ORIGIN ≡ 1

Canonical Correlation Analysis

Canonical Correlation Analysis (CCA) also known as Canonical Variates Analysis (CVA) or just Canonical Analysis (CA), should not be confused with Canonical Correspondence Analysis (also commonly abbreviated CCA) which is an entirely different method. Canonical Correlation Analysis is also sometimes termed "Redundancy Analysis" (when redundancy coefficients are calculated) but the basic method is the same with the latter name is only adding more confusion.

Canonical Correlation Analysis covers the multivariate situation where a data are naturally divided into two sets of variables (X versus Y) covering the same objects. Classic situations include variables collected on individuals before (X) versus after (Y) an event, aptitude measures (X) versus test results (Y), subject behaviors (X) versus environmental conditions (Y), etc. In general, the basic intent is ordination, that is to identify a simple way to judge the relationship between the two sets of variables without specifying independent versus depent relationships (as typically encountered in regression). As with Principal Components Analysis (PCA), CCA may be used primarily as a graphical method to look for patterns in high-dimension data in a simple way that employs a few well-chosen linear combinations with useful summary properties. In CCA, these linear combinations are called "canonical variates" (CV). Similar to PCA, CV's are calculated by decomposition of a covariance or correlation matrix resulting in eigenvector and eigenvalue pairs. Eigenvalues (λ) provide "canonical correlations", a measure of correlation between the variable sets, with the associated eigenvectors (E) leading to coefficients or "loadings" indicating the relative contributions of the original variables to the CV in each set. Data here are derived from a psychological study referenced online by IDRE at UCLA at http://www.ats.ucla.edu/stat/r/dae/canonical.htm. Discussion and calculations are mostly based on R.A. Johnson & D.W Wichern (JW) *Applied Multivariate Statistical Analysis 5th Edition* 2002 with further commentary from A.C. Rencher *Methods of Multivariate Analysis* 1995, excellent explanation and code in the R package {yacca} by C.T. Butts at: cran.r-project.org/web/packages/yacca/yacca.pdf, and clarification by J. Steiger at: http://www.statpower.net/Content/312/Lecture%20Slides/CanonicalCorrelation.pdf.

Data setup:

M := READPRN("c:/DATA/Multivariate/UCLAidreCCAEx.txt")

Data set with 600 cases (rows) with columns 1-3 representing one set of variables and columns 4-7 the other set:

col 1 = control
col 2 = concept
col 3 = motivation

Psychological variables (X)

col 4 = read
col 5 = write
col 6 = math
col 7 = science
col 8 = sex

Academic variables (Y)

$n := \text{length}\left(M^{\langle 1\rangle}\right)$ $n = 600$

X := submatrix(M, 1, 600, 1, 3)

Y := submatrix(M, 1, 600, 4, 8)

i := 1 .. n **< total number of objects**

p := 1 .. 8 **< total number of variables**

M =

	1	2	3	4	5	6	7	8
1	-0.84	-0.24	1	54.8	64.5	44.5	52.6	1
2	-0.38	-0.47	0.67	62.7	43.7	44.7	52.6	1
3	0.89	0.59	0.67	60.6	56.7	70.5	58	0
4	0.71	0.28	0.67	62.7	56.7	54.7	58	0
5	-0.64	0.03	1	41.6	46.3	38.4	36.3	1
6	1.11	0.9	0.33	62.7	64.5	61.4	58	1
7	0.06	0.03	0.67	41.6	39.1	56.3	45	0
8	-0.91	-0.59	0.67	44.2	39.1	46.3	36.3	0
9	0.45	0.03	1	62.7	51.5	54.4	49.8	1
10	0	0.03	0.67	62.7	64.5	38.3	55.8	1
11	0.24	-0.43	0.33	70.7	43.7	58.8	66.1	0
12	-1.09	-0.26	0.33	44.2	41.1	45.1	47.1	0
13	0.46	0.03	0.67	57.4	59.3	53.9	49.8	1
14	0.68	0.06	0.67	49.5	51.5	41.2	41.7	1
15	-0.14	-1.05	1	70.7	65.1	66.4	63.4	1
16	0.1	-0.16	0.33	49.5	59.3	51	47.1	0

Total Covariance Matrix:

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

$l_{vec_i} := 1$

$$S := \frac{1}{n-1} \cdot M^T \cdot \left(I - \frac{1}{n} \cdot l_{vec} \cdot l_{vec}^T\right) \cdot M$$

$$S = \begin{pmatrix} 0.449 & 0.081 & 0.056 & 2.53 & 2.34 & 2.128 & 2.112 & 0.038 \\ 0.081 & 0.498 & 0.07 & 0.432 & 0.133 & 0.356 & 0.478 & -0.044 \\ 0.056 & 0.07 & 0.117 & 0.729 & 0.848 & 0.629 & 0.385 & 0.017 \\ 2.53 & 0.432 & 0.729 & 102.07 & 61.769 & 64.611 & 67.73 & -0.21 \\ 2.34 & 0.133 & 0.848 & 61.769 & 94.604 & 57.935 & 53.732 & 1.184 \\ 2.128 & 0.356 & 0.629 & 64.611 & 57.935 & 88.637 & 59.354 & -0.226 \\ 2.112 & 0.478 & 0.385 & 67.73 & 53.732 & 59.354 & 94.21 & -0.668 \\ 0.038 & -0.044 & 0.017 & -0.21 & 1.184 & -0.226 & -0.668 & 0.248 \end{pmatrix}$$

Total Correlation Matrix:

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

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

$$R = \begin{pmatrix} 1 & 0.171 & 0.245 & 0.374 & 0.359 & 0.337 & 0.325 & 0.113 \\ 0.171 & 1 & 0.289 & 0.061 & 0.019 & 0.054 & 0.07 & -0.126 \\ 0.245 & 0.289 & 1 & 0.211 & 0.254 & 0.195 & 0.116 & 0.098 \\ 0.374 & 0.061 & 0.211 & 1 & 0.629 & 0.679 & 0.691 & -0.042 \\ 0.359 & 0.019 & 0.254 & 0.629 & 1 & 0.633 & 0.569 & 0.244 \\ 0.337 & 0.054 & 0.195 & 0.679 & 0.633 & 1 & 0.65 & -0.048 \\ 0.325 & 0.07 & 0.116 & 0.691 & 0.569 & 0.65 & 1 & -0.138 \\ 0.113 & -0.126 & 0.098 & -0.042 & 0.244 & -0.048 & -0.138 & 1 \end{pmatrix}$$

Partitioning Covariance and Correlation Matrices:

$Sxx := \text{submatrix}(S, 1, 3, 1, 3)$

$Syy := \text{submatrix}(S, 4, 8, 4, 8)$

$Syx := \text{submatrix}(S, 4, 8, 1, 3)$

$Sxy := \text{submatrix}(S, 1, 3, 4, 8)$

$Rxx := \text{submatrix}(R, 1, 3, 1, 3)$

$Ryy := \text{submatrix}(R, 4, 8, 4, 8)$

$Ryx := \text{submatrix}(R, 4, 8, 1, 3)$

$Rxy := \text{submatrix}(R, 1, 3, 4, 8)$

$$Sxx = \begin{pmatrix} 0.4493 & 0.081 & 0.0563 \\ 0.081 & 0.4977 & 0.0698 \\ 0.0563 & 0.0698 & 0.1175 \end{pmatrix}$$

$$Rxx = \begin{pmatrix} 1 & 0.1712 & 0.2451 \\ 0.1712 & 1 & 0.2886 \\ 0.2451 & 0.2886 & 1 \end{pmatrix}$$

$$Sxy = \begin{pmatrix} 2.5297 & 2.3397 & 2.1283 & 2.112 & 0.0379 \\ 0.4323 & 0.1335 & 0.356 & 0.4782 & -0.0443 \\ 0.7293 & 0.8475 & 0.6293 & 0.3848 & 0.0168 \end{pmatrix}$$

$$Rxy = \begin{pmatrix} 0.3736 & 0.3589 & 0.3373 & 0.3246 & 0.1134 \\ 0.0607 & 0.0194 & 0.0536 & 0.0698 & -0.126 \\ 0.2106 & 0.2542 & 0.195 & 0.1157 & 0.0981 \end{pmatrix}$$

$$Syy = \begin{pmatrix} 102.0703 & 61.7692 & 64.6106 & 67.7303 & -0.2102 \\ 61.7692 & 94.6039 & 57.9345 & 53.7316 & 1.1844 \\ 64.6106 & 57.9345 & 88.6373 & 59.3544 & -0.2262 \\ 67.7303 & 53.7316 & 59.3544 & 94.2099 & -0.6685 \\ -0.2102 & 1.1844 & -0.2262 & -0.6685 & 0.2484 \end{pmatrix}$$

$$Ryy = \begin{pmatrix} 1 & 0.6286 & 0.6793 & 0.6907 & -0.0417 \\ 0.6286 & 1 & 0.6327 & 0.5691 & 0.2443 \\ 0.6793 & 0.6327 & 1 & 0.6495 & -0.0482 \\ 0.6907 & 0.5691 & 0.6495 & 1 & -0.1382 \\ -0.0417 & 0.2443 & -0.0482 & -0.1382 & 1 \end{pmatrix}$$

$$Syx = \begin{pmatrix} 2.5297 & 0.4323 & 0.7293 \\ 2.3397 & 0.1335 & 0.8475 \\ 2.1283 & 0.356 & 0.6293 \\ 2.112 & 0.4782 & 0.3848 \\ 0.0379 & -0.0443 & 0.0168 \end{pmatrix}$$

$$Ryx = \begin{pmatrix} 0.3736 & 0.0607 & 0.2106 \\ 0.3589 & 0.0194 & 0.2542 \\ 0.3373 & 0.0536 & 0.195 \\ 0.3246 & 0.0698 & 0.1157 \\ 0.1134 & -0.126 & 0.0981 \end{pmatrix}$$

Eigenvalues & Eigenvectors of Partitioned CCA Matrix:

Using partitioned covariance matrix S:

For variable set X:

$$Sxx^{-1}\cdot Sxy\cdot Syy^{-1}\cdot Syx=\begin{pmatrix}0.169 & 0.0091 & 0.0486\\ -0.0382 & 0.02 & -0.0164\\ 0.1474 & -0.0137 & 0.0652\end{pmatrix}$$

$$SSx := Sxx^{-1}\cdot Sxy\cdot Syy^{-1}\cdot Syx$$

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

$$\lambda_{Sx} := \text{reverse}(\text{sort}(\text{eigenvals}(SSx)))$$

$$E_{Sx}^{\langle i\rangle} := \text{eigenvec}\left(SSx,\lambda_{Sx_i}\right)$$

$$\lambda_{Sx}=\begin{pmatrix}0.2154\\ 0.0281\\ 0.0108\end{pmatrix}\qquad E_{Sx}=\begin{pmatrix}0.6913 & 0.2557 & -0.2924\\ -0.1937 & 0.4887 & 0.3653\\ 0.6961 & -0.8341 & 0.8838\end{pmatrix}$$

Using partitioned correlation matrix R:

$$Rxx^{-1}\cdot Rxy\cdot Ryy^{-1}\cdot Ryx=\begin{pmatrix}0.169 & 0.0087 & 0.095\\ -0.0402 & 0.02 & -0.0338\\ 0.0753 & -0.0066 & 0.0652\end{pmatrix}$$

$$RRx := Rxx^{-1}\cdot Rxy\cdot Ryy^{-1}\cdot Ryx$$

$$\lambda_{Rx} := \text{reverse}(\text{sort}(\text{eigenvals}(RRx)))$$

$$E_{Rx}^{\langle i\rangle} := \text{eigenvec}\left(RRx,\lambda_{Rx_i}\right)$$

$$\lambda_{Rx}=\begin{pmatrix}0.2154\\ 0.0281\\ 0.0108\end{pmatrix}\qquad E_{Rx}=\begin{pmatrix}0.86 & -0.3574 & -0.442\\ -0.2537 & -0.7189 & 0.5813\\ 0.4428 & 0.5961 & 0.6832\end{pmatrix}$$

For variable set Y:

$$Syy^{-1}\cdot Syx\cdot Sxx^{-1}\cdot Sxy=\begin{pmatrix}0.0824 & 0.0824 & 0.0698 & 0.0618 & 0.0016\\ 0.0627 & 0.0691 & 0.0536 & 0.0394 & 0.0016\\ 0.0425 & 0.0433 & 0.0361 & 0.0312 & 0.0009\\ 0.0162 & 0.0048 & 0.0128 & 0.0253 & -0.0004\\ 1.0191 & 1.1319 & 0.8668 & 0.7135 & 0.0414\end{pmatrix}$$

$$SSy := Syy^{-1}\cdot Syx\cdot Sxx^{-1}\cdot Sxy$$

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

$$\lambda_{Sy} := \text{reverse}(\text{sort}(\text{eigenvals}(SSy)))$$

$$E_{Sy}^{\langle i\rangle} := \text{eigenvec}\left(SSy,\lambda_{Sy_i}\right)$$

$$\lambda_{Sy}=\begin{pmatrix}0.2154\\ 0.0281\\ 0.0108\\ 0\\ 0\end{pmatrix}\qquad E_{Sy}=\begin{pmatrix}0.0703 & 0.0045 & 0.0119 & -0.6883 & -0.6883\\ 0.0565 & -0.0386 & 0.0507 & 0.0331 & 0.0331\\ 0.0369 & -0.0039 & 0.0052 & 0.7119 & 0.7119\\ 0.0079 & 0.0782 & -0.061 & 0.0725 & 0.0724\\ 0.9952 & -0.9962 & -0.9968 & -0.1148 & -0.1147\end{pmatrix}$$

$$Ryy^{-1}\cdot Ryx\cdot Rxx^{-1}\cdot Rxy=\begin{pmatrix}0.0824 & 0.0856 & 0.0749 & 0.0643 & 0.033\\ 0.0603 & 0.0691 & 0.0554 & 0.0395 & 0.0308\\ 0.0396 & 0.0419 & 0.0361 & 0.0303 & 0.0176\\ 0.0156 & 0.0048 & 0.0132 & 0.0253 & -0.0074\\ 0.0503 & 0.058 & 0.0459 & 0.0366 & 0.0414\end{pmatrix}$$

$$RRy := Ryy^{-1}\cdot Ryx\cdot Rxx^{-1}\cdot Rxy$$

$$\lambda_{Ry} := \text{reverse}(\text{sort}(\text{eigenvals}(RRy)))$$

$$E_{Ry}^{\langle i\rangle} := \text{eigenvec}\left(RRy,\lambda_{Ry_i}\right)$$

$$\lambda_{Ry}=\begin{pmatrix}0.2154\\ 0.0281\\ 0.0108\\ 0\\ 0\end{pmatrix}$$

$$E_{Ry}=\begin{pmatrix}0.6539 & -0.0463 & -0.1296 & -0.7539 & -0.7539\\ 0.5062 & 0.3821 & -0.5327 & 0.09 & 0.09\\ 0.3198 & 0.0372 & -0.0531 & 0.6417 & 0.6417\\ 0.0708 & -0.7718 & 0.6394 & 0.1073 & 0.1073\\ 0.457 & 0.5047 & 0.5365 & -0.0167 & -0.0166\end{pmatrix}$$

For each of variable sets X & Y, eigenvalues are identical regardless of whether the covariance or correlation matrix is decomposed. This is a unique result for CCA not shared with PCA.

Canonical Correlations:

$$\lambda_{Sx} = \begin{pmatrix} 0.2153759 \\ 0.0280593 \\ 0.0108142 \end{pmatrix} \qquad \sqrt{\lambda_{Sx}} = \begin{pmatrix} 0.4640861 \\ 0.1675092 \\ 0.1039911 \end{pmatrix}$$

< Eigenvalues give shared variance on each CV.

$$\lambda_{Sy} = \begin{pmatrix} 0.2153759 \\ 0.0280593 \\ 0.0108142 \\ 0 \\ 0 \end{pmatrix} \qquad \sqrt{\lambda_{Sy}} = \begin{pmatrix} 0.4640861 \\ 0.1675092 \\ 0.1039911 \\ 0 \\ 0i \end{pmatrix}$$

< square root of eigenvalues give cannonical correlations. When X & Y sets have different numbers of variables, the number of canonical correlations and canonical variates corresponds only to the smaller set.

Raw Canonical Variate Coefficients:

For variable set X:

$$E_{Sx} = \begin{pmatrix} 0.6913 & 0.2557 & -0.2924 \\ -0.1937 & 0.4887 & 0.3653 \\ 0.6961 & -0.8341 & 0.8838 \end{pmatrix}$$

Note: from this point on, calculations are based on the partitioned covariance matrix S only. See the section below covering "Standardized Data" with calculations using the partitioned correlation matrix R.

For variable set Y:

$$E_{Sy} = \begin{pmatrix} 0.0703 & 0.0045 & 0.0119 & -0.6883 & -0.6883 \\ 0.0565 & -0.0386 & 0.0507 & 0.0331 & 0.0331 \\ 0.0369 & -0.0039 & 0.0052 & 0.7119 & 0.7119 \\ 0.0079 & 0.0782 & -0.061 & 0.0725 & 0.0724 \\ 0.9952 & -0.9962 & -0.9968 & -0.1148 & -0.1147 \end{pmatrix}$$

< Eigenvectors (columns) are the raw canonical variate coefficients. However, eigenvectors never represent unique solutions (lengths). So the vectors are typically scaled to unit variance.

Scaled Canonical Variate Coefficients:

For variable X:

$N_x := X \cdot E_{Sx}$ **< projects each X object onto CV axis**

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

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

$l_{vec_i} := 1$

$$S_{Nx} := \frac{1}{n-1} \cdot N_x^T \cdot \left(I - \frac{1}{n} \cdot l_{vec} \cdot l_{vec}^T\right) \cdot N_x \qquad S_{Nx} = \begin{pmatrix} 0.304 & 0 & 0 \\ 0 & 0.1693 & 0 \\ 0 & 0 & 0.1952 \end{pmatrix}$$

< covariance of all objects in CV coordinates

$$D_{Nx_{i,i}} := \sqrt{S_{Nx_{i,i}}} \qquad D_{Nx} = \begin{pmatrix} 0.5514 & 0 & 0 \\ 0 & 0.4115 & 0 \\ 0 & 0 & 0.4418 \end{pmatrix}$$

< standard deviation

$$E_{Sx} \cdot D_{Nx}^{-1} = \begin{pmatrix} 1.2538339 & 0.6214776 & -0.6616896 \\ -0.3513499 & 1.1876866 & 0.826721 \\ 1.2624204 & -2.0272641 & 2.0002283 \end{pmatrix}$$

< X scaled coefficients - scaled by recriprocal of the calculated standard deviation. Scaled coefficients will be used to calculate scores for variables in the last section below.

For variable Y:

$N_y := Y \cdot E_{Sy}$ **< projects each Y object onto CV axis**

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

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

$l_{vec_i} := 1$

$$S_{Ny} := \frac{1}{n-1} \cdot N_y^T \cdot \left(I - \frac{1}{n} \cdot l_{vec} \cdot l_{vec}^T\right) \cdot N_y$$

$$S_{Ny} = \begin{pmatrix} 2.4788 & 0 & -0 & -0 & -0 \\ 0 & 0.8435 & -0 & -0 & -0 \\ -0 & -0 & 0.3085 & 0 & 0 \\ -0 & -0 & 0 & 30.1051 & 30.1053 \\ 0 & -0 & 0 & 30.1053 & 30.1056 \end{pmatrix}$$

^ variance of all objcects in CV coordinates

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

$$D_{Ny} = \begin{pmatrix} 1.5744 & 0 & 0 & 0 & 0 \\ 0 & 0.9184 & 0 & 0 & 0 \\ 0 & 0 & 0.5554 & 0 & 0 \\ 0 & 0 & 0 & 5.4868 & 0 \\ 0 & 0 & 0 & 0 & 5.4869 \end{pmatrix}$$

^ standard deviation

$$E_{Sy} \cdot D_{Ny}^{-1} = \begin{pmatrix} 0.0446206 & 0.00491 & 0.0213806 & -0.1254435 & -0.1254408 \\ 0.0358771 & -0.0420715 & 0.0913073 & 0.00603 & 0.0060261 \\ 0.0234172 & -0.0042295 & 0.0093982 & 0.1297393 & 0.1297425 \\ 0.0050252 & 0.0851622 & -0.109835 & 0.0132066 & 0.0132042 \\ 0.6321192 & -1.0846423 & -1.794647 & -0.0209249 & -0.0209105 \end{pmatrix}$$

< Y scaled coefficients - scaled by recriprocal of the calculated standard deviation

Structural Correlations (Loadings):

Structural Correlations are the Pearson product correlation between original variables in sets X & Y with each CV.

For variable X:

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

$j := 1 .. \text{rows}(E_{Sx})$

$X_S^{\langle i \rangle} := X \cdot E_{Sx}^{\langle i \rangle}$

$$E_{Sx} = \begin{pmatrix} 0.6913 & 0.2557 & -0.2924 \\ -0.1937 & 0.4887 & 0.3653 \\ 0.6961 & -0.8341 & 0.8838 \end{pmatrix}$$

$r_{X_{i,j}} := \text{corr}\left(X^{\langle i \rangle}, X_S^{\langle j \rangle}\right)$

$$r_X = \begin{pmatrix} 0.9040463 & 0.3896883 & -0.1756227 \\ 0.0208433 & 0.7087386 & 0.7051632 \\ 0.5671511 & -0.3508882 & 0.7451289 \end{pmatrix}$$

^ Pearson product correlations = "Structural correlations"

$$\frac{\text{cvar}\left(X^{\langle 1 \rangle}, X_S^{\langle 2 \rangle}\right)}{\text{stdev}\left(X^{\langle 1 \rangle}\right) \cdot \text{stdev}\left(X_S^{\langle 2 \rangle}\right)} = 0.3896883$$

< example calculation

$X_S =$

	1	2	3
1	0.1618	-1.1662	1.0417
2	0.2947	-0.8857	0.5316
3	0.9674	-0.043	0.5475
4	0.903	-0.2405	0.4868
5	0.2478	-0.9831	1.0819
6	0.8227	0.4484	0.2959
7	0.502	-0.5289	0.5856
8	-0.0485	-1.0799	0.6427
9	1.0014	-0.7044	0.7632
10	0.4606	-0.5442	0.6031
11	0.4789	-0.424	0.0644
12	-0.4735	-0.6811	0.5154
13	0.7786	-0.4266	0.4686
14	0.9249	-0.3557	0.4153
15	0.8027	-1.3831	0.5412
16	0.3298	-0.3279	0.204

For variable Y:

$i := 1 .. rows(E_{Sy})$

$j := 1 .. rows(E_{Sy})$

$Y_S^{\langle i \rangle} := Y \cdot E_{Sy}^{\langle i \rangle}$

$$E_{Sy} = \begin{pmatrix} 0.0703 & 0.0045 & 0.0119 & -0.6883 & -0.6883 \\ 0.0565 & -0.0386 & 0.0507 & 0.0331 & 0.0331 \\ 0.0369 & -0.0039 & 0.0052 & 0.7119 & 0.7119 \\ 0.0079 & 0.0782 & -0.061 & 0.0725 & 0.0724 \\ 0.9952 & -0.9962 & -0.9968 & -0.1148 & -0.1147 \end{pmatrix}$$

$rY_{i,j} := corr\left(Y^{\langle i \rangle}, Y_S^{\langle j \rangle}\right)$

$$rY = \begin{pmatrix} 0.840448 & 0.3588254 & 0.1353635 & -0.3118055 & -0.311798 \\ 0.8765429 & -0.0648367 & 0.2545608 & 0.1051897 & 0.1051765 \\ 0.7639483 & 0.2979488 & 0.1477611 & 0.4814476 & 0.4814569 \\ 0.6584139 & 0.6767976 & -0.2303551 & 0.0810262 & 0.0810187 \\ 0.3641127 & -0.7549281 & -0.5434036 & -0.0198058 & -0.0198072 \end{pmatrix}$$

^ Pearson product correlations = "Structural correlations"

$$\frac{cvar\left(Y^{\langle 1 \rangle}, Y_S^{\langle 2 \rangle}\right)}{stdev\left(Y^{\langle 1 \rangle}\right) \cdot stdev\left(Y_S^{\langle 2 \rangle}\right)} = 0.3588254$$

< example calculation

Redundancy Coefficients:

Redundancy may be used as a measure of the amount of variance in each set (X or Y) that may be "explained" by the other set. The concept of "explained" is always problematic in multivariate statistics and should be treated with caution as causal relations are not actually evaluated.
Calculations shown here were recovered by reading the R code for cca() in package {yacca}.

For variable X:

see in R:

```
>cca
>?cca
```

$i := 1 .. rows(r_X)$

$j := 1 .. rows(r_X)$

$$r_X = \begin{pmatrix} 0.904 & 0.3897 & -0.1756 \\ 0.0208 & 0.7087 & 0.7052 \\ 0.5672 & -0.3509 & 0.7451 \end{pmatrix}$$

< X structural correlations

$R_{X_{i,j}} := \left(r_{X_{i,j}}\right)^2$

$$R_X = \begin{pmatrix} 0.8173 & 0.1519 & 0.0308 \\ 0.0004 & 0.5023 & 0.4973 \\ 0.3217 & 0.1231 & 0.5552 \end{pmatrix}$$

< squarred structural correlations = fraction variance "explained" by each variable

$Xcanvad_i := mean\left(R_X^{\langle i \rangle}\right)$

$$Xcanvad = \begin{pmatrix} 0.3797982 \\ 0.2590966 \\ 0.3611052 \end{pmatrix}$$

< canonical variable adequacies = fraction of total X variance "explained" by each CV

$xvrd_i := Xcanvad_i \cdot \lambda_{Sx_i}$

$$xvrd = \begin{pmatrix} 0.081799367 \\ 0.007270075 \\ 0.003905049 \end{pmatrix}$$

< canonical redundancies = total fraction of X variance "explained" by the Y variables, through EACH CV

$xrd := \sum_i xvrd_i$

$xrd = 0.09297449$

< Total canonical redundancy = total fraction of X variance "explained" by the Y variables, through ALL CV's

For variable Y:

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

$j := 1 .. \text{rows}(r_Y)$

$$r_Y = \begin{pmatrix} 0.8404 & 0.3588 & 0.1354 & -0.3118 & -0.3118 \\ 0.8765 & -0.0648 & 0.2546 & 0.1052 & 0.1052 \\ 0.7639 & 0.2979 & 0.1478 & 0.4814 & 0.4815 \\ 0.6584 & 0.6768 & -0.2304 & 0.081 & 0.081 \\ 0.3641 & -0.7549 & -0.5434 & -0.0198 & -0.0198 \end{pmatrix}$$

< Y structural correlations

$$R_{Y_{i,j}} := \left(r_{Y_{i,j}}\right)^2 \qquad R_Y = \begin{pmatrix} 0.7064 & 0.1288 & 0.0183 & 0.0972 & 0.0972 \\ 0.7683 & 0.0042 & 0.0648 & 0.0111 & 0.0111 \\ 0.5836 & 0.0888 & 0.0218 & 0.2318 & 0.2318 \\ 0.4335 & 0.4581 & 0.0531 & 0.0066 & 0.0066 \\ 0.1326 & 0.5699 & 0.2953 & 0.0004 & 0.0004 \end{pmatrix}$$

< squarred structural correlations
= fraction variance "explained"
by each variable

$$Ycanvad_i := \text{mean}\left(R_Y^{\langle i \rangle}\right) \qquad Ycanvad = \begin{pmatrix} 0.5248768 \\ 0.2499409 \\ 0.0906618 \\ 0.0694074 \\ 0.0694074 \end{pmatrix}$$

< canonical variable adequacies
= fraction of total Y variance
"explained" by each CV

$$yvrd_i := Ycanvad_i \cdot \lambda_{Sy_i} \qquad yvrd = \begin{pmatrix} 0.113045817 \\ 0.007013171 \\ 0.000980431 \\ 0 \\ 0 \end{pmatrix}$$

< canonical redundancies
= total fraction of Y variance "explained" by
the X variables, through EACH CV

$$yrd := \sum_i yvrd_i \qquad yrd = 0.12103942$$

< Total canonical redundancy
= total fraction of Y variance "explained" by
the X variables, through ALL CV's

Note: Rencher pp. 407-408 argues that the only proper "multivariate" measure for associations between variable set X & Y are the scaled canonical variate coefficients described above or, preferably standardized scaled canonical variate coefficients using standardized data values and partioned correlation matrix R shown below. He specifically recommends against using the structural correlations (loadings) or redundancy coefficients for any purpose. However, it is unclear to what extent his recommendations are generally followed. What one is looking for in making these comparisons are the relative contributions of original variables to each canonical variate (CV). As far as I can tell, these relative contributions are usually reflected in both coefficients and loadings. Far more important, it seems to me, is the decision whether to standardize the data and use the correlation matrix R as opposed to covariance matrix S. Standardizing the data gives a scale-less estimate of variable association. On the other hand when using the orignal data, scale of measurement of the original variables plays an important part in the values of S. This decision mirrors a similar choice in PCA.

Prototype in R:

```
#CANONICAL CORRELATION ANALYSIS
#data for prototype from:
#http://www.ats.ucla.edu/stat/r/dae/canonical.htm
getwd()
setwd("c:/DATA/Multivariate")
getwd()
library(CCA)
library(yacca)
M=read.table("UCLAidreCCAEx.txt")
M
summary(M)
psyc=M[,1:3]
acad=M[,4:8]

#using {yacca}
CCA=cca(psyc,acad)
CCA
summary(CCA)
```

Since the data involves the original variable sets X & Y and covariance matrix S, X & Y coefficients reported by cca() are scaled to unit variance for each variable as shown in calcuations above.

Structural correlations = "loadings" match calculations above, as do the aggregate Redundancy Coefficients.

Note: more information is available about the CCA fit and derived measures by using summary().

```
>CCA
Canonical Correlation Analysis

Canonical Correlations:
     CV 1      CV 2      CV 3
0.4640861 0.1675092 0.1039911

X Coefficients:
                 CV 1       CV 2       CV 3
Control     1.2538339 -0.6214776 -0.6616896
Concept    -0.3513499 -1.1876866  0.8267210
Motivation  1.2624204  2.0272641  2.0002283

Y Coefficients:
               CV 1         CV 2         CV 3
Read    0.044620600 -0.004910024  0.021380576
Write   0.035877112  0.042071478  0.091307329
Math    0.023417185  0.004229478  0.009398182
Science 0.005025152 -0.085162184 -0.109835014
Sex     0.632119234  1.084642326 -1.794647036

Structural Correlations (Loadings) - X Vars:
                 CV 1       CV 2       CV 3
Control    0.90404631 -0.3896883 -0.1756227
Concept    0.02084327 -0.7087386  0.7051632
Motivation 0.56715106  0.3508882  0.7451289

Structural Correlations (Loadings) - Y Vars:
             CV 1        CV 2       CV 3
Read    0.8404480 -0.35882541  0.1353635
Write   0.8765429  0.06483674  0.2545608
Math    0.7639483 -0.29794884  0.1477611
Science 0.6584139 -0.67679761 -0.2303551
Sex     0.3641127  0.75492811 -0.5434036

Aggregate Redundancy Coefficients
(Total Variance Explained):
       X | Y: 0.09297449
       Y | X: 0.1210394
```

Using Standardized Data and Correlation in CCA:

Ms := READPRN("c:/DATA/Multivariate/Mstand.txt") **< Standardized data was created in R (see script).**

$n := \mathrm{length}\left(M^{\langle 1\rangle}\right)$ $\quad n = 600$

$Xs := \mathrm{submatrix}(Ms, 1, 600, 1, 3)$ $\quad Ys := \mathrm{submatrix}(Ms, 1, 600, 4, 8)$

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

$l_{vec_i} := 1$

$$S_S := \frac{1}{n-1} \cdot Ms^T \cdot \left(I - \frac{1}{n} \cdot l_{vec} \cdot l_{vec}^T\right) \cdot Ms$$

Total Correlation matrix R:
(Note that the covariance matrix S of standardized data is equivalent to correlation matrix R)

$$S_S = \begin{pmatrix} 1 & 0.171 & 0.245 & 0.374 & 0.359 & 0.337 & 0.325 & 0.113 \\ 0.171 & 1 & 0.289 & 0.061 & 0.019 & 0.054 & 0.07 & -0.126 \\ 0.245 & 0.289 & 1 & 0.211 & 0.254 & 0.195 & 0.116 & 0.098 \\ 0.374 & 0.061 & 0.211 & 1 & 0.629 & 0.679 & 0.691 & -0.042 \\ 0.359 & 0.019 & 0.254 & 0.629 & 1 & 0.633 & 0.569 & 0.244 \\ 0.337 & 0.054 & 0.195 & 0.679 & 0.633 & 1 & 0.65 & -0.048 \\ 0.325 & 0.07 & 0.116 & 0.691 & 0.569 & 0.65 & 1 & -0.138 \\ 0.113 & -0.126 & 0.098 & -0.042 & 0.244 & -0.048 & -0.138 & 1 \end{pmatrix}$$

$$\lambda_{Rx} = \begin{pmatrix} 0.2154 \\ 0.0281 \\ 0.0108 \end{pmatrix} \quad E_{Rx} = \begin{pmatrix} 0.86 & -0.3574 & -0.442 \\ -0.2537 & -0.7189 & 0.5813 \\ 0.4428 & 0.5961 & 0.6832 \end{pmatrix}$$

Standardized Raw Canonical Variate Coefficients:

For variable set X:

$$E_{Rx} = \begin{pmatrix} 0.86 & -0.3574 & -0.442 \\ -0.2537 & -0.7189 & 0.5813 \\ 0.4428 & 0.5961 & 0.6832 \end{pmatrix}$$

< Eigenvectors (columns) are the raw canonical variate coefficients. However, eigenvectores never represent unique solutions (lengths). So the vectors are normally scaled to unit variance.

For variable set Y:

$$E_{Ry} = \begin{pmatrix} 0.6539 & -0.0463 & -0.1296 & -0.7539 & -0.7539 \\ 0.5062 & 0.3821 & -0.5327 & 0.09 & 0.09 \\ 0.3198 & 0.0372 & -0.0531 & 0.6417 & 0.6417 \\ 0.0708 & -0.7718 & 0.6394 & 0.1073 & 0.1073 \\ 0.457 & 0.5047 & 0.5365 & -0.0167 & -0.0166 \end{pmatrix}$$

Standardized Scaled Canonical Variate Coefficients:

For variable X:

$N_x := Xs \cdot E_{Rx}$ **< projects each X object onto CV axis**

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

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

$l_{vec_i} := 1$

$$S_{Nx} := \frac{1}{n-1} \cdot N_x^{T} \cdot \left(I - \frac{1}{n} \cdot l_{vec} \cdot l_{vec}^{T}\right) \cdot N_x$$

$$S_{Nx} = \begin{pmatrix} 1.0472 & -1.2511 \times 10^{-15} & -1.2467 \times 10^{-15} \\ -1.2246 \times 10^{-15} & 0.7362 & 3.5353 \times 10^{-15} \\ -1.248 \times 10^{-15} & 3.5234 \times 10^{-15} & 0.9932 \end{pmatrix}$$

^ co variance of all objcects in CV coordinates

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

$$D_{Nx} = \begin{pmatrix} 1.0233 & 0 & 0 \\ 0 & 0.858 & 0 \\ 0 & 0 & 0.9966 \end{pmatrix}$$

< standard deviation

$$E_{Rx} \cdot D_{Nx}^{-1} = \begin{pmatrix} 0.8404196 & -0.4165639 & -0.4435172 \\ -0.2478818 & -0.8379278 & 0.583262 \\ 0.4326685 & 0.6948029 & 0.685537 \end{pmatrix}$$

< X standardized coefficients - scaled by recriprocal of the calculated standard deviation

For variable Y:

$N_y := Ys \cdot E_{Ry}$ **< projects each Y object onto CV axis**

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

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

$l_{vec_i} := 1$

$$S_{Ny} := \frac{1}{n-1} \cdot N_y^{T} \cdot \left(I - \frac{1}{n} \cdot l_{vec} \cdot l_{vec}^{T}\right) \cdot N_y$$

variance of all objcects in CV coordinates

$$S_{Ny} = \begin{pmatrix} 2.1043 & -0 & 0 & 0 & 0 \\ -0 & 0.8719 & -0 & 0 & 0 \\ 0 & -0 & 0.3597 & 0 & 0 \\ 0 & 0 & 0 & 0.319 & 0.319 \\ 0 & 0 & 0 & 0.319 & 0.319 \end{pmatrix}$$

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

$$D_{Ny} = \begin{pmatrix} 1.4506 & 0 & 0 & 0 & 0 \\ 0 & 0.9337 & 0 & 0 & 0 \\ 0 & 0 & 0.5998 & 0 & 0 \\ 0 & 0 & 0 & 0.5648 & 0 \\ 0 & 0 & 0 & 0 & 0.5648 \end{pmatrix}$$

^ standard deviation

$$E_{Ry} \cdot D_{Ny}^{-1} = \begin{pmatrix} 0.4508012 & -0.0496059 & -0.2160076 & -1.3348643 & -1.3348382 \\ 0.3489571 & 0.4092063 & -0.8880966 & 0.1593076 & 0.1592656 \\ 0.2204666 & 0.0398194 & -0.0884814 & 1.1360946 & 1.1361325 \\ 0.048775 & -0.8265994 & 1.0660783 & 0.1899452 & 0.1899196 \\ 0.3150396 & 0.540571 & 0.8944276 & -0.0294858 & -0.0294778 \end{pmatrix}$$

< Y standardized coefficients - scaled by recriprocal of the calculated standard deviation

```
#standardized dataset:
Ms=as.matrix(scale(M))
summary(Ms)
write.table(Ms,"Mstand.txt")

#CCA standardized analysis:
CCAs=cca(psyc,acad,xscale=TRUE,yscale=TRUE)
CCAs
```

With standardized data:

- **canonical correlations stay the same.**
- **X & Y Standardized Scaled Coefficients are different from unstandardized coefficients calculated shown above.**
- **Structural Correlations stay the same.**
- **Redundancy Coefficients stay the same.**

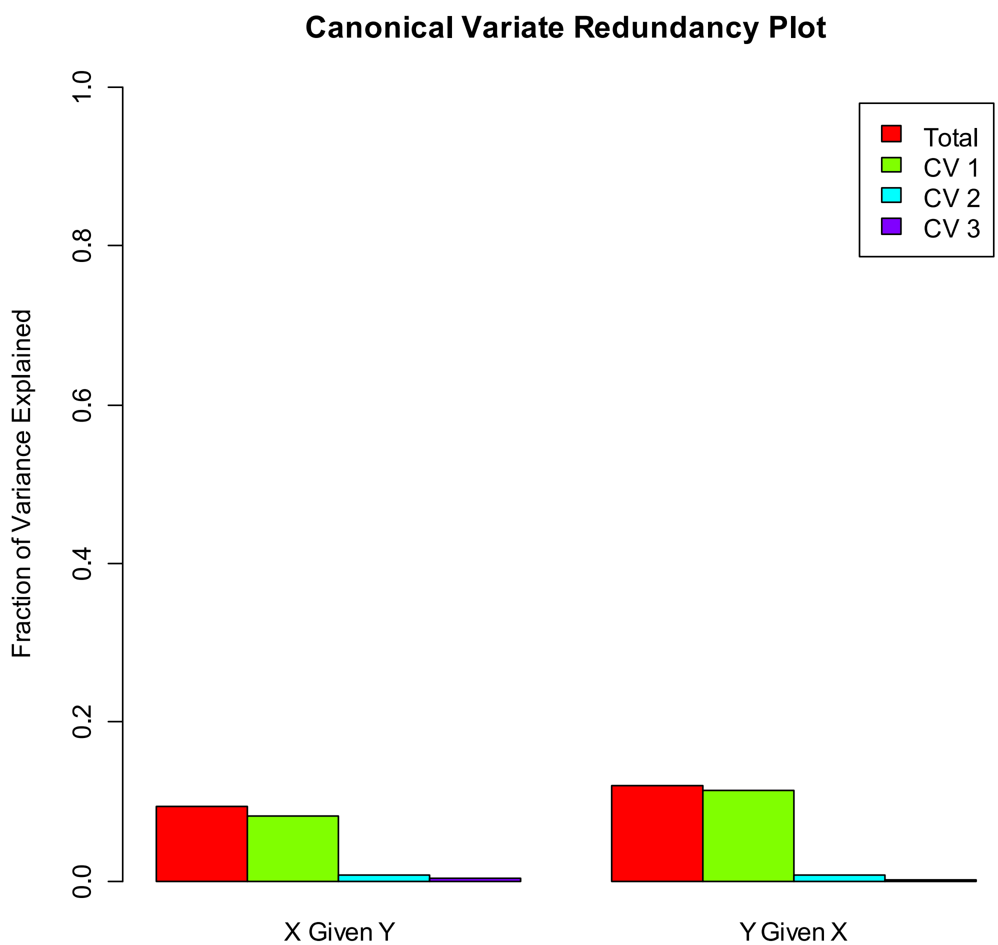
```
> CCAs
Canonical Correlation Analysis

Canonical Correlations:
     CV 1      CV 2      CV 3
0.4640861 0.1675092 0.1039911

X Coefficients:
                CV 1       CV 2       CV 3
Control    0.8404196  0.4165639  0.4435172
Concept   -0.2478818  0.8379278 -0.5832620
Motivation 0.4326685 -0.6948029 -0.6855370

Y Coefficients:
             CV 1        CV 2        CV 3
Read    0.45080116  0.04960589 -0.21600760
Write   0.34895712 -0.40920634 -0.88809662
Math    0.22046662 -0.03981942 -0.08848141
Science 0.04877502  0.82659938  1.06607828
Sex     0.31503962 -0.54057096  0.89442764

Structural Correlations (Loadings) - X Vars:
                 CV 1       CV 2       CV 3
Control    0.90404631  0.3896883  0.1756227
Concept    0.02084327  0.7087386 -0.7051632
Motivation 0.56715106 -0.3508882 -0.7451289

Structural Correlations (Loadings) - Y Vars:
            CV 1        CV 2       CV 3
Read    0.8404480  0.35882541 -0.1353635
Write   0.8765429 -0.06483674 -0.2545608
Math    0.7639483  0.29794884 -0.1477611
Science 0.6584139  0.67679761  0.2303551
Sex     0.3641127 -0.75492811  0.5434036

Aggregate Redundancy Coefficients
(Total Variance Explained):
        X | Y: 0.09297449
        Y | X: 0.1210394
```

Plotting CCA:

Scree plots:

In practical work, I have mostly observed CCA used as a heuristic device to summarize data with high dimensionality with as few canonical variates (CV) as possible. Because two sets of variables (X and Y) are involved, canonical variate "scores" are calculated for each set. The scores as linear combinations of the original variables (as specified by the scaled coefficients) represent the location of each object (row of the original data set) along axes of "maximum correlation" between the two sets X & Y. As in PCA, because the original data have high dimension, there are multiple directions of "maximum correlation", each direction uncorrelated with the others. The first CV provides information about the direction of largest "maxium correlation". CV 2, CV 3, etc. in order represent successively lesser directions. "Scree" plots of Redundancy ("variance explained") give a good indication of the relative importance of the CV's (CV's with low "variance explained" values are often ignored.) In this example CV1 "explains" by far the most variance, with CV2 perhaps deserving only a little scrutiny (especially if scaled coefficients or loadings can be interpreted). PC3 can be safely ignored.

```
>plot(CCAs)
```



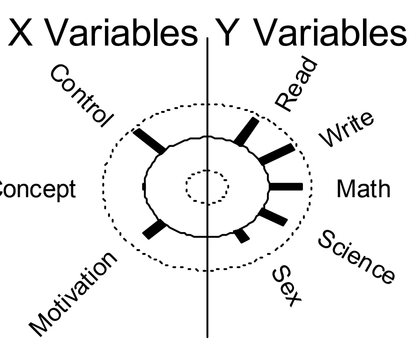
Note that the "fraction of variance explained" in a redundancy plot involves calculations related to two conditional statments:

- **"Variance explained" in X given Y, and**
- **"Variance explained" in Y given X.**

Helio plots:

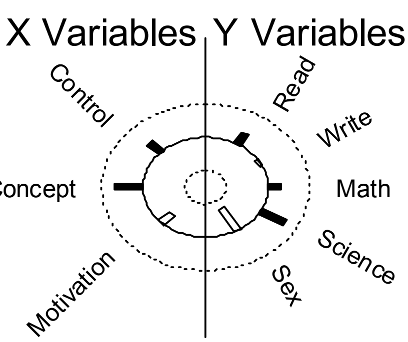
The relative importance to the CV's of individual variables in groups X & Y can be appraised graphically by consulting (standardized) scaled coefficient values as recommended by Rencher or by looking at "structural correlations" or "variance explained" as utilized in R package {yacca}. Although graph type is mostly a matter of style, "helio plots" appear popular. In these, the two variable groups are placed on the left versus right sides of the graph (which side is arbitrary). For each variable, positive correlations with the CV is typically indicated by solid bars oriented in the centrifugal direction (i.e., going away from the center), and negative correlations by empty bars in the centripetal direction (i.e., going toward the center). The amount of correlation in each case is shown by the relative lengths of each bar.

Structural Correlations for CV 1



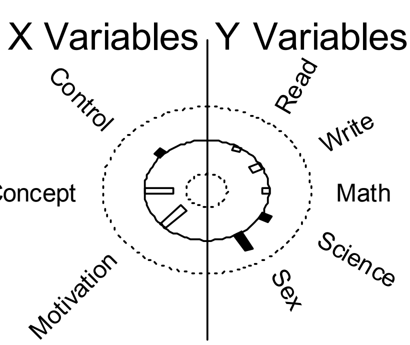
Canonical Variate1

Structural Correlations for CV 2



Canonical Variate2

Structural Correlations for CV 3



Canonical Variate3

In this example, positive correlations in CV1 are noted between all acad (Y) variables with "control" and "motivation" psyc (X) variables, but not with "concept". Perhaps this can be interpreted mostly as a "general" performance relationship of individuals on both batteries of tests. (Is this IQ??). As a result, CV2 may also be of interest although it must be admitted that this represents a much lesser effect. Here we see that "read", "science" and "sex" among acad variables specifically relate mostly to "concept" and perhaps to a lesser degree "control" with an opposite but strong effect on "motivation". What's needed now is informed interpretation based on the meaning (or not) of these measures.

Optimizing Helio Plots:

```
#helio plots with more control:
helio.plot(CCAs,cv=1,type="correlation")
helio.plot(CCAs,cv=2,type="variance",x.name="psyc",y.name="acad",main="HP Variance")
?helio.plot
```

The plotting function helio.plot() in {yacca} provides a very nice set of tools for optimizing helio plots.

One choice involves a "type" switch specifying choice between "explained variance" or "structural correlation".

Note: "explained variance" is merely the square of "structural correlations". Examples are from calculations above:

$$r_X = \begin{pmatrix} 0.904 & 0.3897 & -0.1756 \\ 0.0208 & 0.7087 & 0.7052 \\ 0.5672 & -0.3509 & 0.7451 \end{pmatrix}$$

^ this column for the first plot X

$$R_X = \begin{pmatrix} 0.8173 & 0.1519 & 0.0308 \\ 0.0004 & 0.5023 & 0.4973 \\ 0.3217 & 0.1231 & 0.5552 \end{pmatrix}$$

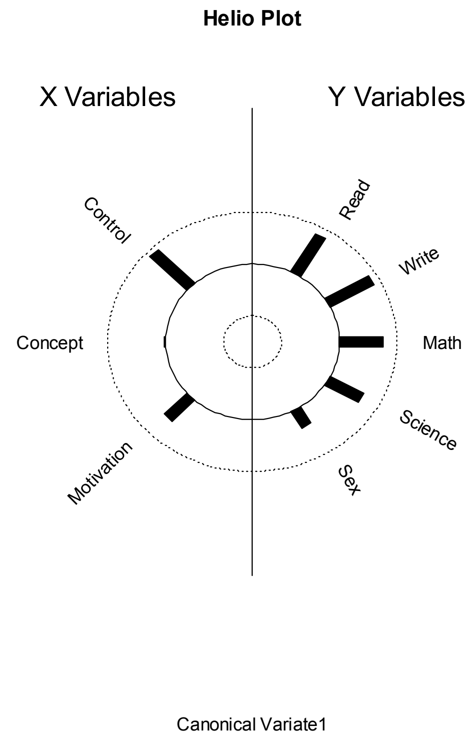
^ this column for the second plot X

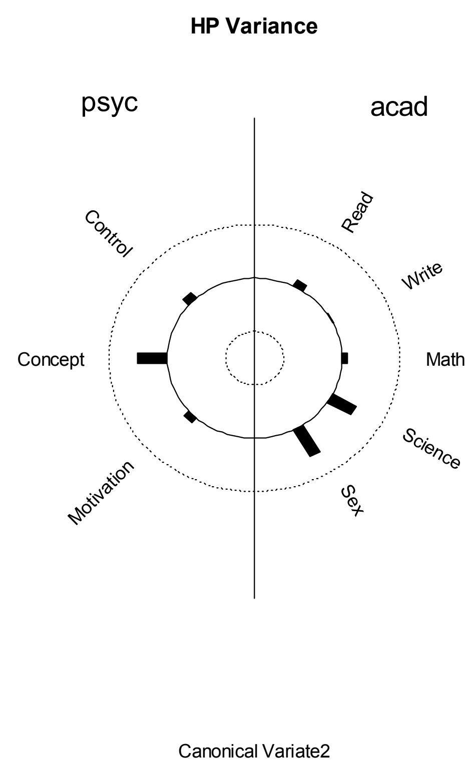
$$r_Y = \begin{pmatrix} 0.8404 & 0.3588 & 0.1354 & -0.3118 & -0.3118 \\ 0.8765 & -0.0648 & 0.2546 & 0.1052 & 0.1052 \\ 0.7639 & 0.2979 & 0.1478 & 0.4814 & 0.4815 \\ 0.6584 & 0.6768 & -0.2304 & 0.081 & 0.081 \\ 0.3641 & -0.7549 & -0.5434 & -0.0198 & -0.0198 \end{pmatrix}$$

^this column for the first plot Y

$$R_Y = \begin{pmatrix} 0.7064 & 0.1288 & 0.0183 & 0.0972 & 0.0972 \\ 0.7683 & 0.0042 & 0.0648 & 0.0111 & 0.0111 \\ 0.5836 & 0.0888 & 0.0218 & 0.2318 & 0.2318 \\ 0.4335 & 0.4581 & 0.0531 & 0.0066 & 0.0066 \\ 0.1326 & 0.5699 & 0.2953 & 0.0004 & 0.0004 \end{pmatrix}$$

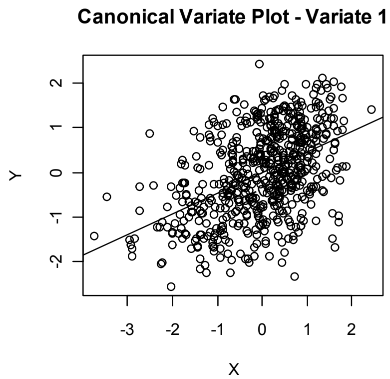
^ this column for the second plot Y

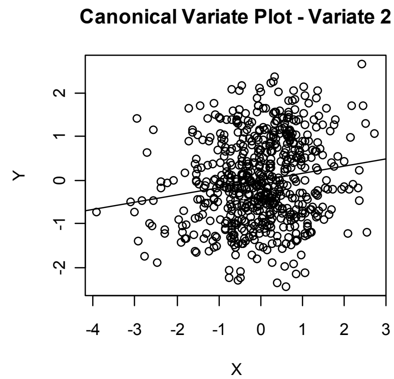




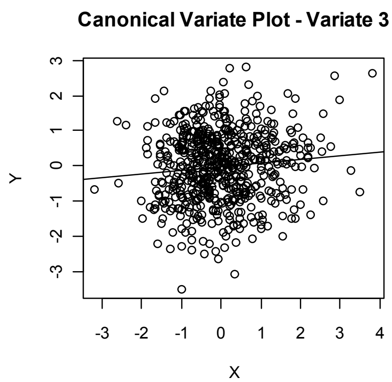
Plots of CV Scores:

Scores for each object on each CV may be plotted to look for substructure in the data. Scores are calculated separately for each variable group (X or Y), and plots are commonly made comparing X versus Y for each CV.





< output from plot.cv in R



#plot) will give also give output...
#plot ing CV scores directly:
#look ing at standardized scaled coefficients:
CCAs$ xcoef
CCAs$ ycoef
#look ing at the scores:
CCAs$ canvarx
CCAs$ canvary

Calculation of CV Scores:

standardized scaled coefficients

For variable X:

$$Coef_X := E_{Rx} \cdot D_{Nx}^{-1} \qquad Coef_X = \begin{pmatrix} 0.8404196 & -0.4165639 & -0.4435172 \\ -0.2478818 & -0.8379278 & 0.583262 \\ 0.4326685 & 0.6948029 & 0.685537 \end{pmatrix}$$

$$Score_X := Xs \cdot Coef_X$$

^ Scores for X

> CCAs$xcoef

```
                  CV 1       CV 2       CV 3
Control      0.8404196  0.4165639  0.4435172
Concept     -0.2478818  0.8379278 -0.5832620
Motivation   0.4326685 -0.6948029 -0.6855370
```

For variable Y:

$$Coef_Y := E_{Ry} \cdot D_{Ny}^{-1} \qquad Coef_Y = \begin{pmatrix} 0.4508012 & -0.0496059 & -0.2160076 & -1.3348643 & -1.3348382 \\ 0.3489571 & 0.4092063 & -0.8880966 & 0.1593076 & 0.1592656 \\ 0.2204666 & 0.0398194 & -0.0884814 & 1.1360946 & 1.1361325 \\ 0.048775 & -0.8265994 & 1.0660783 & 0.1899452 & 0.1899196 \\ 0.3150396 & 0.540571 & 0.8944276 & -0.0294858 & -0.0294778 \end{pmatrix}$$

$$Score_Y := Ys \cdot Coef_Y$$

^ Scores for Y

> CCAs$ycoef

```
              CV 1        CV 2        CV 3
Read    0.45080116  0.04960589 -0.21600760
Write   0.34895712 -0.40920634 -0.88809662
Math    0.22046662 -0.03981942 -0.08848141
Science 0.04877502  0.82659938  1.06607828
Sex     0.31503962 -0.54057096  0.89442764
```

Example scores for X:

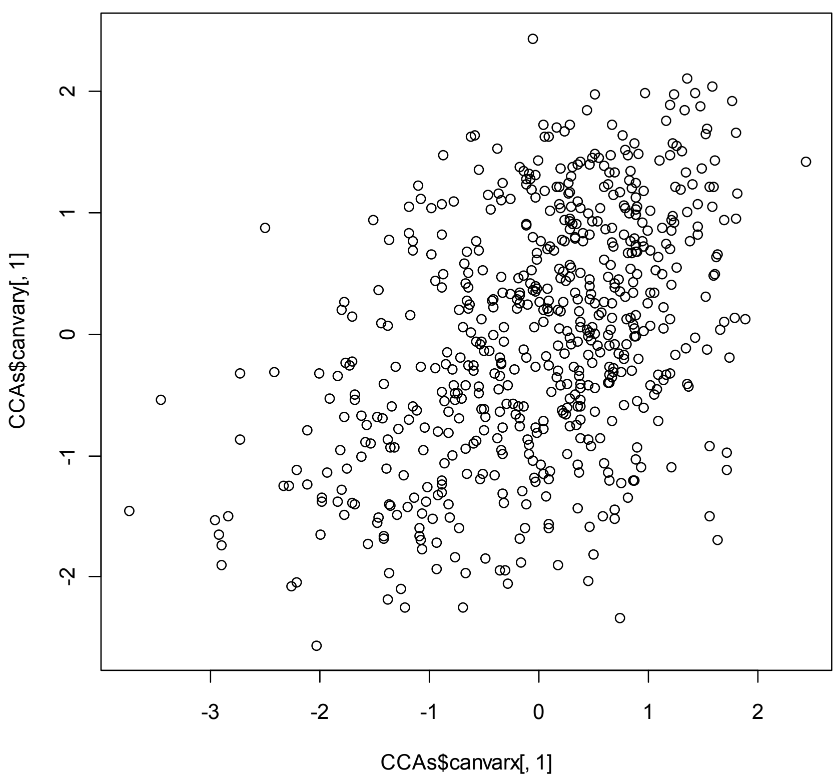
Score_X =

	1	2	3
1	-0.66003488	1.56049911	1.09562739
2	-0.41905952	0.8787902	-0.058971
3	0.80087859	-1.16943412	-0.02299254
4	0.68410697	-0.68938531	-0.16017192
5	-0.50413258	1.11552821	1.18650414
6	0.53858065	-2.36361183	-0.59235837
7	-0.04304757	0.01149676	0.06324607
8	-1.04142949	1.3506957	0.19251796
9	0.86254637	0.43811766	0.46526248
10	-0.11827761	0.04878542	0.10294745
11	-0.08495942	-0.24330315	-1.11622735
12	-1.81228799	0.38135529	-0.09563762
13	0.45848599	-0.23709426	-0.20142977
14	0.72378895	-0.40944993	-0.32219985
15	0.50224231	2.08749097	-0.03719933
16	-0.35536065	-0.47697168	-0.80037614

```
> CCAs$canvarx
          CV 1          CV 2          CV 3
1  -0.66003488 -1.5604991085 -1.095627388
2  -0.41905952 -0.8787901977  0.058971003
3   0.80087859  1.1694341207  0.022992536
4   0.68410697  0.6893853085  0.160171916
5  -0.50413258 -1.1155282096 -1.186504137
6   0.53858065  2.3636118290  0.592358375
7  -0.04304757 -0.0114967620 -0.063246071
8  -1.04142949 -1.3506957005 -0.192517962
9   0.86254637 -0.4381176646 -0.465262482
10 -0.11827761 -0.0487854159 -0.102947447
11 -0.08495942  0.2433031504  1.116227346
12 -1.81228799 -0.3813552856  0.095637620
13  0.45848599  0.2370942637  0.201429766
14  0.72378895  0.4094499263  0.322199846
15  0.50224231 -2.0874909715  0.037199327
16 -0.35536065  0.4769716774  0.800376136
...
```

```
#plot for CV 1:
plot(CCAs$canvarx[,1],CCAs$canvary[,1])
#plot for CV 2:
plot(CCAs$canvarx[,2],CCAs$canvary[,2])
```

Example plot for CV1:



All plots here suggest little evidence for substructure in the data, although it might be useful to further investigate any difference in scatter between males versus females (Y variable "sex").