 \\ 0 & 0 & 0 & 0 & 0.00685 & 0.02397 & 0.02055 & 0.02055 & 0.02397 & 0.02055 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.02397 & 0.0137 & 0.02055 & 0.01712 \\ 0 & 0 & 0 & 0 & 0 & 0.01712 & 0.0137 & 0.0274 & 0.0274 & 0.03082 \end{pmatrix}$$

Calculating Row (R) and Column (C) vectors of sums of matrix P:

$R := P \cdot l_c$

$C := P^T \cdot l_r$

$D_r := \mathrm{diag}(R)$

$D_c := \mathrm{diag}(C)$

$$R = \begin{pmatrix} 0.12329 \\ 0.10959 \\ 0.18151 \\ 0.17123 \\ 0.10616 \\ 0.11644 \\ 0.07534 \\ 0.11644 \end{pmatrix} \qquad C = \begin{pmatrix} 0.09247 \\ 0.09247 \\ 0.08219 \\ 0.10959 \\ 0.10274 \\ 0.09247 \\ 0.13014 \\ 0.10274 \\ 0.09247 \\ 0.10274 \end{pmatrix}$$

CalculAting diagonal inverse squaret matrices $D_r^{-1/2}$ & $D_c^{-1/2}$ and matrix A:

$$\mathrm{Dinvsqrt}_r := \mathrm{diag}\left(\frac{1}{\sqrt{R}}\right)$$

$$\mathrm{Dinvsqrt}_c := \mathrm{diag}\left(\frac{1}{\sqrt{C}}\right)$$

$$A := \mathrm{Dinvsqrt}_r \cdot \left(P - R \cdot C^T\right) \cdot \mathrm{Dinvsqrt}_c$$

$$A = \begin{pmatrix} 0.18 & 0.15 & 0 & 0.03 & 0.07 & -0.11 & 0.01 & -0.11 & -0.11 & -0.11 \\ 0.17 & 0.21 & 0.19 & 0.11 & -0.11 & -0.1 & -0.12 & -0.11 & -0.1 & -0.11 \\ 0 & -0.02 & 0.13 & 0.08 & 0.04 & 0.06 & -0.06 & 0.01 & -0.13 & -0.09 \\ -0.04 & -0.02 & -0.12 & 0.01 & 0.1 & 0.09 & 0.01 & 0.02 & -0.02 & -0.06 \\ -0.03 & -0.03 & 0.05 & 0.05 & 0.09 & -0.1 & 0.03 & -0.1 & -0.03 & 0.06 \\ -0.1 & -0.1 & -0.1 & -0.11 & -0.05 & 0.13 & 0.04 & 0.08 & 0.13 & 0.08 \\ -0.08 & -0.08 & -0.08 & -0.09 & -0.09 & -0.08 & 0.14 & 0.07 & 0.16 & 0.11 \\ -0.1 & -0.1 & -0.1 & -0.11 & -0.11 & 0.06 & -0.01 & 0.14 & 0.16 & 0.17 \end{pmatrix}$$

2005 jw715

Singular Value Decomposition of a Matrix (from Matrix Algebra Toolbox jw 084.mcd):

$m := \text{rows}(A) \quad k := \text{cols}(A)$

$\gamma_U := \text{reverse}\left(\text{sort}\left(\text{eigenvals}\left(A \cdot A^T\right)\right)\right)$

$\gamma_V := \text{reverse}\left(\text{sort}\left(\text{eigenvals}\left(A^T \cdot A\right)\right)\right)$

$i := 1 .. \text{rows}\left(A \cdot A^T\right) \quad j := 1 .. \text{rows}\left(A^T \cdot A\right)$

$U^{\langle i \rangle} := \text{eigenvec}\left(A \cdot A^T, \gamma_{U_i}\right)$

$V^{\langle j \rangle} := \text{eigenvec}\left(A^T \cdot A, \gamma_{V_j}\right)$

$\lambda := \sqrt{\gamma_U}$

$i := 1 .. k - m \quad k - m = 2 \quad Z_{m,i} := 0$

$\Lambda := \text{augment}(\text{diag}(\lambda), Z)$

$$\gamma_U = \begin{pmatrix} 0.53673 \\ 0.09618 \\ 0.0721 \\ 0.04555 \\ 0.01107 \\ 0.00454 \\ 0.00388 \\ 0 \end{pmatrix} \quad \gamma_V = \begin{pmatrix} 0.53673 \\ 0.09618 \\ 0.0721 \\ 0.04555 \\ 0.01107 \\ 0.00454 \\ 0.00388 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \lambda = \begin{pmatrix} 0.73262 \\ 0.31013 \\ 0.26851 \\ 0.21342 \\ 0.10519 \\ 0.06737 \\ 0.06231 \\ 3.00387 \times 10^{-9} \end{pmatrix} \quad Z = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$U = \begin{pmatrix} -0.39042 & 0.08311 & 0.47814 & -0.4562 & -0.03768 & -0.33691 & -0.40707 & 0.35112 \\ -0.53268 & 0.4985 & -0.40802 & -0.09251 & -0.07379 & 0.34196 & 0.24645 & 0.33104 \\ -0.19986 & -0.38891 & -0.40889 & 0.36216 & 0.43908 & -0.32169 & -0.18076 & 0.42604 \\ 0.06979 & -0.5382 & 0.17259 & -0.31809 & -0.05439 & 0.1596 & 0.61224 & 0.4138 \\ -0.08195 & 0.01512 & 0.42706 & 0.7086 & -0.41601 & 0.16851 & -0.03071 & 0.32583 \\ 0.40046 & -0.08307 & -0.14776 & -0.18661 & -0.00416 & 0.58951 & -0.55865 & 0.34123 \\ 0.36336 & 0.48503 & 0.32324 & 0.09371 & 0.62981 & -0.01639 & 0.21722 & 0.27449 \\ 0.46892 & 0.24757 & -0.315 & -0.07263 & -0.47713 & -0.51421 & 0.07627 & 0.34123 \end{pmatrix}$$

$$V = \begin{pmatrix} 0.38774 & -0.21076 & 0.06164 & -0.40292 & -0.0582 & -0.32693 & 0.42471 & 0.16413 & 0.16413 & 0.16417 \\ 0.38558 & -0.24276 & 0.01057 & -0.43452 & -0.19499 & 0.19683 & -0.26351 & 0.43824 & 0.43824 & 0.4382 \\ 0.34953 & -0.18206 & -0.40795 & 0.57182 & 0.23432 & 0.11673 & 0.32936 & 0.34774 & 0.34774 & 0.34774 \\ 0.30064 & 0.13553 & -0.05396 & 0.26465 & 0.00063 & 0.08255 & -0.66442 & 0.17881 & 0.17882 & 0.17884 \\ 0.11083 & 0.58167 & 0.48562 & 0.15977 & -0.23333 & -0.16067 & 0.07718 & 0.39287 & 0.39287 & 0.39288 \\ -0.20218 & 0.54005 & -0.46257 & -0.26868 & -0.09782 & 0.39427 & 0.26684 & 0.2543 & 0.25431 & 0.25431 \\ -0.18519 & -0.07562 & 0.50896 & 0.02909 & 0.60261 & 0.19554 & 0.15201 & 0.32748 & 0.32748 & 0.32747 \\ -0.31395 & 0.06439 & -0.33937 & -0.15671 & 0.33659 & -0.65731 & -0.25071 & 0.36671 & 0.36671 & 0.36671 \\ -0.42003 & -0.34836 & 0.0394 & -0.11651 & -0.06249 & 0.37722 & -0.14565 & 0.29052 & 0.29052 & 0.29056 \\ -0.35495 & -0.28969 & 0.0345 & 0.33932 & -0.59944 & -0.20021 & 0.12616 & 0.28805 & 0.28805 & 0.28804 \end{pmatrix}$$

2005 jw715

Initeria:

$i := 1 .. \text{rows}(\lambda)$

$TI := \sum_i \gamma U_i \qquad TI = 0.77005$

$$\gamma U = \begin{pmatrix} 0.53673 \\ 0.09618 \\ 0.0721 \\ 0.04555 \\ 0.01107 \\ 0.00454 \\ 0.00388 \\ 0 \end{pmatrix} \quad \lambda = \begin{pmatrix} 0.73262 \\ 0.31013 \\ 0.26851 \\ 0.21342 \\ 0.10519 \\ 0.06737 \\ 0.06231 \\ 3.00387 \times 10^{-9} \end{pmatrix} \quad \frac{1}{TI}\cdot\lambda^2 = \begin{pmatrix} 0.69701 \\ 0.1249 \\ 0.09363 \\ 0.05915 \\ 0.01437 \\ 0.00589 \\ 0.00504 \\ 0 \end{pmatrix}$$

< Proportion of Inertia for each singular value squared $\lambda^2 = \gamma U$

Calculating Correspondence analysis coordinates:

$Dsqrt_r := \text{diag}(\sqrt{R})$

$Dsqrt_c := \text{diag}(\sqrt{C})$

$F := D_r^{-1}\cdot(Dsqrt_r \cdot U)\cdot\Lambda$

$G := D_c^{-1}\cdot(Dsqrt_c \cdot V)\cdot\Lambda^T$

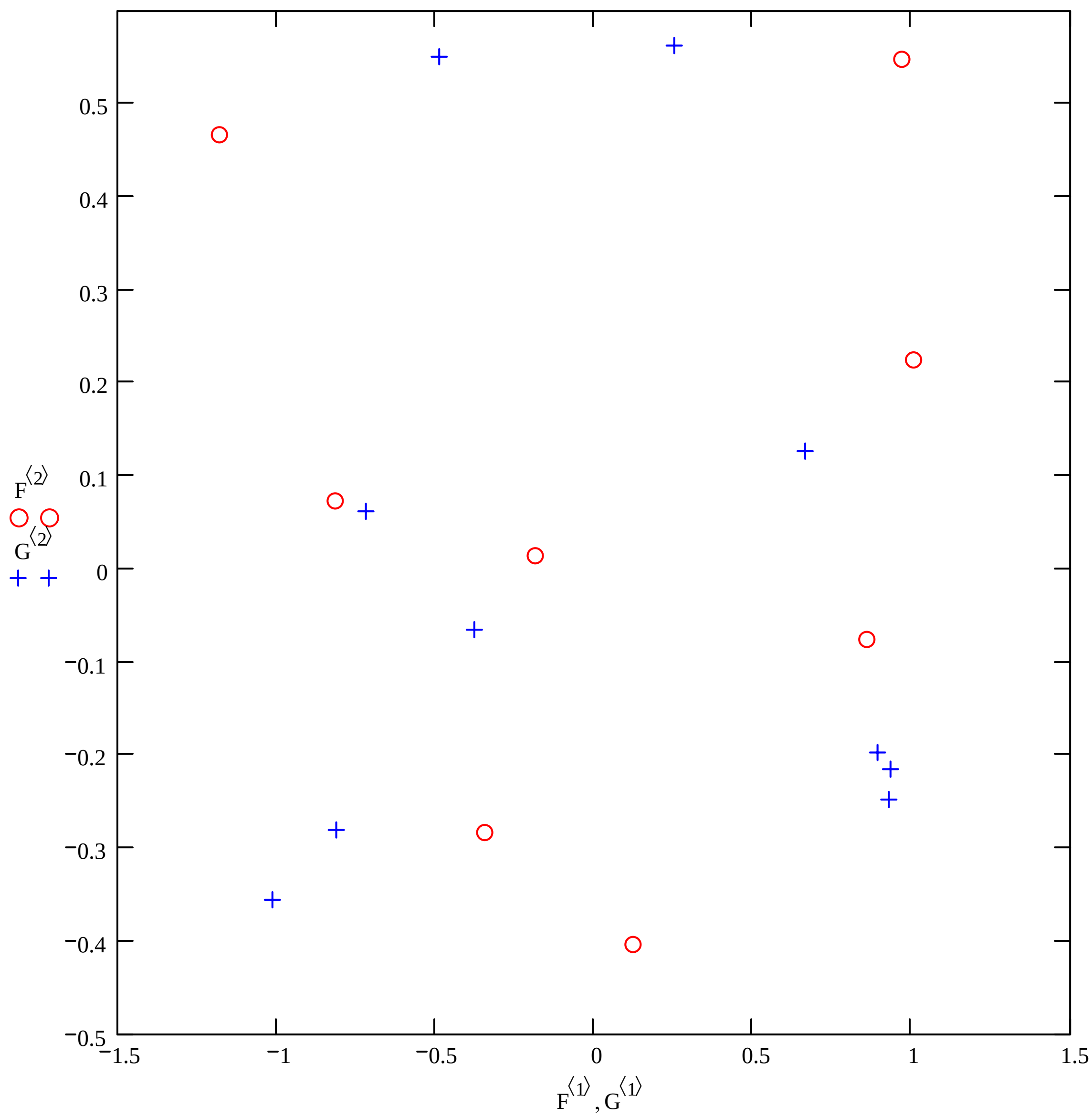
Matrix F represents vector positions for ROWS of M

$$F = \begin{pmatrix} -0.81462 & 0.07341 & 0.36564 & -0.27728 & -0.01129 & -0.06465 & -0.07224 & 3.00387 \times 10^{-9} & 0 \\ -1.17887 & 0.46701 & -0.33094 & -0.05964 & -0.02345 & 0.0696 & 0.04639 & 3.00387 \times 10^{-9} & 0 \\ -0.34368 & -0.28311 & -0.2577 & 0.18142 & 0.10841 & -0.05087 & -0.02644 & 3.00387 \times 10^{-9} & 0 \\ 0.12357 & -0.40337 & 0.11199 & -0.16405 & -0.01383 & 0.02599 & 0.09219 & 3.00387 \times 10^{-9} & 0 \\ -0.18426 & 0.01439 & 0.35194 & 0.46413 & -0.13431 & 0.03484 & -0.00587 & 3.00387 \times 10^{-9} & 0 \\ 0.85979 & -0.0755 & -0.11627 & -0.11671 & -0.00128 & 0.1164 & -0.10201 & 3.00387 \times 10^{-9} & 0 \\ 0.96982 & 0.54802 & 0.3162 & 0.07286 & 0.24137 & -0.00402 & 0.04931 & 3.00387 \times 10^{-9} & 0 \\ 1.00677 & 0.22501 & -0.24787 & -0.04543 & -0.14709 & -0.10153 & 0.01393 & 3.00387 \times 10^{-9} & 0 \end{pmatrix}$$

Note: typically the first two (or three) columns of F and G are plotted together... See plot below

Matrix G represents vector positions for COLUMNS of M

$$G = \begin{pmatrix} 0.93418 & -0.21495 & 0.05443 & -0.28279 & -0.02013 & -0.07244 & 0.08703 & 1.62132 \times 10^{-9} \\ 0.92897 & -0.24759 & 0.00933 & -0.30497 & -0.06745 & 0.04361 & -0.054 & 4.32915 \times 10^{-9} \\ 0.8932 & -0.19695 & -0.38208 & 0.42567 & 0.08598 & 0.02743 & 0.07158 & 3.64349 \times 10^{-9} \\ 0.66534 & 0.12697 & -0.04377 & 0.17062 & 0.0002 & 0.0168 & -0.12506 & 1.62253 \times 10^{-9} \\ 0.25331 & 0.5628 & 0.40681 & 0.10638 & -0.07658 & -0.03377 & 0.015 & 3.68181 \times 10^{-9} \\ -0.4871 & 0.55079 & -0.40846 & -0.18857 & -0.03384 & 0.08736 & 0.05468 & 2.51214 \times 10^{-9} \\ -0.3761 & -0.06501 & 0.37883 & 0.01721 & 0.17572 & 0.03652 & 0.02626 & 2.72685 \times 10^{-9} \\ -0.71758 & 0.0623 & -0.28429 & -0.10434 & 0.11047 & -0.13816 & -0.04874 & 3.43666 \times 10^{-9} \\ -1.01196 & -0.35529 & 0.03479 & -0.08177 & -0.02162 & 0.08358 & -0.02985 & 2.86993 \times 10^{-9} \\ -0.81128 & -0.28029 & 0.0289 & 0.22593 & -0.19673 & -0.04208 & 0.02453 & 2.69945 \times 10^{-9} \end{pmatrix}$$



Compare this plot with jw fig. 12.26 - It's the same!

F = ROWS of M = Trees
G = COLUMNS of M = Sites

CORRESPONDENCE ANALYSIS
jw715A.mcd

prepared by:
Wm Stein

ORIGIN ≡ 1

Verifying calculations in Example 12.18 jw p. 715 Contingency table with two variables: A (three levels) & B (two levels)

M := READPRN("\DATA\Ex12-18.txt")

$$M = \begin{pmatrix} 24 & 12 \\ 16 & 48 \\ 60 & 40 \end{pmatrix}$$

r := rows(M)   r = 3   c := cols(M)   c = 2

i := 1 .. r   j := 1 .. c

$l_{r_i} := 1 \quad l_{c_j} := 1$

Calculating Total number of observations in matrix M:

$$T := \sum_i \sum_j M_{i,j} \qquad T = 200$$

Matrix P of frequencies:

$$P := \frac{1}{T}M \qquad P = \begin{pmatrix} 0.12 & 0.06 \\ 0.08 & 0.24 \\ 0.3 & 0.2 \end{pmatrix}$$

Calculating Row (R) and Column (C) vectors of sums of matrix P:

$$R := P \cdot l_c \qquad R = \begin{pmatrix} 0.18 \\ 0.32 \\ 0.5 \end{pmatrix} \qquad D_r := \text{diag}(R) \qquad D_r = \begin{pmatrix} 0.18 & 0 & 0 \\ 0 & 0.32 & 0 \\ 0 & 0 & 0.5 \end{pmatrix}$$

$$C := P^T \cdot l_r \qquad C = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix} \qquad D_c := \text{diag}(C) \qquad D_c = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix}$$

Calculting diagonal inverse squaret matrices $D_r^{-1/2}$ & $D_c^{-1/2}$:

$$\text{Dinvsqrt}_r := \text{diag}\left(\frac{1}{\sqrt{R}}\right) \qquad \text{Dinvsqrt}_r = \begin{pmatrix} 2.35702 & 0 & 0 \\ 0 & 1.76777 & 0 \\ 0 & 0 & 1.41421 \end{pmatrix} \qquad \frac{\sqrt{2}}{.6} = 2.35702$$

$$\text{Dinvsqrt}_c := \text{diag}\left(\frac{1}{\sqrt{C}}\right) \qquad \text{Dinvsqrt}_c = \begin{pmatrix} 1.41421 & 0 \\ 0 & 1.41421 \end{pmatrix} \qquad \sqrt{2} = 1.41421$$

Verifying matrices $P-RC^T$ and A:

$$P - R \cdot C^T = \begin{pmatrix} 0.03 & -0.03 \\ -0.08 & 0.08 \\ 0.05 & -0.05 \end{pmatrix} \qquad A := \text{Dinvsqrt}_r \cdot \left(P - R \cdot C^T\right) \cdot \text{Dinvsqrt}_c \qquad A = \begin{pmatrix} 0.1 & -0.1 \\ -0.2 & 0.2 \\ 0.1 & -0.1 \end{pmatrix}$$

$$A \cdot A^T = \begin{pmatrix} 0.02 & -0.04 & 0.02 \\ -0.04 & 0.08 & -0.04 \\ 0.02 & -0.04 & 0.02 \end{pmatrix} \qquad A^T \cdot A = \begin{pmatrix} 0.06 & -0.06 \\ -0.06 & 0.06 \end{pmatrix}$$

2005 jw715A

Singular Value Decomposition of a Matrix (from Matrix Algebra Toolbox jw 084.mcd):

$A := A^T$

< Note: I transposed matrix A here to conform to the example in the toolbox. This allows all calculations to follow automatically but changes the definition of U & V compared with Example 12.18 ...

$m := \text{rows}(A) \qquad k := \text{cols}(A)$

$\gamma_U := \text{reverse}\left(\text{sort}\left(\text{eigenvals}\left(A \cdot A^T\right)\right)\right) \qquad \gamma_U = \begin{pmatrix} 0.12 \\ 0 \end{pmatrix}$

$\gamma_V := \text{reverse}\left(\text{sort}\left(\text{eigenvals}\left(A^T \cdot A\right)\right)\right) \qquad \gamma_V = \begin{pmatrix} 0.12 \\ 0 \\ 0 \end{pmatrix}$

Note: as in the text, Eigenvalues for the different products of A are designated γ (gamma) γ_U & γ_V

$i := 1 .. \text{rows}\left(A \cdot A^T\right) \qquad j := 1 .. \text{rows}\left(A^T \cdot A\right)$

Associated Eigenvectors are extracted:

$U^{\langle i \rangle} := \text{eigenvec}\left(A \cdot A^T, \gamma_{U_i}\right) \qquad U = \begin{pmatrix} -0.70711 & 0.70711 \\ 0.70711 & 0.70711 \end{pmatrix}$

$V^{\langle j \rangle} := \text{eigenvec}\left(A^T \cdot A, \gamma_{V_j}\right) \qquad V = \begin{pmatrix} 0.40825 & -0.05241 & -0.05241 \\ -0.8165 & 0.42551 & 0.42551 \\ 0.40825 & 0.90343 & 0.90343 \end{pmatrix}$

$\lambda := \sqrt{\gamma_U} \qquad \lambda = \begin{pmatrix} 0.34641 \\ 0 \end{pmatrix}$

Singular values λ (lambda) are calculated.

$i := 1 .. k - m \qquad k - m = 1 \qquad Z_{m,i} := 0 \qquad Z = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

(k-m) Extra columns of zeros are needed for singular value matrix Λ

$\Lambda := \text{augment}(\text{diag}(\lambda), Z) \qquad \Lambda = \begin{pmatrix} 0.34641 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

$U \cdot \Lambda \cdot V^T = \begin{pmatrix} -0.1 & 0.2 & -0.1 \\ 0.1 & -0.2 & 0.1 \end{pmatrix} \qquad A = \begin{pmatrix} 0.1 & -0.2 & 0.1 \\ -0.1 & 0.2 & -0.1 \end{pmatrix}$

Λ is the matrix of singular values and zeros

Singular Value Decomposition confirmed!

Comparing values of the singular value decomposition functions svd & sdvs in Mathcad:

$\text{svd}\left(A^T \cdot A\right) = \begin{pmatrix} 0.40825 & -0.23961 & -0.88086 \\ -0.8165 & -0.52738 & -0.23496 \\ 0.40825 & -0.81514 & 0.41094 \\ 0.40825 & 0 & 0.91287 \\ -0.8165 & 0.44721 & 0.36515 \\ 0.40825 & 0.89443 & -0.18257 \end{pmatrix} \qquad \text{svd}\left(A \cdot A^T\right) = \begin{pmatrix} -0.70711 & -0.70711 \\ 0.70711 & -0.70711 \\ -0.70711 & 0.70711 \\ 0.70711 & 0.70711 \end{pmatrix}$

$\text{svds}\left(A^T \cdot A\right) = \begin{pmatrix} 0.12 \\ 0 \\ 0 \end{pmatrix} \qquad \text{svds}\left(A \cdot A^T\right) = \begin{pmatrix} 0.12 \\ 0 \end{pmatrix}$

$U = \begin{pmatrix} -0.70711 & 0.70711 \\ 0.70711 & 0.70711 \end{pmatrix} \qquad \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 0.70711 \\ 0.70711 \end{pmatrix}$

$V = \begin{pmatrix} 0.40825 & -0.05241 & -0.05241 \\ -0.8165 & 0.42551 & 0.42551 \\ 0.40825 & 0.90343 & 0.90343 \end{pmatrix} \qquad \begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{pmatrix} = \begin{pmatrix} 0.40825 \\ -0.8165 \\ 0.40825 \end{pmatrix}$

< Confirming values given for U & V (first column) jw p. 716

2005 jw715A

Initeria:

$i := 1 .. \text{rows}(\lambda)$

$TI := \sum_i \gamma U_i \qquad TI = 0.12$

$\gamma U = \begin{pmatrix} 0.12 \\ 0 \end{pmatrix} \qquad \lambda = \begin{pmatrix} 0.34641 \\ 0 \end{pmatrix} \qquad \frac{1}{TI}\cdot\lambda^2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

< Proportion of Inertia for each singular value squared $\lambda^2 = \gamma U$

Calculating Correspondence analysis coordinates:

$Dsqrt_r := \text{diag}(\sqrt{R})$

$Dsqrt_c := \text{diag}(\sqrt{C})$

$Dsqrt_r = \begin{pmatrix} 0.42426 & 0 & 0 \\ 0 & 0.56569 & 0 \\ 0 & 0 & 0.70711 \end{pmatrix}$

$F := D_r^{-1}\cdot(Dsqrt_r\cdot V)\cdot\Lambda^T$

$G := D_c^{-1}\cdot(Dsqrt_c\cdot U)\cdot\Lambda$

$F = \begin{pmatrix} 0.33333 & 0 \\ -0.5 & 0 \\ 0.2 & 0 \end{pmatrix} \qquad Dsqrt_c = \begin{pmatrix} 0.70711 & 0 \\ 0 & 0.70711 \end{pmatrix} \qquad U = \begin{pmatrix} -0.70711 & 0.70711 \\ 0.70711 & 0.70711 \end{pmatrix}$

Matrix F represents vector positions for ROWS of M

Matrix G represents vector positions for COLUMNS of M

$D_r^{-1} = \begin{pmatrix} 5.55556 & 0 & 0 \\ 0 & 3.125 & 0 \\ 0 & 0 & 2 \end{pmatrix} \qquad V = \begin{pmatrix} 0.40825 & -0.05241 & -0.05241 \\ -0.8165 & 0.42551 & 0.42551 \\ 0.40825 & 0.90343 & 0.90343 \end{pmatrix}$

$G = \begin{pmatrix} -0.34641 & 0 & 0 \\ 0.34641 & 0 & 0 \end{pmatrix}$

$D_c^{-1} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \qquad \Lambda = \begin{pmatrix} 0.34641 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

Comparing Results jw p. 717

Simplified Calculation of F & G:

$Dinvsqrt_r\cdot V\cdot\Lambda^T = \begin{pmatrix} 0.33333 & 0 \\ -0.5 & 0 \\ 0.2 & 0 \end{pmatrix}$

$Dinvsqrt_c\cdot U\cdot\Lambda = \begin{pmatrix} -0.34641 & 0 & 0 \\ 0.34641 & 0 & 0 \end{pmatrix}$

jw Report on p. 717:

$\sqrt{.12}\cdot\begin{pmatrix} \frac{.3}{\sqrt{3}} \\ \frac{-.8}{\sqrt{3}} \\ \frac{.5}{\sqrt{3}} \end{pmatrix} = \begin{pmatrix} 0.06 \\ -0.16 \\ 0.1 \end{pmatrix}$

$\sqrt{.12}\cdot\begin{pmatrix} \frac{1}{2} \\ \frac{-1}{2} \end{pmatrix} = \begin{pmatrix} 0.17321 \\ -0.17321 \end{pmatrix}$

I can calculate their report by replacing $D^{-1/2}$ with $D^{1/2}$ - but is this an error on their part?

$Dsqrt_r\cdot V\cdot\Lambda^T = \begin{pmatrix} 0.06 & 0 \\ -0.16 & 0 \\ 0.1 & 0 \end{pmatrix}$

$Dsqrt_c\cdot U\cdot\Lambda = \begin{pmatrix} -0.17321 & 0 & 0 \\ 0.17321 & 0 & 0 \end{pmatrix}$

Note: typically the first two (or three) columns of F and G are plotted together...

However this case is degenerate: only a single axis of variation for F or G can be plotted (one dimensional plot).

Formulas for Matrices F & G on jw p. 714 clearly use $D^{-1/2}$