1 Question 1 (FA)

Given the standardized variables Z_1, Z_2, and Z_3 with the correlation matrix R,

$$R = \begin{bmatrix} 1 & 0.63 & 0.45 \\ 0.63 & 1 & 0.35 \\ 0.45 & 0.35 & 1 \end{bmatrix}$$

can be generated by the following one factor model,

$$\begin{cases} Z_1 = 0.9F_1 + \epsilon_1 \\ Z_2 = 0.7F_1 + \epsilon_2 \\ Z_3 = 0.5F_1 + \epsilon_3 \end{cases}$$

where $E(F_1) = 0, Var(F_1) = 1, E(\epsilon) = 0$, and $Cov(\epsilon, F_1) = 0$. Prove that

$$Cov(\epsilon) = \Sigma_\epsilon = \begin{bmatrix} 0.19 & 0 & 0 \\ 0 & 0.51 & 0 \\ 0 & 0 & 0.75 \end{bmatrix}.$$

2 Question 2 (CCA, Question 12.2 in textbook 1)

Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_q > 0$ be the positive roots of

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

where Σ_{11} and Σ_{22} are $q \times q$ nonsingular matrices.
(a) What is the rank of Σ_{12}?
(b) Write $\prod_{i=1}^{q} \lambda_i^2$ as the determinant of a rational function of Σ_{11}, Σ_{12}, Σ_{21} and Σ_{22}. Justify your answer.

3 Question 3 (CCA, Question 12.4 in textbook 1)

Let

$$X^{(1)} = AZ + Y^{(1)},$$
$$X^{(2)} = BZ + Y^{(2)},$$

where $Y^{(1)}, Y^{(2)}, Z$ are independent with mean zero and covariance matrices I with appropriate dimensionalities. Let $A = (a_1, \ldots, a_k), B = (b_1, \ldots, b_k)$, and suppose that $A^T A, B^T B$ are diagonal with positive diagonal elements. Show that the canonical variables for nonzero canonical correlations are proportional to $a_i^T X^{(1)}, b_i^T X^{(2)}$.

1 Question 1 (FA)

Given the standardized variables Z_1, Z_2, and Z_3 with the correlation matrix R,

$$R = \begin{bmatrix} 1 & 0.63 & 0.45 \\ 0.63 & 1 & 0.35 \\ 0.45 & 0.35 & 1 \end{bmatrix}$$

can be generated by the following one factor model,

$$\begin{cases} Z_1 = 0.9F_1 + \epsilon_1 \\ Z_2 = 0.7F_1 + \epsilon_2 \\ Z_3 = 0.5F_1 + \epsilon_3 \end{cases}$$

where $E(F_1) = 0, Var(F_1) = 1, E(\epsilon) = 0$, and $Cov(\epsilon, F_1) = 0$. Prove that

$$Cov(\epsilon) = \Sigma_\epsilon = \begin{bmatrix} 0.19 & 0 & 0 \\ 0 & 0.51 & 0 \\ 0 & 0 & 0.75 \end{bmatrix}.$$

1.1 Solution

Proof.

$$R = \begin{pmatrix} 1 & r_{12} & r_{13} \\ r_{12} & 1 & r_{23} \\ r_{13} & r_{23} & 1 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} (\lambda_1, \lambda_2, \lambda_3) + \begin{pmatrix} \psi_1 & 0 & 0 \\ 0 & \psi_2 & 0 \\ 0 & 0 & \psi_3 \end{pmatrix}.$$

$$\Longrightarrow \begin{pmatrix} \psi_1 & 0 & 0 \\ 0 & \psi_2 & 0 \\ 0 & 0 & \psi_3 \end{pmatrix} = R - \begin{pmatrix} 0.9 \\ 0.7 \\ 0.5 \end{pmatrix} (0.9, 0.7, 0.5)\,.$$

We have $\psi_1 = 0.19$, $\psi_2 = 0.51$, and $\psi_3 = 0.75$ and $Cov(\epsilon) = \Sigma_\epsilon = diag(\psi_1, \psi_2, \psi_3)$. □

2 Canonical Correlation Analysis (CCA)

- The purpose of CCA?
  - In PCA technique, to find the new component with maximum variance.
  - In CCA technique, study correlation between two sets of variables (or components).

Prob (1)

$$\begin{aligned} \max \quad & Cov(U, V) \\ s.t. \quad & Var(U) = 1 \\ & Var(V) = 1. \end{aligned}$$

- If $Var(U) = 1$ and $Var(V) = 1$, then $Cov(U, V) = Cor(U, V)$, that is covariance=correlation.

$$X = \begin{pmatrix} X^{(1)} \\ X^{(2)} \end{pmatrix} \quad \Sigma_X = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

- We can just blockwise Σ_X directly, then we can prove that $\Sigma_{X^{(1)}} = \Sigma_{11}$ and $\Sigma_{X^{(2)}} = \Sigma_{22}$.
- $U = \alpha^T X^{(1)}$, $V = \gamma^T X^{(2)}$.
- $Var(U) = Var(\alpha^T X^{(1)}) = \alpha^T \Sigma_{11} \alpha$.
- $Var(V) = Var(\gamma^T X^{(2)}) = \gamma^T \Sigma_{22} \gamma$.
- $Cov(U, V) = Cov(\alpha^T X^{(1)}, \gamma^T X^{(2)}) = \alpha^T \Sigma_{12} \gamma$. $\Sigma_{12}^T = \Sigma_{21}$, Σ_X is symmetric.
- Rewrite Prob(1):

$$\begin{aligned} \max \quad & \alpha^T \Sigma_{12} \gamma \\ s.t. \quad & \alpha^T \Sigma_{11} \alpha = 1 \\ & \gamma^T \Sigma_{22} \gamma = 1. \end{aligned}$$

- $\phi = \alpha^T \Sigma_{12} \gamma - \frac{1}{2}\lambda(\alpha^T \Sigma_{11} \alpha - 1) - \frac{1}{2}\mu(\gamma^T \Sigma_{22} \gamma - 1)$, where λ and μ are Lagrangian multipliers.
- $\frac{\partial \phi}{\partial \alpha} = 0$ and $\frac{\partial \phi}{\partial \gamma} = 0$.
- The solution of Prob (1) is $U^{(1)}$ and $V^{(1)}$.
- $Cov(U, U^{(1)}) = 0$ means that U and $U^{(1)}$ are uncorrelated.

$$\begin{cases} \frac{\partial \phi}{\partial \alpha} = \Sigma_{12}\gamma - \lambda \Sigma_{11} \alpha = 0. \\ \frac{\partial \phi}{\partial \gamma} = \Sigma_{12}^T \alpha - \mu \Sigma_{22} \gamma = 0. \end{cases} \Longrightarrow \begin{cases} \Sigma_{11}^{-1}\Sigma_{12}\gamma - \lambda \alpha = 0. \\ \Sigma_{22}^{-1}\Sigma_{12}^T \alpha - \mu \gamma = 0. \end{cases} \tag{1}$$

- Now we prove that $\lambda = \mu$.

$$\begin{cases} \Sigma_{12}\gamma - \lambda \Sigma_{11} \alpha = 0. \\ \Sigma_{12}^T \alpha - \mu \Sigma_{22} \gamma = 0. \end{cases} \Longrightarrow \begin{cases} \alpha^T \Sigma_{12}\gamma - \lambda \alpha^T \Sigma_{11} \alpha = 0. \\ \gamma^T \Sigma_{12}^T \alpha - \mu \gamma^T \Sigma_{22} \gamma = 0. \end{cases} \Longrightarrow \begin{cases} \alpha^T \Sigma_{12}\gamma = \lambda. \\ \gamma^T \Sigma_{12}^T \alpha = \mu. \end{cases}$$

  - According to the constraints, we have $\alpha^T \Sigma_{11} \alpha = 1$ and $\gamma^T \Sigma_{22} \gamma = 1$.
  - $\lambda = \alpha^T \Sigma_{12} \gamma = \lambda^T = (\alpha^T \Sigma_{12} \gamma)^T = \gamma^T \Sigma_{12}^T \alpha = \mu$.

$$\begin{aligned} \begin{pmatrix} \Sigma_{11}^{-1} & 0 \\ 0 & \Sigma_{22}^{-1} \end{pmatrix} \begin{pmatrix} 0 & \Sigma_{12} \\ \Sigma_{21} & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \gamma \end{pmatrix} &= \begin{pmatrix} 0 & \Sigma_{11}^{-1}\Sigma_{12} \\ \Sigma_{22}^{-1}\Sigma_{21} & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \gamma \end{pmatrix} \\ &= \begin{pmatrix} \Sigma_{11}^{-1}\Sigma_{12}\gamma \\ \Sigma_{22}^{-1}\Sigma_{21}\alpha \end{pmatrix} = \lambda \begin{pmatrix} \alpha \\ \gamma \end{pmatrix}. \end{aligned}$$

- Thus, λ is a eigenvalue of A, where

$$A = \begin{pmatrix} \Sigma_{11}^{-1} & 0 \\ 0 & \Sigma_{22}^{-1} \end{pmatrix} \begin{pmatrix} 0 & \Sigma_{12} \\ \Sigma_{21} & 0 \end{pmatrix}$$

- Note that $\Sigma_{21} = \Sigma_{12}^T$.
- $\max Cov(U, V) = \max \alpha^T \Sigma_{12} \gamma = \max \lambda$.

3 Question 2 (CCA, Question 12.2 in textbook 1)

Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_q > 0$ be the positive roots of

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

where Σ_{11} and Σ_{22} are $q \times q$ nonsingular matrices.
(a) What is the rank of Σ_{12}?
(b) Write $\prod_{i=1}^{q} \lambda_i^2$ as the determinant of a rational function of Σ_{11}, Σ_{12}, Σ_{21} and Σ_{22}. Justify your answer.

3.1 Solution

(a) λ_i also the eigenvalues of

$$A = \begin{pmatrix} \Sigma_{11}^{-1} & 0 \\ 0 & \Sigma_{22}^{-1} \end{pmatrix} \begin{pmatrix} 0 & \Sigma_{12} \\ \Sigma_{21} & 0 \end{pmatrix}$$

. Since $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_q > 0$, A is nonsingular. Thus, we have Σ_{12} is full rank matrix.
(b)

$$\begin{aligned} &\det(\lambda^2\Sigma_{11}\Sigma_{22} - \Sigma_{12}\Sigma_{21}) = 0 \\ \Longrightarrow &\det(\lambda^2 I - \Sigma_{22}^{-1}\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{21}) = \det(\Sigma_{22}^{-1}\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{21} - \lambda^2 I) = 0. \end{aligned}$$

Thus, λ_i^2 is eigenvalues of matrix $\Sigma_{22}^{-1}\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{21}$.

$$\prod_{i=1}^{q} \lambda_i^2 = \det\left(\Sigma_{22}^{-1}\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{21}\right).$$

4 Question 3 (CCA, Question 12.4 in textbook 1)

Let

$$\begin{aligned} X^{(1)} &= AZ + Y^{(1)}, \\ X^{(2)} &= BZ + Y^{(2)}, \end{aligned}$$

where $Y^{(1)}, Y^{(2)}, Z$ are independent with mean zero and covariance matrices I with appropriate dimensionalities. Let $A = (a_1, \ldots, a_k), B = (b_1, \ldots, b_k)$, and suppose that $A^T A, B^T B$ are diagonal with positive diagonal elements. Show that the canonical variables for nonzero canonical correlations are proportional to $a_i^T X^{(1)}, b_i^T X^{(2)}$.

4.1 Reform the question in a mathematical form

Consider

$$X = \begin{pmatrix} X^{(1)} \\ X^{(2)} \end{pmatrix}$$

$$\mathbb{E}(X^{(1)}) = A\mathbb{E}(Z) + \mathbb{E}(Y^{(1)}) = 0,$$
$$\mathbb{E}(X^{(2)}) = B\mathbb{E}(Z) + \mathbb{E}(Y^{(2)}) = 0.$$

$$\begin{aligned}
\Sigma_{11} = Cov(X^{(1)}) &= \mathbb{E}(X^{(1)} - \mathbb{E}X^{(1)})(X^{(1)} - \mathbb{E}X^{(1)})^T = \mathbb{E}(AZ + Y^{(1)})(AZ + Y^{(1)})^T \\
&= ACov(Z)A^T + 2ACov(Z, Y^{(1)}) + Cov(Y^{(1)}) \\
&= AA^T + I.
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
&\Sigma_{22} = Cov(X^{(2)}) = BB^T + I, \\
&\Sigma_{12} = AB^T, \\
&\Sigma_{21} = BA^T, \text{thus we have } \Sigma_{21} = \Sigma_{12}^T.
\end{aligned}$$

Recall CCA, the problem is to solve the following equation:

$$\begin{pmatrix} 0 & \Sigma_{11}^{-1}\Sigma_{12} \\ \Sigma_{22}^{-1}\Sigma_{21} & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \gamma \end{pmatrix} = \begin{pmatrix} 0 & (AA^T + I)^{-1}AB^T \\ (BB^T + I)^{-1}BA^T & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \gamma \end{pmatrix} = \lambda \begin{pmatrix} \alpha \\ \gamma \end{pmatrix},$$

where $A = (a_1, \ldots, a_k), B = (b_1, \ldots, b_k)$, and suppose that A^TA, B^TB are diagonal with positive diagonal elements. Recall that the canonical variables are $U^{(i)} = (\alpha^{(i)})^T X^{(1)}$ and $V^{(i)} = (\gamma^{(i)})^T X^{(2)}$. We need to prove that $\alpha^{(i)}$ is proportional to a_i and $\gamma^{(i)}$ is proportional to b_i.

4.2 Solution

Hint: $(I + AA^T)^{-1} = I - A(I + A^TA)^{-1}A^T$ and $(I + A^TA)^{-1} = I - (I + A^TA)^{-1}A^TA$ for any matrix A.

We first prove this hint by showing that $(I + AA^T)[I - A(I + A^TA)^{-1}A^T] = I$ and $(I + A^TA)[I - (I + A^TA)^{-1}A^TA] = I$.

Proof.

$$\begin{aligned}
&(I + AA^T)[I - A(I + A^TA)^{-1}A^T] \\
&= (I + AA^T) - (I + AA^T)A(I + A^TA)^{-1}A^T \\
&= (I + AA^T) - (A + AA^TA)(I + A^TA)^{-1}A^T \\
&= (I + AA^T) - A(I + A^TA)(I + A^TA)^{-1}A^T \\
&= (I + AA^T) - AA^T = I, \\
&(I + A^TA)[I - (I + A^TA)^{-1}A^TA] = (I + A^TA) - A^TA = I.
\end{aligned}$$

According to this hint, we have

$$\begin{aligned}
&(AA^T + I)^{-1}AB^T \\
&= [I - A(I + A^TA)^{-1}A^T]AB^T \\
&= AB^T - A(I + A^TA)^{-1}A^TAB^T \\
&= A[I - (I + A^TA)^{-1}A^TA]B^T \\
&= A(I + A^TA)^{-1}B^T
\end{aligned}$$

Similarly, we have $(BB^T + I)^{-1}BA^T = B(I + B^T B)^{-1}A^T$.

$$\lambda \begin{pmatrix} \alpha \\ \gamma \end{pmatrix} = \begin{pmatrix} 0 & (AA^T + I)^{-1}AB^T \\ (BB^T + I)^{-1}BA^T & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \gamma \end{pmatrix},$$

That is,

$$\begin{cases} \lambda\alpha = (AA^T + I)^{-1}AB^T\gamma = A(I + A^T A)^{-1}B^T\gamma, \\ \lambda\gamma = (BB^T + I)^{-1}BA^T\alpha = B(I + B^T B)^{-1}A^T\alpha. \end{cases}$$

$\Longrightarrow \lambda^2\alpha = \lambda(\lambda\alpha) = A(I + A^T A)^{-1}B^T(\lambda\gamma) = A(I + A^T A)^{-1}B^T B(I + B^T B)^{-1}A^T\alpha,$ multiply A^T

$\Longrightarrow \lambda^2 A^T\alpha = A^T A(I + A^T A)^{-1}B^T B(I + B^T B)^{-1}A^T\alpha.$

- Since $A^T A = diag(a_{11}, \dots, a_{kk})$, that is $A^T(a_1, \dots, a_k) = diag(a_{11}, \dots, a_{kk})$, we have $A^T a_i = (0, \dots, 0, a_{ii}, 0, \dots, 0)^T$.
- Since $A^T A, B^T B$ are diagonal, $A^T A(I + A^T A)^{-1}B^T B(I + B^T B)^{-1}$ which is denoted by C is diagonal.
- Since $C(A^T\alpha) = \lambda^2(A^T\alpha)$ with diagonal matrix C and constant λ, all diagonal elements of C must be the same (in fact must be λ^2).
- Let $D = (I + A^T A)^{-1}B^T B(I + B^T B)^{-1} = diag(d_{11}, \dots, d_{kk})$.
- $\lambda^2\alpha = ADA^T\alpha = (a_1, \dots, a_k)diag(d_{11}, \dots, d_{kk})A^T\alpha = (d_{11}a_1, \dots, d_{kk}a_k)A^T\alpha$.
- The ith term of $\lambda^2\alpha$, that is $\lambda^2\alpha_i$, should be $[d_{11}a_1a_1^T + \dots + d_{kk}a_ka_k^T]_i\alpha$.
- $[d_{11}a_1a_1^T + \dots + d_{kk}a_ka_k^T]_{i,j}\alpha_j = 0$ for $i \neq j$.
- $A^T\alpha = (0, \dots, 0, t, 0, \dots, 0)^T$.

$\square$