Canonical Correlations

Dr. (Mrs.) Kirtee Kiran Kamalja
Department of Statistics, School of Mathematical Sciences,
Kavayitri Bahinbai Chaudhari North Maharashtra University,
Jalgaon

Purpose of Canonical Correlation Analysis (CCA)

i) **Data Reduction**: Canonical correlation analysis provides a way for explaining the relationship between 2 sets of variables using linear combinations of these variables.

ii) **Data interpretation:** This multivariate analysis technique is also of particular interest when it comes to finding features (canonical variates) that are statistically significant in terms of explaining the covariance between the given sets of variables.

Canonical Correlations analysis

- Canonical correlation analysis seeks to identify and quantify the associations between two sets of variable
- Among the two large sets of variates an investigator may want to study the interrelations.
- If the two sets are very large, she may want to consider only a few linear combinations of each set.
- An investigator may study those linear combinations most highly correlated.

Objectives of Canonical Correlations analysis…

- To determine the pair of linear combinations having the largest correlation
- To determine the another pair of linear combinations having the largest correlation among all pairs uncorrelated with the initially selected pair, and so on.
- The pairs of linear combinations are called the canonical variables, and their correlations are called canonical correlations.

Example 1:

Set 1: Measurements of physical characteristics, such as various lengths and breadths of skulls.

Set 2: Measurements of mental characteristics, such as scores on intelligence tests

Objective: Study/relate the variables in Set 1 and Set 2

Examples…

- Relating governmental policy variables with economic goal variables
- Relating college "performance" variables with precollege "achievement" variables
- Relating arithmetic speed and arithmetic power to reading speed and reading power

Examples…

- A researcher has collected data on three psychological variables, four academic variables (standardized test scores) and gender for 600 college freshman.
- She is interested in how the set of psychological variables relates to the academic variables and gender.
- In particular, the researcher is interested in how many dimensions (canonical variables) are necessary to understand the association between the two sets of variables

What is Canonical Correlations?

- We find linear combinations of variables in the sets that have maximum correlation; these linear combinations are the first coordinates in the new systems.
- Then a second linear combination in each set is sought such that the correlation between these is the maximum of correlations between such linear combinations as are uncorrelated with the first linear combinations.
- The procedure is continued until the two new coordinate systems are completely specified.

Definition:

- The **first pair of canonical variables**, or first canonical variate pair, is the pair of linear combinations U_1, V_1
  i) having unit variances
  ii) which maximize the correlation.
- The **second pair of canonical variables**, or second canonical variate pair, is the pair of linear combinations U_2, V_2
  i) having unit variances,
  ii) which maximize the correlation among all choices
  iii) that are uncorrelated with the first pair of canonical variables

Definition…

- The **k^{th} pair of canonical variables**, or k^{th} canonical variate pair, is the pair of linear combinations U_k, V_k
  i) having unit variances,
  ii) which maximize the correlation among all choices
  iii) that are uncorrelated with the previous $k-1$ pair of canonical variables.
- The correlation between the k^{th} pair of canonical variables is called the **k^{th} canonical correlation coefficient**.

First pair of canonical variables

- Let $\underline{X}$, $\underline{Y}$ be p and q $(p \leq q)$ component vectors (i.e. two sets of random variables) respectively such that $E\begin{pmatrix}\underline{X}\\ \underline{Y}\end{pmatrix} = \begin{pmatrix}\underline{0}\\ \underline{0}\end{pmatrix}$ and $Cov\begin{pmatrix}\underline{X}\\ \underline{Y}\end{pmatrix} = \begin{pmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{pmatrix} > 0.$
- The first pair of canonical variables $(\mathrm{U}_1, \mathrm{V}_1)$ is such that

i) $U_1 = \underline{a}'\underline{X}$ and $V_1 = \underline{b}'\underline{Y}$ (i.e. U_1 is linear combination of X-variables and V_1 is linear combination of Y-variables)

ii) $\mathrm{corr}(\mathrm{U}_1, \mathrm{V}_1)$ is maximum and

iii) $var(U_1) = 1$ and $var(V_1) = 1$

First pair of canonical variables

- $\text{cov}(\text{U}_1, \text{V}_1) = cov\left(\underline{a}'\underline{X}, \underline{b}'\underline{Y}\right)$

$$= \underline{a}'\Sigma_{12}\underline{b}$$

- $\text{Va}r(\text{U}_1) = \underline{a}'\Sigma_{11}\underline{a} = 1$

- $\text{Va}r(\text{V}_1) = \underline{b}'\Sigma_{22}\underline{b} = 1$

- $Corr\left(\underline{a}'\underline{X}, \underline{b}'\underline{Y}\right) = \dfrac{\underline{a}'\Sigma_{12}\underline{b}}{\sqrt{\underline{a}'\Sigma_{11}\underline{a}}\sqrt{\underline{b}'\Sigma_{22}\underline{b}}}.$

Derivation of first pair of canonical variables

Maximize function ϕ where,

$$\phi = \underline{a}'\Sigma_{12}\underline{b} - \frac{1}{2}\lambda_1\left(\underline{a}'\Sigma_{11}\underline{a} - 1\right) - \frac{1}{2}\lambda_2\left(\underline{b}'\Sigma_{22}\underline{b} - 1\right)$$

with respect to $\underline{a}$ and $\underline{b}$ where λ_1 and λ_2 are Langrangian multipliers.

$$\frac{d\phi}{d\underline{a}} = \Sigma_{12}\underline{b} - \lambda_1\Sigma_{11}\underline{a} = \underline{0} \qquad (1)$$

$$\frac{d\phi}{d\underline{b}} = \Sigma_{21}\underline{a} - \lambda_2\Sigma_{22}\underline{b} = \underline{0} \qquad (2)$$

First pair of canonical variables…

$$\frac{d\phi}{d\lambda_1} = \left(\underline{a}'\Sigma_{11}\underline{a} - 1\right) = 0 \qquad \text{(a)}$$

$$\frac{d\phi}{d\lambda_2} = \left(\underline{b}'\Sigma_{22}\underline{b} - 1\right) = 0 \qquad \text{(b)}$$

(1)⇒ $\underline{a}'\Sigma_{12}\underline{b} - \lambda_1\underline{a}'\Sigma_{11}\underline{a} = \underline{0}$ (3)

(2)⇒ $\underline{b}'\Sigma_{21}\underline{a} - \lambda_2\underline{b}'\Sigma_{22}\underline{b} = \underline{0}$ (4)

From (3) and (4)

$$\lambda_1 = \lambda_2 = \underline{a}'\Sigma_{12}\underline{b} = \lambda$$

First pair of canonical variables…

Use $\lambda_1 = \lambda_2 = \underline{a}'\Sigma_{12}\underline{b} = \lambda$ in (1) and (2):

$$\Sigma_{12}\underline{b} - \lambda_1\Sigma_{11}\underline{a} = \underline{0} \qquad (1)$$

$$\Sigma_{21}\underline{a} - \lambda_2\Sigma_{22}\underline{b} = \underline{0} \qquad (2)$$

Hence (1) and (2) can be rewritten as

$$-\lambda\Sigma_{11}\underline{a} + \Sigma_{12}\underline{b} = \underline{0} \qquad (5)$$

$$\Sigma_{21}\underline{a} - \lambda\Sigma_{22}\underline{b} = \underline{0} \qquad (6)$$

First pair of canonical variables…

Equations (5) and (6) can be rewritten as

$$\begin{pmatrix} -\lambda\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\lambda\Sigma_{22} \end{pmatrix} \begin{pmatrix} \underline{a} \\ \underline{b} \end{pmatrix} = \begin{pmatrix} 0 \\ \underline{0} \end{pmatrix}$$

Now $\underline{a}, \underline{b}$ are non-null if and only if

$$\begin{vmatrix} -\lambda\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\lambda\Sigma_{22} \end{vmatrix} = 0$$

The determinant on left is polynomial of degree p

First pair of canonical variables…

Consider the transformation

$$\Sigma_{11}^{\frac{1}{2}}\underline{a} = \underline{c} \qquad \text{i.e.} \qquad \underline{a} = \Sigma_{11}^{\frac{-1}{2}}\underline{c}$$

and $$\Sigma_{22}^{\frac{1}{2}}\underline{b} = \underline{d} \qquad \text{i.e.} \qquad \underline{b} = \Sigma_{22}^{\frac{-1}{2}}\underline{d}$$

(5)⇒ $$-\lambda\Sigma_{11}^{\frac{1}{2}}\underline{c} + \Sigma_{12}\Sigma_{22}^{\frac{-1}{2}}\underline{d} = \underline{0} \qquad (7)$$

(6)⇒ $$\Sigma_{21}\Sigma_{11}^{\frac{-1}{2}}\underline{c} - \lambda\Sigma_{22}^{\frac{1}{2}}\underline{d} = \underline{0} \qquad (8)$$

(8)⇒ $$\lambda\underline{d} = \Sigma_{22}^{\frac{-1}{2}}\Sigma_{21}\Sigma_{11}^{\frac{-1}{2}}\underline{c}$$

First pair of canonical variables…

Substituting $\lambda \underline{d} = \Sigma_{22}^{\frac{-1}{2}} \Sigma_{21} \Sigma_{11}^{\frac{-1}{2}} \underline{c}$ in $\lambda \times(7)$

$$-\lambda^2 \Sigma_{11}^{\frac{1}{2}} \underline{c} + \lambda \Sigma_{12} \Sigma_{22}^{\frac{-1}{2}} \underline{d} = \underline{0}$$

$$-\lambda^2 \Sigma_{11}^{\frac{1}{2}} \underline{c} + \Sigma_{12} \Sigma_{22}^{\frac{-1}{2}} \boldsymbol{\Sigma_{22}^{\frac{-1}{2}} \Sigma_{21} \Sigma_{11}^{\frac{-1}{2}} \underline{c}} = \underline{0}$$

$$-\lambda^2 \underline{c} + \Sigma_{11}^{-\frac{1}{2}} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1/2} \underline{c} = \underline{0}$$

$$-\lambda^2 \underline{c} + \Sigma_{11}^{-\frac{1}{2}} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1/2} \underline{c} = \underline{0}$$

$$\left(\Sigma_{11}^{-\frac{1}{2}} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-\frac{1}{2}} - \lambda^2 I \right) \underline{c} = \underline{0} \qquad (9)$$

First pair of canonical variables…

(9) has non-null solution ($i.e.\ \underline{c} \neq \underline{0}$) if and only if

$$\left|\Sigma_{11}^{-\frac{1}{2}}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-\frac{1}{2}} - \lambda^2 I_p\right| = 0$$

Let $A = \Sigma_{22}^{-\frac{1}{2}}\Sigma_{21}\Sigma_{11}^{-\frac{1}{2}}$ be $q \times p$ matrix. Then,

$$\left|A'A - \lambda^2 I_p\right| = 0$$

If $\rho_1^{*2} \geq \rho_2^{*2} \geq \cdots \geq \rho_p^{*2}$ are the eigen values of $\Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1/2}$ and $\underline{e}_1, \underline{e}_2, \dots, \underline{e}_p$ are the corresponding orthonormalized eigen-vectors then choose $\lambda^2 = \rho_1^{*2}$.

First pair of canonical variables…

(9) becomes

$$\left(\Sigma_{11}^{-\frac{1}{2}}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-\frac{1}{2}} - \rho_1^{*2} I\right)\underline{c} = \underline{0}$$

Observe that $\underline{c} = \underline{e}_1$ (eigen vector corresponding to ρ_1^{*2}) satisfies the above equation.

Thus $\underline{c} = \underline{e}_1$

$$\Sigma_{11}^{\frac{1}{2}}\underline{a} = \underline{e}_1$$

$$\boldsymbol{\underline{a} = \Sigma_{11}^{\frac{-1}{2}}\underline{e}_1}$$

First pair of canonical variables…

(7)⇒ $\lambda \underline{c} = \Sigma_{11}^{\frac{-1}{2}} \Sigma_{12} \Sigma_{22}^{\frac{-1}{2}} \underline{d}$

Substituting this in $\lambda \times$(8)

$$\lambda \Sigma_{21} \Sigma_{11}^{\frac{-1}{2}} \underline{c} - \lambda^2 \Sigma_{22}^{\frac{1}{2}} \underline{d} = \underline{0}$$

$$\Sigma_{21} \Sigma_{11}^{\frac{-1}{2}} \Sigma_{11}^{\frac{-1}{2}} \Sigma_{12} \Sigma_{22}^{\frac{-1}{2}} \underline{d} - \rho_1^{*2} \Sigma_{22}^{\frac{1}{2}} \underline{d} = \underline{0}$$

$$\left(\Sigma_{22}^{-\frac{1}{2}} \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-\frac{1}{2}} - \rho_1^{*2} I_q \right) \underline{d} = \underline{0} \qquad (10)$$

First pair of canonical variables…

$$\left(AA' - {\rho_1^*}^2 I_q\right)\underline{d} = \underline{0}$$

(10) has non-null solution if $\underline{d} = \underline{f_1}$ as ${\rho_1^*}^2$ is the largest eigen value of $\Sigma_{22}^{-1/2}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1/2}$ and $\underline{f_1}$ is the corresponding eigen-vector and is proportional to $\Sigma_{22}^{-1/2}\Sigma_{21}\Sigma_{11}^{-1}\underline{e}_1$.

Thus $$\underline{d} = \underline{f_1}$$

$$\underline{\boldsymbol{b}} = \boldsymbol{\Sigma}_{22}^{\frac{-1}{2}}\underline{\boldsymbol{f}}_{\mathbf{1}}$$

First pair of canonical variables

- $U_1 = \underline{e}_1' \Sigma_{11}^{-\frac{1}{2}} \underline{X}$ and $V_1 = \underline{f_1}' \Sigma_{22}^{-\frac{1}{2}} \underline{Y}$ where

- $({\rho_1^*}^2, \underline{e}_1)$ is the eigen value-eigen vector pair of $A'A = \Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1/2}$ and

- $({\rho_1^*}^2, \underline{f_1})$ is the eigen value-eigen vector pair of $AA' = \Sigma_{22}^{-1/2} \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1/2}$

k^{th} pair of canonical variables

- $U_k = \underline{e}_k' \Sigma_{11}^{-\frac{1}{2}} \underline{X}$ and $V_k = \underline{f}_k' \Sigma_{22}^{-\frac{1}{2}} \underline{Y}$ where

- $({\rho_k^*}^2, \underline{e}_k)$ is the k^{th}-largest eigen value-eigen vector pair of $\mathrm{A}'A = \Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1/2}$ and

- $({\rho_k^*}^2, \underline{f}_k)$ is the k^{th}-largest eigen value-eigen vector pair of $\mathrm{AA}' = \Sigma_{22}^{-1/2}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1/2}$

Pairs of canonical variables

- $\underline{U} = \begin{pmatrix} U_1 \\ U_2 \\ \vdots \\ U_p \end{pmatrix} = \begin{pmatrix} \underline{e}_1' \\ \underline{e}_2' \\ \vdots \\ \underline{e}_p' \end{pmatrix} \Sigma_{11}^{-\frac{1}{2}} \underline{X} = E \Sigma_{11}^{-\frac{1}{2}} \underline{X}$

- $\underline{V} = \begin{pmatrix} V_1 \\ V_2 \\ \vdots \\ V_q \end{pmatrix} = \begin{pmatrix} \underline{f}_1' \\ \underline{f}_2' \\ \vdots \\ \underline{f}_q' \end{pmatrix} \Sigma_{22}^{-\frac{1}{2}} \underline{Y} = F \Sigma_{22}^{-\frac{1}{2}} \underline{Y}$

Pairs of canonical variables…

- $cov(\underline{U},\underline{X}) = cov\left(E\Sigma_{11}^{-\frac{1}{2}}\underline{X},\underline{X}\right) = E\Sigma_{11}^{\frac{1}{2}}$

- $cov(\underline{V},\underline{Y}) = cov\left(F\Sigma_{22}^{-\frac{1}{2}}\underline{Y},\underline{Y}\right) = F\Sigma_{22}^{\frac{1}{2}}$

- $cov(\underline{U},\underline{Y}) = cov\left(E\Sigma_{11}^{-\frac{1}{2}}\underline{X},\underline{Y}\right) = E\Sigma_{11}^{\frac{-1}{2}}\Sigma_{12}$

- $cov(\underline{V},\underline{X}) = cov\left(F\Sigma_{22}^{-\frac{1}{2}}\underline{Y},\underline{X}\right) = F\Sigma_{22}^{\frac{-1}{2}}\Sigma_{21}$

Pairs of canonical variables…

- $cov(\underline{U}) = I_p$

- $cov(\underline{V}) = I_q$

- $cov(\underline{U}, \underline{V}) = diag(\rho_1^*, \rho_2^*, \ldots, \rho_p^*, 0, \ldots, 0)_{p\times q}$ $\boldsymbol{= R}$

- $\boldsymbol{cov\begin{pmatrix}\underline{U}\\ \underline{V}\end{pmatrix} = \begin{pmatrix} I_p & R \\ R' & I_q \end{pmatrix}}$ while $\boldsymbol{cov\begin{pmatrix}\underline{X}\\ \underline{Y}\end{pmatrix} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}}$

Other method

to obtain canonical correlations

Some results:

For $Cov\begin{pmatrix}\underline{X}\\ \underline{Y}\end{pmatrix} = \begin{pmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{pmatrix}$, let $A = \Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1/2}$

- $AA' = \Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1/2}$ is $p \times p$
- $A'A = \Sigma_{22}^{-1/2}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1/2}$ is $q \times q$
- $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$ be the eigen values of AA'
- $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_q \geq 0$ be the eigen values of $A'A$

Some results…

- Non-zero eigen values of AA' are same as non-zero eigen values of $A'A$ and the eigen value 0 has different multiplicities.
- Cauchy-Schwartz Inequality: For any two real vectors $\underline{a}$ and $\underline{b}$

$$\underline{a}'\underline{b} \leq \sqrt{(\underline{a}'\underline{a})(\underline{b}'\underline{b})}$$

Further equality holds if and only if $\underline{a} \propto \underline{b}$

Some results…

- $\underset{\underline{x} \neq \underline{0}}{Max} \frac{\underline{x}' B \underline{x}}{\underline{x}' \underline{x}} = \lambda_1$ (largest eigen-value of B) with equality at $\underline{x} = \underline{e}_1$ (the eigen vector corresponding to λ_1)

- $\underset{\underline{x} \perp \underline{e}_1}{Max} \frac{\underline{x}' B \underline{x}}{\underline{x}' \underline{x}} = \lambda_2$ (second-largest eigen-value of B) with equality at $\underline{x} = \underline{e}_2$ (the eigen vector corresponding to λ_2)

- $\underset{\underline{x} \perp \underline{e}_1, \underline{e}_2, \dots, \underline{e}_k}{Max} \frac{\underline{x}' B \underline{x}}{\underline{x}' \underline{x}} = \lambda_{k+1}$ $(k+1)^{th}$-largest eigen-value of B) with equality at $\underline{x} = \underline{e}_{k+1}$ (the eigen vector corresponding to λ_{k+1})

Derivation of first pair of canonical variables…

Theorem: Let $U_1 = \underline{a}'\underline{X}$ and $V_1 = \underline{b}'\underline{Y}$ where $Cov\begin{pmatrix}\underline{X}\\ \underline{Y}\end{pmatrix} = \begin{pmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{pmatrix}$. Then Max $\operatorname{corr}\left(\underline{a}'\underline{X}, \underline{b}'\underline{Y}\right) = \rho_1^*$ is attained at

$U_1 = \underline{e}_1'\Sigma_{11}^{-\frac{1}{2}}\underline{X}$ and $V_1 = \underline{f}_1'\Sigma_{22}^{-\frac{1}{2}}\underline{Y}$ where

Theorem…

${\rho_1^*}^2 \geq {\rho_2^*}^2 \geq \cdots \geq {\rho_p^*}^2$ are the eigen values of $\Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1/2}$ and $\underline{e}_1, \underline{e}_2, \dots, \underline{e}_p$ are the corresponding orthonormalized eigen-vectors.

Furthermore ${\rho_1^*}^2, {\rho_2^*}^2, \dots, {\rho_p^*}^2$ are the p largest eigen values of $\Sigma_{22}^{-1/2}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1/2}$ with eigen-vectors $\underline{f}_1, \underline{f}_2, \dots, \underline{f}_p$ which are proportional to $\Sigma_{22}^{-1/2}\Sigma_{21}\Sigma_{11}^{-1}\underline{e}_1$.

Proof:

$$\operatorname{corr}(\underline{a}'\underline{X}, \underline{b}'\underline{Y})$$

$$= \frac{\underline{a}'\Sigma_{12}\underline{b}}{\sqrt{(\underline{a}'\Sigma_{11}\underline{a})(\underline{b}'\Sigma_{22}\underline{b})}}$$

$$= \frac{\underline{c}'\Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1/2}\underline{d}}{\sqrt{(\underline{c}'\underline{c})(\underline{d}'\underline{d})}} \qquad \underline{a} = \Sigma_{11}^{\frac{-1}{2}}\underline{c} \quad \text{and} \quad \underline{b} = \Sigma_{22}^{\frac{-1}{2}}\underline{d}$$

$$= \frac{\underline{h}'\underline{d}}{\sqrt{(\underline{c}'\underline{c})(\underline{d}'\underline{d})}} \qquad \text{where } \underline{h}' = \underline{c}'\Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1/2} = \underline{c}'A$$

$$\leq \frac{\sqrt{(\underline{h}'\underline{h})(\underline{d}'\underline{d})}}{\sqrt{(\underline{c}'\underline{c})(\underline{d}'\underline{d})}} \qquad \text{by Cauchy-Schwartz (CS) inequality}$$

Proof…1

Max corr$\left(\underline{a}'\underline{X}, \underline{b}'\underline{Y}\right)$

$$= \frac{\sqrt{(\underline{h}'\underline{h})(\underline{d}'\underline{d})}}{\sqrt{(\underline{c}'\underline{c})(\underline{d}'\underline{d})}}$$

C.S.I. holds if $\underline{d} \propto \underline{h} = \Sigma_{22}^{-\frac{1}{2}}\Sigma_{21}\Sigma_{11}^{-\frac{1}{2}}\underline{c} = A'\underline{c}$

$$= \frac{\sqrt{\underline{c}'(AA')\underline{c}(\underline{d}'\underline{d})}}{\sqrt{(\underline{c}'\underline{c})(\underline{d}'\underline{d})}}$$

as $\underline{h}'\underline{h} = \underline{c}'(AA')\underline{c} \le {\rho_1^*}^2(\underline{c}'\underline{c})$

$$= \frac{\rho_1^*\sqrt{(\underline{c}'\underline{c})(\underline{d}'\underline{d})}}{\sqrt{(\underline{c}'\underline{c})(\underline{d}'\underline{d})}}$$

$$= \rho_1^*$$

$\underline{c}'(AA')\underline{c} = \rho_1^*(\underline{c}'\underline{c})$ if $\underline{c} = \underline{e}_1$

Proof…2

Thus

$$\operatorname{corr}\left(\underline{a}'\underline{X}, \underline{b}'\underline{Y}\right) = \rho_1^*$$

if $\underline{c} = \underline{e}_1$ and $\underline{d} \propto \underline{h} = \Sigma_{22}^{-1/2}\Sigma_{21}\Sigma_{11}^{-1/2}\underline{c}$

if $\Sigma_{11}^{\frac{1}{2}}\underline{a} = \underline{e}_1$ and $\Sigma_{22}^{\frac{1}{2}}\underline{b} = \Sigma_{22}^{-1/2}\Sigma_{21}\Sigma_{11}^{-1/2}\underline{e}_1$

if $\underline{a} = \Sigma_{11}^{-\frac{1}{2}}\underline{e}_1$ and $\underline{b} = \Sigma_{22}^{-1/2}\Sigma_{22}^{-1/2}\Sigma_{21}\Sigma_{11}^{-1/2}\underline{e}_1$

if $\underline{a} = \Sigma_{11}^{-\frac{1}{2}}\underline{e}_1$ and $\underline{b} = \Sigma_{22}^{-1/2}\underline{f}_1$

(as $\Sigma_{22}^{-1/2}\Sigma_{21}\Sigma_{11}^{-1/2}\underline{e}_1 = \underline{f}_1$)

Proof…3

Note:

$\underline{e}_1$ is eigen vector of AA'

$\Rightarrow AA'\underline{e}_1 = \rho_1^{*2}\underline{e}_1$

$\Rightarrow \Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1/2}\underline{e}_1 = \rho_1^{*2}\underline{e}_1$

$\Rightarrow \left(\boldsymbol{\Sigma}_{22}^{-\frac{1}{2}}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-\frac{1}{2}}\right)\Sigma_{11}^{-1/2}\Sigma_{12}\boldsymbol{\Sigma}_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1/2}\underline{e}_1 = \rho_1^{*2}\left(\boldsymbol{\Sigma}_{22}^{-\frac{1}{2}}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-\frac{1}{2}}\right)\underline{e}_1$

$\Rightarrow \left(\Sigma_{22}^{-\frac{1}{2}}\Sigma_{21}\Sigma_{11}^{-\frac{1}{2}}\right)\Sigma_{11}^{-\frac{1}{2}}\Sigma_{12}\boldsymbol{\Sigma}_{22}^{-\frac{1}{2}}\left(\Sigma_{22}^{-\frac{1}{2}}\Sigma_{21}\Sigma_{11}^{-\frac{1}{2}}\right)\underline{e}_1 = \rho_1^{*2}\left(\boldsymbol{\Sigma}_{22}^{-\frac{1}{2}}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-\frac{1}{2}}\right)\underline{e}_1$

$\Rightarrow \Sigma_{22}^{-1/2}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1/2}\underline{f}_1 = \rho_1^{*2}\underline{f}_1$ where $\underline{\boldsymbol{f}}_1 = \left(\boldsymbol{\Sigma}_{22}^{-\frac{1}{2}}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-\frac{1}{2}}\right)\underline{\boldsymbol{e}}_1$

$\Rightarrow \underline{f}_1$ is eigen vector of $\Sigma_{22}^{-1/2}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1/2} = A'A$ corr. to ρ_1^{*2}

Steps to perform canonical analysis

- Partition Σ as $\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$ by deciding p and q

- Calculate $A = \Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1/2}$ and hence

$$AA' = \Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1/2} \text{ is } p \times p$$

$$A'A = \Sigma_{22}^{-1/2}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1/2} \text{ is } q \times q$$

Steps to perform canonical analysis

- Calculate eigen values $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$ and eigen vectors $\underline{e}_1, \underline{e}_2, \ldots, \underline{e}_p$ of AA'
- Calculate eigen values $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_q \geq 0$ and eigen vectors $\underline{f}_1, \underline{f}_2, \ldots, \underline{f}_q$ of $A'A$
- Observe that when $p \leq q$ AA' and $A'A$ have same eigen values i.e. $\lambda_i = \mu_i, i = 1,2, \ldots, p$

Steps to perform canonical analysis

- Calculate the k^{th}pair of canonical variables (U_k, V_k) as

$$U_k = \underline{e}_k' \Sigma_{11}^{-\frac{1}{2}} \underline{X} \text{ and } V_k = \underline{f}_k' \Sigma_{22}^{-\frac{1}{2}} \underline{Y}$$

- Canonical correlation for the k^{th} pair of canonical variables is

$$can\ corr(U_k, V_k) = \sqrt{\lambda_k}, \quad k = 1,2, \ldots, p$$

Sample based CCA

- To find the sample pairs of canonical variables and correlation between them use the same algorithm on MLE $S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}$ of $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$.

Example:

Obtain the CC and CV with $p = q = 2$ and

$$\Sigma = \begin{pmatrix} 100 & 0 & 0 & 0 \\ 0 & 1 & 0.95 & 0 \\ 0 & 0.95 & 1 & 0 \\ 0 & 0 & 0 & 100 \end{pmatrix},$$

$$\Sigma_{11} = \begin{pmatrix} 100 & 0 \\ 0 & 1 \end{pmatrix},$$

$$\Sigma_{22} = \begin{pmatrix} 1 & 0 \\ 0 & 100 \end{pmatrix}, \qquad \Sigma_{12} = \begin{pmatrix} 0 & 0 \\ 0.95 & 0 \end{pmatrix} = \Sigma_{21}'$$

Example:

$$\mathrm{R}_1 = \begin{pmatrix} 0 & 0 \\ 0 & (0.95)^2 \end{pmatrix},$$

$$\underline{e}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$U_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 1/10 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = X_2$$

$$\underline{f}_1 = X_3$$

$$\rho_1 = 0.95$$

$$(U_1, V_1) = (X_2, X_3)$$

$$\mathrm{corr}(U_1, V_1) = 0.95$$

CCA with R : An example

- Consider data which consists of two datasets where each dataset represents measurements in some specific (64-65) river localities in the Czech republic.

i) The first dataset contains measurements of biologial metrics for each locality (17 different metrics and taxons)

ii) The second dataset contains measurements on chemical concentrations and values at the same localities (7 covariates recorded).

The objective is to somehow relate both datasets in order to explain what bio metrics can correlate which chemical concentrations.

R code

```
install.packages("CCA")
rm(list = ls())
bioData <- read.csv("http://msekce.karlin.mff.cuni.cz/~maciak/NMST539/bioData.csv", header = T)
chemData <- read.csv("http://msekce.karlin.mff.cuni.cz/~maciak/NMST539/chemData.csv", header = T)
head(bioData)
head(chemData)
ind <- match(chemData[,1], bioData[,1])
data <- data.frame(bioData[ind, ], chemData[, 2:8])
X <- data[,2:9]
Y <- data[,19:25]
library("CCA")
correl <- matcor(X, Y )
img.matcor(correl, type = 2)
```

R code…

```
cc1 <- cancor(X, Y)  ### function from standard R instalation
cc1$cor
cc1$xcoef  ### function from standard R instalation
par(mfrow = c(1,2))
barplot(cc1$cor, main = "Canonical correlations for 'cancor()'", col = "gray")
cc1$xcoef  ### function from standard R instalation
cc1$ycoef

cc2 <- cc(X, Y)      ### function for the R package 'CCA'
cc2$cor
barplot(cc1$cor, main = "Canonical correlations for 'cancor()'", col = "gray")
cc2$xcoef
cc2$ycoef
plt.cc(cc2, var.label = TRUE, ind.names = data[,1])
```

CCA in Matlab

[A, B, r, U, V] = canoncorr(x, y)

x, y : set of variables in the form of matrices
Each row is an observation
Each column is an attribute/feature

A, B : Matrices containing the correlation coefficient

r : Column matrix containing the canonical correlations (Successively decreasing)

U, V : Canonical variates/basis vectors for A,B respectively

Properties of Canonical Correlations

- Canonical correlations are invariant to changes of scale on either the response variables (*y's*) and the explanatory variables (*x's*).

  In other words, changing the scale of measurement of the 2 sets of variables of interest in the analysis, for instance, from inches to centimetres does not interfere with the canonical correlations that follow.

Properties of Canonical Correlations…

- The first canonical correlation (ρ_1) is the maximum correlation between the linear functions of Y and X.
- In other words, ρ_1 exceeds the (absolute) simple correlation between any Y and any X or the multiple correlation between any Y and all the X's or between any X and all the Y's.

CCA, PCA and Multiple correlation analysis

- The main purpose of the canonical correlation approach is the exploration of sample correlations between two sets of quantitative variables observed on the same experimental units.
- On the other hand PCA method deals with one data set only and it tries to reduce the overall dimensionality of the dataset through having a few linear combinations of the initial variables.

CCA, PCA and Multiple correlation analysis

- CCA is the technique which is an extension of multiple correlation analysis and is often applicable in the same situations in which multivariate regression analysis methods would be applicable.