Lecture 10 Canonical Correlation Analysis

Shiwei Lan[1]

[1]School of Mathematical and Statistical Sciences
Arizona State University

STP533 Multivariate Analysis
Spring 2025

Introduction

- Canonical correlation analysis seeks to *identify and quantify the associations between two sets of variables*.
- It started with H. Hotelling trying to relating the arithmetic speed and arithmetic power to the reading speed and reading power.
- The *idea* is to first determine a pair of linear combinations with the largest correlation, and then to obtain the next pair with the largest correlation among all pairs uncorrelated from previous pairs.
- The pairs of linear combinations are called *canonical variables*, and their correlations are called *canonical correlations*.
- The canonical correlation analysis is a technique to summarize high-dimensional relationship between two sets of variables into a few pairs of canonical variables.

Table of Contents

Canonical Variates and Canonical Correlations

- Consider two groups of variables, represented by $\mathbf{X}^{(1)} \in \mathbb{R}^p$ for p variables and $\mathbf{X}^{(2)} \in \mathbb{R}^q$ for q variables. Assume $p \leq q$.
- We assue

$$\mathrm{E}(\mathbf{X}^{(1)}) = \boldsymbol{\mu}^{(1)}, \quad \mathrm{E}(\mathbf{X}^{(2)}) = \boldsymbol{\mu}^{(2)}$$
$$\mathrm{Cov}(\mathbf{X}^{(1)}) = \boldsymbol{\Sigma}_{11}, \quad \mathrm{Cov}(\mathbf{X}^{(2)}) = \boldsymbol{\Sigma}_{22}, \quad \mathrm{Cov}(\mathbf{X}^{(1)}, \mathbf{X}^{(2)}) = \boldsymbol{\Sigma}_{12} = \boldsymbol{\Sigma}_{21}^T$$

- We denote

$$\mathbf{X}_{(p+q)\times 1} = \begin{bmatrix} \mathbf{X}^{(1)} \\ \mathbf{X}^{(2)} \end{bmatrix}, \ \boldsymbol{\mu}_{(p+q)\times 1} = \begin{bmatrix} \boldsymbol{\mu}^{(1)} \\ \boldsymbol{\mu}^{(2)} \end{bmatrix}, \ \boldsymbol{\Sigma}_{(p+q)\times(p+q)} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}.$$

Canonical Variates and Canonical Correlations

- We want to summarize $\mathbf{\Sigma}_{12}$ in a few chosen covariances (correlations).
- Suppose we have linear combinations $U = \mathbf{a}^T\mathbf{X}^{(1)}, \quad V = \mathbf{b}^T\mathbf{X}^{(2)}$. Then

$$\mathrm{Var}(U) = \mathbf{a}^T\mathbf{\Sigma}_{11}\mathbf{a},\ \mathrm{Var}(V) = \mathbf{b}^T\mathbf{\Sigma}_{22}\mathbf{b},\ \mathrm{Cov}(U, V) = \mathbf{a}^T\mathbf{\Sigma}_{12}\mathbf{b}$$

- We seek vectors $\mathbf{a} \in \mathbb{R}^p$ and $\mathbf{b} \in \mathbb{R}^q$ to maximize the correlation:

$$\max_{\mathbf{a},\mathbf{b}} \mathrm{Corr}(U, V) = \max_{\mathbf{a},\mathbf{b}} \frac{\mathbf{a}^T\mathbf{\Sigma}_{12}\mathbf{b}}{\sqrt{\mathbf{a}^T\mathbf{\Sigma}_{11}\mathbf{a}}\sqrt{\mathbf{b}^T\mathbf{\Sigma}_{22}\mathbf{b}}}$$

- We define the *first pair of canonical variables* as the pair of the linear combinations U_1, V_1 with unit variances that maximize the above correlation.
- The *k-th canonical variate pair*, is the linear combinations U_k, V_k with unit variances that maximize the correlation among pairs uncorrelated to previous $k-1$ canonical variable pairs.

Canonical Variates and Canonical Correlations

- Suppose $p \leq q$ and $\boldsymbol{\Sigma}$ has full rank. For the linear combinations $U = \mathbf{a}^T \mathbf{X}^{(1)}$, $V = \mathbf{b}^T \mathbf{X}^{(2)}$, we have

$$\max_{\mathbf{a},\mathbf{b}} \mathrm{Corr}(U, V) = \lambda_1^*$$

attained by $\mathbf{a}_1 = \boldsymbol{\Sigma}_{11}^{-\frac{1}{2}} \mathbf{v}_1$ and $\mathbf{b}_1 = \boldsymbol{\Sigma}_{22}^{-\frac{1}{2}} \mathbf{w}_1$.

- The k-th pair of canonical variates are $U_k = \mathbf{v}_k^T \boldsymbol{\Sigma}_{11}^{-\frac{1}{2}} \mathbf{X}^{(1)}$, $V_k = \mathbf{w}_k^T \boldsymbol{\Sigma}_{22}^{-\frac{1}{2}} \mathbf{X}^{(2)}$

$$\max_{\mathbf{a},\mathbf{b}} \mathrm{Corr}(U_k, V_k) = \lambda_k^*$$

where λ_k^{*2} is the k-th leading eigenvalue of $\boldsymbol{\Sigma}_{11}^{-\frac{1}{2}} \boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_{22}^{-1} \boldsymbol{\Sigma}_{21} \boldsymbol{\Sigma}_{11}^{-\frac{1}{2}}$ with $\mathbf{v}_k$ the associated eigenvector, and $\mathbf{w}_k$ is the k-th leading eigenvector of $\boldsymbol{\Sigma}_{22}^{-\frac{1}{2}} \boldsymbol{\Sigma}_{21} \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_{22}^{-\frac{1}{2}}$ whose corresponding eigenvalue is also λ_k^{*2}.

- Finally, we have

$$\begin{aligned} \mathrm{Var}(U_k) &= \mathrm{Var}(V_k) = 1 \\ \mathrm{Cov}(U_k, U_\ell) &= \mathrm{Corr}(U_k, U_\ell) = \delta_{k\ell} \\ \mathrm{Cov}(V_k, V_\ell) &= \mathrm{Corr}(V_k, V_\ell) = \delta_{k\ell} \\ \mathrm{Cov}(U_k, V_\ell) &= \mathrm{Corr}(U_k, V_\ell) = 0, \; k \neq \ell, \quad k, \ell = 1, \cdots, p. \end{aligned}$$

- For the standardized variables $\mathbf{Z}^{(1)}$ and $\mathbf{Z}^{(2)}$, the k-th canonical variates are $U_k = \mathbf{a}_k^T \mathbf{Z}^{(1)}$ and $V_k = \mathbf{b}_k^T \mathbf{Z}^{(2)}$ with $\mathbf{a}_k = \boldsymbol{P}_{11}^{-\frac{1}{2}} \mathbf{v}_k$ and $\mathbf{b}_k = \boldsymbol{P}_{22}^{-\frac{1}{2}} \mathbf{w}_k$ and

$$\max_{\mathbf{a},\mathbf{b}} \mathrm{Corr}(U_k, V_k) = \rho_k^*$$

where ρ_k^{*2} is the k-th leading eigenvalue of $\boldsymbol{P}_{11}^{-\frac{1}{2}} \boldsymbol{P}_{12} \boldsymbol{P}_{22}^{-1} \boldsymbol{P}_{21} \boldsymbol{P}_{11}^{-\frac{1}{2}}$ with $\mathbf{v}_k$ the associated eigenvector, and $\mathbf{w}_k$ is the k-th leading eigenvector of $\boldsymbol{P}_{22}^{-\frac{1}{2}} \boldsymbol{P}_{21} \boldsymbol{P}_{11}^{-1} \boldsymbol{P}_{12} \boldsymbol{P}_{22}^{-\frac{1}{2}}$ whose corresponding eigenvalue is also ρ_k^{*2}.

Example 10.1 (Calculating canonical variates and canonical correlations for standardized variables) Suppose $\mathbf{Z}^{(1)} = [Z_1^{(1)}, Z_2^{(1)}]'$ are standardized variables and $\mathbf{Z}^{(2)} = [Z_1^{(2)}, Z_2^{(2)}]'$ are also standardized variables. Let $\mathbf{Z} = [\mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}]'$ and

$$\text{Cov}(\mathbf{Z}) = \left[\begin{array}{c|c} \boldsymbol{\rho}_{11} & \boldsymbol{\rho}_{12} \\ \hline \boldsymbol{\rho}_{21} & \boldsymbol{\rho}_{22} \end{array}\right] = \left[\begin{array}{cc|cc} 1.0 & .4 & .5 & .6 \\ .4 & 1.0 & .3 & .4 \\ \hline .5 & .3 & 1.0 & .2 \\ .6 & .4 & .2 & 1.0 \end{array}\right]$$

Then

$$\boldsymbol{\rho}_{11}^{-1/2} = \begin{bmatrix} 1.0681 & -.2229 \\ -.2229 & 1.0681 \end{bmatrix}$$

$$\boldsymbol{\rho}_{22}^{-1} = \begin{bmatrix} 1.0417 & -.2083 \\ -.2083 & 1.0417 \end{bmatrix}$$

and

$$\boldsymbol{\rho}_{11}^{-1/2}\boldsymbol{\rho}_{12}\boldsymbol{\rho}_{22}^{-1}\boldsymbol{\rho}_{21}\boldsymbol{\rho}_{11}^{-1/2} = \begin{bmatrix} .4371 & .2178 \\ .2178 & .1096 \end{bmatrix}$$

The eigenvalues, ρ_1^{*2}, ρ_2^{*2}, of $\boldsymbol{\rho}_{11}^{-1/2}\boldsymbol{\rho}_{12}\boldsymbol{\rho}_{22}^{-1}\boldsymbol{\rho}_{21}\boldsymbol{\rho}_{11}^{-1/2}$ are obtained from

$$0 = \begin{vmatrix} .4371 - \lambda & .2178 \\ .2178 & .1096 - \lambda \end{vmatrix} = (.4371 - \lambda)(.1096 - \lambda) - (2.178)^2$$
$$= \lambda^2 - .5467\lambda + .0005$$

Table of Contents

Identifying the Canonical Variables

- Canonical variables are in general artificial and have no physical meaning. Often they are interpreted in terms of standardized variables.
- Let $\mathbf{A}_{p\times p} = [\mathbf{a}_1, \cdots, \mathbf{a}_p]^T$ and $\mathbf{B}_{q\times q} = [\mathbf{b}_1, \cdots, \mathbf{b}_q]^T$. Then we have vectors of canonical variables represented as $\mathbf{U}_{p\times 1} = \mathbf{A}\mathbf{X}^{(1)}$ and $\mathbf{V}_{q\times 1} = \mathbf{B}\mathbf{X}^{(2)}$, where we are interested in the first p canonical variables in $\mathbf{V}$.
- $\mathrm{Cov}(\mathbf{U}, \mathbf{X}^{(1)}) = \mathbf{A}\boldsymbol{\Sigma}_{11}$, hence $\rho_{\mathbf{U},\mathbf{X}^{(1)}} = \mathrm{Corr}(\mathbf{U}, \mathbf{X}^{(1)}) = \mathrm{Cov}(\mathbf{U}, \mathbf{D}_{\boldsymbol{\Sigma}_{11}}^{-\frac{1}{2}}\mathbf{X}^{(1)})$.
- Similar calculations yield

$$\rho_{\mathbf{U},\mathbf{X}^{(1)}} = \mathbf{A}\boldsymbol{\Sigma}_{11}\mathbf{D}_{\boldsymbol{\Sigma}_{11}}^{-\frac{1}{2}},\ \rho_{\mathbf{V},\mathbf{X}^{(2)}} = \mathbf{B}\boldsymbol{\Sigma}_{22}\mathbf{D}_{\boldsymbol{\Sigma}_{22}}^{-\frac{1}{2}},\ \rho_{\mathbf{U},\mathbf{X}^{(2)}} = \mathbf{A}\boldsymbol{\Sigma}_{12}\mathbf{D}_{\boldsymbol{\Sigma}_{22}}^{-\frac{1}{2}},\ \rho_{\mathbf{V},\mathbf{X}^{(1)}} = \mathbf{B}\boldsymbol{\Sigma}_{21}\mathbf{D}_{\boldsymbol{\Sigma}_{11}}^{-\frac{1}{2}}.$$

- Canonical variables for standardized variables are interpreted by correlations.

$$\rho_{\mathbf{U},\mathbf{Z}^{(1)}} = \mathbf{A}_{\mathbf{Z}}P_{11},\ \rho_{\mathbf{V},\mathbf{Z}^{(2)}} = \mathbf{B}_{\mathbf{Z}}P_{22},\ \rho_{\mathbf{U},\mathbf{Z}^{(2)}} = \mathbf{A}_{\mathbf{Z}}P_{12},\ \rho_{\mathbf{V},\mathbf{Z}^{(1)}} = \mathbf{B}_{\mathbf{Z}}P_{21}.$$

where $\mathbf{A}_{\mathbf{Z}} = \mathbf{A}\mathbf{D}_{\boldsymbol{\Sigma}_{11}}^{\frac{1}{2}}$, and $P_{11} = \mathbf{D}_{\boldsymbol{\Sigma}_{11}}^{-\frac{1}{2}}\boldsymbol{\Sigma}_{11}\mathbf{D}_{\boldsymbol{\Sigma}_{11}}^{-\frac{1}{2}}$.

Example 10.2 (Computing correlations between canonical variates and their component variables) Compute the correlations between the first pair of canonical variates and their component variables for the situation considered in Example 10.1.

The variables in Example 10.1 are already standardized, so equation (10-15) is applicable. For the standardized variables,

$$\boldsymbol{\rho}_{11} = \begin{bmatrix} 1.0 & .4 \\ .4 & 1.0 \end{bmatrix} \qquad \boldsymbol{\rho}_{22} = \begin{bmatrix} 1.0 & .2 \\ .2 & 1.0 \end{bmatrix}$$

Canonical Correlations as Generalizations

- If $p = q = 1$, then $|\mathrm{Corr}(X^{(1)}, X^{(2)})| = |\mathrm{Corr}(aX^{(1)}, bX^{(2)})|$ for all $a, b \neq 0$.
- For $\mathbf{a} = \mathbf{e}_i$ and $\mathbf{b} = \mathbf{e}_k$, we have

$$|\mathrm{Corr}(X_i^{(1)}, X_k^{(2)})| = |\mathrm{Corr}(\mathbf{a}^T\mathbf{X}^{(1)}, \mathbf{b}^T\mathbf{X}^{(2)})| \leq \max_{\mathbf{a},\mathbf{b}} \mathrm{Corr}(\mathbf{a}^T\mathbf{X}^{(1)}, \mathbf{b}^T\mathbf{X}^{(2)}) = \lambda_1^*$$

- The multiple correlation coefficient $\rho_{1,\mathbf{X}^{(2)}}$ is a special case when $p = 1$:

$$\rho_{1,\mathbf{X}^{(2)}} = \max_{\mathbf{b}} \mathrm{Corr}(X_1^{(1)}, \mathbf{b}^T\mathbf{X}^{(2)}) = \lambda_1^*$$

- Finally, we have canonical correlations as multiple correlation coefficients:

$$\rho_{U_k,\mathbf{X}^{(2)}} = \max_{\mathbf{b}} \mathrm{Corr}(U_k, \mathbf{b}^T\mathbf{X}^{(2)}) = \mathrm{Corr}(U_k, V_k) = \lambda_k^* = \rho_{\mathbf{X}^{(1)},V_k}, \quad k = 1, \cdots, p.$$

- Canonical variables are ment for summarizing correlations, not variation.

A Geometrical Interpretation

- Recall that the canonical variables are selected as $\mathbf{U} = \mathbf{A}\mathbf{X}^{(1)}$ such that $\mathrm{Cov}(\mathbf{U}) = \mathbf{I}$.
- From the previous discussion, we have

$$\mathbf{A} = \mathbf{V}^T \boldsymbol{\Sigma}_{11}^{-\frac{1}{2}} = \mathbf{V}^T \boldsymbol{\Gamma}_1 \boldsymbol{\Lambda}^{-\frac{1}{2}} \boldsymbol{\Gamma}_1^T$$

  where $\mathbf{V}$ is formed by principal eigenvectors of $\boldsymbol{\Sigma}_{11}^{-\frac{1}{2}}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-\frac{1}{2}}$, and $\boldsymbol{\Sigma}_{11} = \boldsymbol{\Gamma}_1\boldsymbol{\Lambda}\boldsymbol{\Gamma}_1^T$
- $\boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{\Gamma}_1^T$ has the effect of de-correlating $\mathbf{X}^{(1)}$: $\mathrm{Cov}(\boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{\Gamma}_1^T\mathbf{X}^{(1)}) = \mathbf{I}$.
- Therefore, $\mathbf{U} = \mathbf{V}^T\boldsymbol{\Gamma}_1\boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{\Gamma}_1^T\mathbf{X}^{(1)}$ can be interpreted as (1) a transformation of $\mathbf{X}^{(1)}$ to uncorrelated standardized PCs, followed by (2) a rigid rotation by Γ_1 and (3) another rotation $\mathbf{V}^T$ determined by $\boldsymbol{\Sigma}$.

Table of Contents

Sample Canonical Variates and Correlations

- Now we combine random samples $\mathbf{X}^{(1)}_{n\times p}$ and $\mathbf{X}^{(2)}_{n\times q}$ to $\mathbf{X}_{n\times(p+q)} = [\mathbf{X}^{(1)}, \mathbf{X}^{(2)}]$.
- Sample mean $\bar{\mathbf{x}}_{(p+q)\times 1} = \begin{bmatrix} \bar{\mathbf{x}}^{(1)} \\ \bar{\mathbf{x}}^{(2)} \end{bmatrix}$ and covariance $\mathbf{S}_{(p+q)\times(p+q)} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{bmatrix}$.
- The linear combinations $\hat{U} = \hat{\mathbf{a}}^T\mathbf{x}^{(1)}$, $\hat{V} = \hat{\mathbf{b}}^T\mathbf{x}^{(2)}$ have sample correlation

$$r_{\hat{U},\hat{V}} = \frac{\hat{\mathbf{a}}^T\mathbf{S}_{12}\hat{\mathbf{b}}}{\sqrt{\hat{\mathbf{a}}^T\mathbf{S}_{11}\hat{\mathbf{a}}}\sqrt{\hat{\mathbf{b}}^T\mathbf{S}_{22}\hat{\mathbf{b}}}}$$

- We define the *k-th sample canonical variates*, $\hat{U}_k, \hat{V}_k$ by maximizing the above sample correlation, similarly as k-th canonical variates.

Sample Canonical Variates and Canonical Correlations

- Suppose $p \le q$ and $\mathbf{S}$ has full rank. $\hat{\lambda}_k^{*2}$ is the k-th leading eigenvalue of $\mathbf{S}_{11}^{-\frac{1}{2}}\mathbf{S}_{12}\mathbf{S}_{22}^{-1}\mathbf{S}_{21}\mathbf{S}_{11}^{-\frac{1}{2}}$ with $\hat{\mathbf{v}}_k$ the associated eigenvector, and $\hat{\mathbf{w}}_k$ is the k-th leading eigenvector of $\mathbf{S}_{22}^{-\frac{1}{2}}\mathbf{S}_{21}\mathbf{S}_{11}^{-1}\mathbf{S}_{12}\mathbf{S}_{22}^{-\frac{1}{2}}$.
- The k-th pair of canonical variates are $\hat{U}_k = \hat{\mathbf{a}}_k^T\mathbf{x}^{(1)}$, $\hat{V}_k = \hat{\mathbf{b}}_k^T\mathbf{x}^{(2)}$

$$\max_{\mathbf{a},\mathbf{b}} r_{\hat{U},\hat{V}} = \hat{\lambda}_k^*, \quad \textit{with } \hat{\mathbf{a}}_k = \mathbf{S}_{11}^{-\frac{1}{2}}\hat{\mathbf{v}}_k, \quad \hat{\mathbf{b}}_k = \mathbf{S}_{22}^{-\frac{1}{2}}\hat{\mathbf{w}}_k$$

- We also have

$$r_{U_k,U_\ell} = r_{V_k,V_\ell} = \delta_{k\ell}$$
$$r_{U_k,V_\ell} = 0, \quad k \neq \ell, \quad k,\ell = 1,\cdots,p.$$

- Similarly, let $\hat{\mathbf{A}}_{p\times p} = [\hat{\mathbf{a}}_1,\cdots,\hat{\mathbf{a}}_p]^T$ and $\hat{\mathbf{B}}_{q\times q} = [\hat{\mathbf{b}}_1,\cdots,\hat{\mathbf{b}}_q]^T$. Then we have $\hat{\mathbf{U}}_{p\times 1} = \hat{\mathbf{A}}\mathbf{x}^{(1)}$ and $\hat{\mathbf{V}}_{q\times 1} = \hat{\mathbf{B}}\mathbf{x}^{(2)}$.

- Sample correlations between sample canonical variates and original variables:

$$\mathbf{R}_{\hat{\mathbf{U}},\mathbf{x}^{(1)}} = \hat{\mathbf{A}}\mathbf{S}_{11}\mathbf{D}_{\mathbf{S}_{11}}^{-\frac{1}{2}}, \ \mathbf{R}_{\hat{\mathbf{V}},\mathbf{x}^{(2)}} = \hat{\mathbf{B}}\mathbf{S}_{22}\mathbf{D}_{\mathbf{S}_{22}}^{-\frac{1}{2}}, \ \mathbf{R}_{\hat{\mathbf{U}},\mathbf{x}^{(2)}} = \hat{\mathbf{A}}\mathbf{S}_{12}\mathbf{D}_{\mathbf{S}_{22}}^{-\frac{1}{2}}, \ \mathbf{R}_{\hat{\mathbf{V}},\mathbf{x}^{(1)}} = \hat{\mathbf{B}}\mathbf{S}_{21}\mathbf{D}_{\mathbf{S}_{11}}^{-\frac{1}{2}}.$$

- Sample canonical variables for standardized observations $\mathbf{Z}$ are obtained

$$\hat{\mathbf{U}} = \hat{\mathbf{A}}_{\mathbf{Z}}\mathbf{z}^{(1)}, \quad \hat{\mathbf{V}} = \hat{\mathbf{B}}_{\mathbf{Z}}\mathbf{z}^{(2)}$$

  where $\hat{\mathbf{A}}_{\mathbf{Z}} = \hat{\mathbf{A}}\mathbf{D}_{\mathbf{S}_{11}}^{\frac{1}{2}}$, and $\hat{\mathbf{B}}_{\mathbf{Z}} = \hat{\mathbf{B}}\mathbf{D}_{\mathbf{S}_{22}}^{\frac{1}{2}}$.
- Similar sample correlations remain unchanged by standardization, i.e. $\mathbf{R}_{\hat{\mathbf{U}},\mathbf{x}^{(1)}} = \mathbf{R}_{\hat{\mathbf{U}},\mathbf{z}^{(1)}} = \hat{\mathbf{A}}_{\mathbf{Z}}\mathbf{R}_{11}$.

Example 10.4 (Canonical correlation analysis of the chicken-bone data) In Example 9.14, data consisting of bone and skull measurements of white leghorn fowl were described. From this example, the chicken-bone measurements for

$$\text{Head } (\mathbf{X}^{(1)}): \begin{cases} X_1^{(1)} = \text{skull length} \\ X_2^{(1)} = \text{skull breadth} \end{cases}$$

$$\text{Leg } (\mathbf{X}^{(2)}): \begin{cases} X_1^{(2)} = \text{femur length} \\ X_2^{(2)} = \text{tibia length} \end{cases}$$

Table of Contents

Large Sample Inferences

- When $\mathbf{\Sigma}_{12} = \mathbf{0}$, then $\mathrm{Cov}(\mathbf{a}^T\mathbf{X}^{(1)}, \mathbf{b}^T\mathbf{X}^{(2)}) = 0$ for all $\mathbf{a}$, $\mathbf{b}$. How to test it?
- Assume $\mathbf{X}_i \overset{iid}{\sim} N_{p+q}(\boldsymbol{\mu}, \mathbf{\Sigma})$. Consider the likelihood ratio test for $H_0 : \mathbf{\Sigma}_{12} = \mathbf{0}$ vs $H_1 : \mathbf{\Sigma}_{12} \neq 0$.
- The LRT statistic $-2\log\Lambda = n\log\left(\frac{|\mathbf{S}_{11}\mathbf{S}_{22}|}{|\mathbf{S}|}\right)$ leads to

$$-(n - 1 - \frac{1}{2}(p + q + 1))\log\prod_{i=1}^{p}(1 - \hat{\lambda}_i^{*2}) \dot{\sim} \chi^2(pq)$$

- We reject $H_0(\lambda_1^* = \cdots = \lambda_p^* = 0)$ at significance level α if the above test statistic is bigger than $\chi^2_{1-\alpha}(pq)$.
- When H_0 is rejected, we could further test the individual canonical correlations $H_0^k : \rho_1^* \neq 0, \cdots, \rho_k^* \neq 0, \lambda_{k+1}^* = \cdots = \lambda_p^* = 0$.

Example 10.8 (Testing the significance of the canonical correlations for the job satisfaction data) Test the significance of the canonical correlations exhibited by the job characteristics–job satisfaction data introduced in Example 10.5.

All the test statistics of immediate interest are summarized in the table on page 566. From Example 10.5, $n = 784$, $p = 5$, $q = 7$, $\widehat{\rho_1^*} = .55$, $\widehat{\rho_2^*} = .23$, $\widehat{\rho_3^*} = .12$, $\widehat{\rho_4^*} = .08$, and $\widehat{\rho_5^*} = .05$.

Assuming multivariate normal data, we find that the first two canonical correlations, ρ_1^* and ρ_2^*, appear to be nonzero, although with the very large sample size, small deviations from zero will show up as statistically significant. From a practical point of view, the second (and subsequent) sample canonical correlations can probably be ignored, since (1) they are reasonably small in magnitude and (2) the corresponding canonical variates explain *very* little of the sample variation in the variable sets $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$. ■

Canonical Correlation Analysis Examples

- Canonical correlation analysis can be done using `R` packages `CCA` and `CCP`.
- Here are some more exemplary analyses done in `R`:
  - `https://stats.oarc.ucla.edu/r/dae/canonical-correlation-analysis/`
  - `https://medium.com/@heyamit10/a-comprehensive-guide-to-canonical-correlation-analysis-in-r-8904`
  - `https://www.karlin.mff.cuni.cz/~maciak/NMST539/cvicenie11.html`