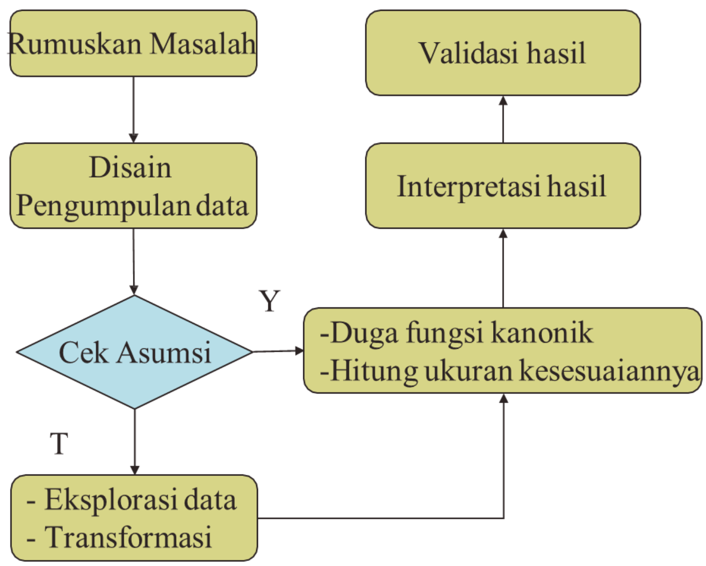
MODUL TUTORIAL ANALISIS MULTIVARIAT
ANALISIS KORELASI KANONIK

1. Pendahuluan

Analisis Korelasi Kanonik (*Canonical Correlation Analysis*/CCA) merupakan pengembangan dari analisis korelasi sederhana yang tidak hanya mengukur hubungan antara dua variabel tunggal, tetapi antara dua kelompok variabel secara simultan. Melalui pendekatan ini, diperoleh pasangan kombinasi linier (kanonik) dari masing-masing himpunan variabel yang memiliki korelasi maksimum satu sama lain. Analisis korelasi kanonik banyak digunakan di berbagai bidang, seperti psikologi, ekonomi, pendidikan, biostatistik, dan ilmu sosial lainnya, karena mampu memberikan pemahaman yang lebih komprehensif tentang hubungan multidimensional antar konsep yang kompleks. Tahapan analisis korelasi kanonik terlihat pada Gambar 1.



Gambar 1 Tahapan Analisis Korelasi Kanonik

2. Asumsi

Beberapa asumsi yang harus diperhatikan dalam analisis korelasi kanonik yaitu:

- Linieritas.
- Tidak terjadi multikolinieritas.
- Multivariate normal (normal ganda).

3. Pendugaan Fungsi Kanonik

Misal, gugus peubah dependen Y1, Y2, …, Yp dan gugus peubah independen X1, X2, …, Xq . Misalkan, karakteristik dari vektor peubah acak X dan Y adalah sebagai berikut:

$$E(X) = \mu_x$$
$$E(Y) = \mu_y$$
$$Cov(X,Y) = \Sigma_{xy} = (\Sigma_{yx} \quad)'$$
$$Cov(Y) = \Sigma_{yy}$$
$$Cov(X) = \Sigma_{xx}$$

Vektor rata-ratanya:

$$\underset{((p+q)x1)}{\mu} = E(X) = \begin{bmatrix} E(X^{(1)}) \\ E(X^{(2)}) \end{bmatrix} = \begin{bmatrix} \mu^{(1)} \\ \mu^{(2)} \end{bmatrix}$$

Vektor acaknya:

$$\underset{((p+q)x1)}{\mu} = \begin{bmatrix} X^{(1)} \\ X^{(2)} \end{bmatrix} = \begin{bmatrix} X_1^{(1)} \\ X_2^{(1)} \\ \vdots \\ X_p^{(1)} \\ \cdots \\ X_1^{(2)} \\ X_2^{(2)} \\ \vdots \\ X_p^{(2)} \end{bmatrix}$$

- Kombinasi linier dari kedua gugus peubah tersebut dapat dituliskan sebagai berikut:

$$\mathbf{U} = \mathbf{a}'\,\mathbf{X} = a1\,X1 + a2\,X2 + \ldots. + aq\,Xq$$
$$\mathbf{V} = \mathbf{b}'\,\mathbf{Y} = b1\,Y1 + b2\,Y2 + \ldots. + bp\,Yp$$

Sehingga,

$$\mathrm{Var}\,(\mathbf{U}) = \mathbf{a}'\mathrm{Cov}(\mathbf{X})\mathbf{a} = \mathbf{a}' \sum\nolimits_{\mathrm{XX}} \mathbf{a}$$
$$\mathrm{Var}\,(\mathbf{V}) = \mathbf{b}'\mathrm{Cov}(\mathbf{Y})\mathbf{b} = \mathbf{b}' \sum\nolimits_{\mathrm{YY}} \mathbf{b}$$
$$\mathrm{Cov}(\mathrm{U},\mathrm{V}) = \mathbf{a}'\,\mathrm{Cov}(\mathbf{X},\mathbf{Y})\,\mathbf{b} = \mathbf{a}' \sum\nolimits_{\mathrm{XY}} \mathbf{b}$$

- Dari sini kita mencari vector koefesien **a** dan **b** sehingga korelasinya maksimum,

$$Corr(U,V) = \frac{a' \Sigma_{XY}\, b}{\sqrt{a' \Sigma_{XX}\, a}\sqrt{b' \Sigma_{YY}\, b}}$$

Definisi:

- **Peubah kanonik ke-k**

  $U_2 = \mathbf{a}_{\boldsymbol{k}}'\,\mathbf{X}\,\mathrm{Var}(U_k) = 1$

  $V_2 = \mathbf{b}_{\boldsymbol{k}}'\,\mathbf{Y}\,\mathrm{Var}(V_k) = 1$

  Maksimum $\mathrm{Corr}(\mathrm{U}_1, \mathrm{V}_1) = \rho_1$

- **Peubah kanonik kedua: korelasi terbesar kedua**

  $\mathrm{U}_2 = \mathbf{a}_2'\,\mathbf{X}\,\mathrm{Var}(\mathrm{U}_2) = 1$

  $V_2 = \mathbf{b}_2'\,\mathbf{Y}\,\mathrm{Var}(V_2) = 1$

$Cov(U_1, U_2) = 0$ $Cov(V_1, V_2) = 0$

$Cov(U_1, V_2) = Cov(U_2, V_1) = 0$

Maksimum $\mathrm{Corr}(U_1, V_1) = \rho_2$

- **Peubah kanonik pertama: korelasi terbesar pertama**

  $U_1 = \mathbf{a_1}'\mathbf{X}\ \mathrm{Var}(U_1) = 1$

  $V_1 = \mathbf{b_1}'\mathbf{Y}\ \mathrm{Var}(V_1) = 1$

  $Cov(U_k, U_l) = 0$

  $Cov = (V_l, V_k) = 0$

  $Cov(U_k, V_l) = 0$

  Dengan $k \neq l$

  Maksimum $\mathrm{Corr}(U_k, V_k) = \rho_k$

Dengan menggunakan ketaksamaan Cauchy-Schwarz atau metode langrange maka diperoleh (**lihat buku Johnson hal 546 untuk pembuktiannya**) koefisien $\boldsymbol{a}$ dan $\boldsymbol{b}$ dengan cara:

- $\rho_1^2 > \rho_2^2 > \cdots > \rho_p^2$ adalah akar ciri (nilai eigen) dari matriks $\sum_{YY}^{-1/2} \sum_{YX} \sum_{XX}^{-1} \sum_{XY} \sum_{YY}^{-1/2}$ yang berpadanan dengan vector ciri $\boldsymbol{f_1}, \boldsymbol{f_2}, \dots, \boldsymbol{f_p}$.

- $\rho_1^2 > \rho_2^2 > \cdots > \rho_p^2$ juga merupakan akar ciri (nilai eigen) dari matriks $\sum_{XX}^{-1/2} \sum_{XY} \sum_{YY}^{-1} \sum_{YX} \sum_{XX}^{-1/2}$ yang berpadanan dengan vector ciri $\boldsymbol{e_1}, \boldsymbol{e_2}, \dots, \boldsymbol{e_p}$.

Sehingga vector koefisien $\boldsymbol{a}$ dan $\boldsymbol{b}$ diperoleh sebagai berikut:

$$
\begin{array}{ll}
a_1 = e_1 \sum_{XX}^{-\frac{1}{2}} & b_1 = f_1 \sum_{YY}^{-\frac{1}{2}} \\
a_2 = e_2 \sum_{XX}^{-\frac{1}{2}} & b_2 = f_2 \sum_{YY}^{-\frac{1}{2}} \\
\vdots & \vdots \\
a_p = e_p \sum_{XX}^{-\frac{1}{2}} & b_p = f_p \sum_{YY}^{-\frac{1}{2}}
\end{array}
$$

4. Uji Hipotesis

a. Uji Simultan

1) Hipotesis:

   a. $H_0 = \rho_1 = \rho_2 = \cdots = \rho_k = 0$ (semua korelasi kanonik bernilai 0)

   b. $H_1 = ada\ \rho_i \neq 0$ (paling tidak ada satu korelasi kanonik yang tidak bernilai 0)

2) Taraf signifikansi

3) Statistik uji

$$B = -\left[n - 1 - \frac{1}{2}(p + q + 1)\right] \ln \Lambda$$

$$\Lambda = \prod_{i=1}^{k} (1 - \rho_i^2)$$

4) Keputusan:

H_0 ditolak pada taraf signifikan α jika $B > \chi_\alpha^2$ dengan derajat bebas $p.q$

b. Uji Parsial

1) Hipotesis:

$H_0 = \rho_1 = \rho_2 = \cdots = \rho_k = 0$ (semua korelasi kanonik bernilai 0)

$H_1 = ada\ \rho_i \neq 0$ (paling tidak ada satu korelasi kanonik yang tidak bernilai 0)

2) Taraf signifikansi
3) Statistik uji

$$B = -\left[n - 1 - \frac{1}{2}(p + q + 1)\right] \ln \Lambda_r$$

$$\Lambda = \prod_{i=1}^{k} (1 - \rho_i^2)$$

4) Keputusan:

H_0 ditolak pada taraf signifikan α jika $B > \chi_\alpha^2$ dengan derajat bebas $(p - r)(q - r)$

5. Redundansi

Mengukur proporsi varians dari satu set variabel (misalnya Y) yang bisa dijelaskan oleh variabel kanonik dari set lain (misalnya X).

✓ **Proporsi Keragaman Y**

Yang diterapkan oleh peubah kanonik V

$$R^2_{Z^{(2)}|V_1,V_2,\ldots,V_r} = \frac{\sum_{i=1}^{r} \sum_{k=1}^{q} r^2_{V_i.Z_k^{(2)}}}{q}$$

✓ **Proporsi Keragaman Y**

Yang diterapkan oleh peubah kanonik W

$$R^2_{(k)Y|X} = \rho_k^2 . R^2_{(k)Y}$$

✓ **Proporsi Keragaman X**

Yang diterapkan oleh peubah kanonik V

$$R^2_{(k)X|Y} = \rho_k^2 . R^2_{(k)X}$$

✓ **Proporsi Keragaman X**

Yang diterapkan oleh peubah kanonik W

$$R^2_{Z^{(1)}|W_1,W_2,\ldots,W_r} = \frac{\sum_{i=1}^{r}\sum_{k=1}^{q} r^2_{W_i.Z_k^{(1)}}}{p}$$

6. Interpretasi Fungsi Kanonik

❑ Weight Kanonik → koefisien kanonik yang dibakukan (a dan b) diinterpretasikan sebagai besarnya konstribusi peubah asal terhadap variat kanonik.

❑ Loadings Kanonik, dibedakan menjadi 3 yaitu:

1) Loadings kanonik peubah independent → $R_{XW} = R_{XX}A_Z$

   → yaitu korelasi antara peubah independen dengan variat kanonik pada himpunan X.

2) Loadings kanonik peubah dependen → $R_{YW} = R_{YY}B_Z$

   → yaitu korelasi antara peubah dependen dengan variat kanonik pada himpunan Y.

3) Cross loadings kanonik → korelasi antara peubah asal dengan fungsi kanonik dari himpunan lainnya (bukan himpunannya sendiri).

   Dengan kata lain:

   - Untuk peubah independen: korelasi antara X dengan variat kanonik Y
   - Untuk peubah dependen: korelasi antara Y dengan variat kanonik X

7. Validasi

- **Membagi sampel menjadi dua bagian**, bagian pertama digunakan untuk menduga fungsi kanonik dan bagian kedua digunakan untuk validasi.
- **Analisis Sensivitas** untuk peubah independent, yaitu dengan membandingkan loading kanonik apabila salah satu dari peubah independent disisihkan dari analisis

8. Contoh Kasus

10.12. A random sample of $n = 70$ families will be surveyed to determine the association between certain "demographic" variables and certain "consumption" variables.

Let

$$\text{Criterion set} \begin{cases} X_1^{(1)} = \text{annual frequency of dining at a restaurant} \\ X_2^{(1)} = \text{annual frequency of attending movies} \end{cases}$$

$$\text{Predictor set} \begin{cases} X_1^{(2)} = \text{age of head of household} \\ X_2^{(2)} = \text{annual family income} \\ X_3^{(2)} = \text{educational level of head of household} \end{cases}$$

Suppose 70 observations on the preceding variables give the sample correlation matrix

$$\mathbf{R} = \left[\begin{array}{c|c} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \hline \mathbf{R}_{21} & \mathbf{R}_{22} \end{array}\right] = \left[\begin{array}{cc|ccc} 1.0 & & & & \\ .80 & 1.0 & & & \\ \hline .26 & .33 & 1.0 & & \\ .67 & .59 & .37 & 1.0 & \\ .34 & .34 & .21 & .35 & 1.0 \end{array}\right]$$

(a) Determine the sample canonical correlations, and test the hypothesis $H_0: \boldsymbol{\Sigma}_{12} = \mathbf{0}$ (or, equivalently, $\boldsymbol{\rho}_{12} = \mathbf{0}$) at the $\alpha = .05$ level. If H_0 is rejected, test for the significance ($\alpha = .05$) of the first canonical correlation.

(b) Using standardized variables, construct the canonical variates corresponding to the "significant" canonical correlation(s).

(c) Using the results in Parts a and b, prepare a table showing the canonical variate coefficients (for "significant" canonical correlations) and the sample correlations of the canonical variates with their component variables.

Penyelesaian:

Diketahui:

- Terdapat 5 variabel.
- Matriks korelasi:

$$R = \begin{bmatrix} 1.00 & 0.80 & 0.26 & 0.67 & 0.34 \\ 0.80 & 1.00 & 0.33 & 0.59 & 0.34 \\ 0.26 & 0.33 & 1.00 & 0.37 & 0.21 \\ 0.67 & 0.59 & 0.37 & 1.00 & 0.35 \\ 0.34 & 0.34 & 0.21 & 0.35 & 1.00 \end{bmatrix}$$

Dari matriks korelasi diperoleh beberapa matriks yang mendukung dalam analisis korelasi kanonik, yaitu:

$$R_{11} = \begin{bmatrix} 1 & 0.8 \\ 0.8 & 1 \end{bmatrix}, \quad R_{12} = \begin{bmatrix} 0.26 & 0.67 & 0.34 \\ 0.33 & 0.59 & 0.34 \end{bmatrix}$$

$$R_{22} = \begin{bmatrix} 1 & 0.37 & 0.21 \\ 0.37 & 1 & 0.35 \\ 0.21 & 0.35 & 1 \end{bmatrix} \qquad R_{21} = R_{12}'$$

a) Menentukan korelasi kanonik sampel & uji hipotesis $H_0: \Sigma_{12} = 0$

Menghitung matriks kanonik

$$M = R_{11}^{-1} R_{12} R_{22}^{-1} R_{21}$$

$$M = \begin{bmatrix} 2.78 & -2.22 \\ -2.22 & 2.78 \end{bmatrix} \begin{bmatrix} 0.26 & 0.67 & 0.34 \\ 0.33 & 0.59 & 0.34 \end{bmatrix} \begin{bmatrix} 1.17 & -0.39 & -0.11 \\ -0.39 & 1.27 & -0.36 \\ -0.11 & -0.36 & 1.15 \end{bmatrix} \begin{bmatrix} 0.26 & 0.33 \\ 0.67 & 0.59 \\ 0.34 & 0.34 \end{bmatrix} = \begin{bmatrix} 0.37 & 0.3 \\ 0.12 & 0.14 \end{bmatrix}$$

Menghitung nilai eigen

$$det(M - \lambda I) = 0$$

$$\begin{vmatrix} 0.37 - \lambda & 0.3 \\ 0.12 & 0.14 - \lambda \end{vmatrix} = 0$$

$$(0.37 - \lambda)(0.14 - \lambda) - (0.12 \times 0.3) = 0$$

$$0.0518 - 0.14\lambda - 0.37\lambda + \lambda^2 - 0.036 = 0$$

$$\lambda^2 - 0.51\lambda + 0.0158 = 0$$

Dengan rumus abc diperoleh nilai eigen:

$\lambda_1 = 0.473$ dan $\lambda_2 = 0.035$

Menghitung nilai korelasi kanonik

$$\rho_1 = \sqrt{\lambda_1} = \sqrt{0.473} = 0.688$$

$$\rho_2 = \sqrt{\lambda_2} = \sqrt{0.035} = 0.187$$

Uji Hipotesis (Simultan)

Hipotesis:

$H_0 = \rho_1 = \rho_2 = \cdots = \rho_k = 0$ (semua korelasi kanonik bernilai 0)

$H_1 = ada\ \rho_i \neq 0$ (paling tidak ada satu korelasi kanonik yang tidak bernilai 0)

Statistik uji:

$$\Lambda = \prod_{i=1}^{k} (1 - \rho_i^2) = \left(1 - (0.688^2)\right)\left(1 - (0.187^2)\right) = (0.526656)(0.965031) = 0.5082$$

Gunakan aproksimasi Bartlett:

$$\chi^2 = -\left[n - 1 - \frac{1}{2}(p + q + 1)\right] \ln \Lambda$$

$$\chi^2 = -\left[70 - 1 - \frac{1}{2}(2 + 3 + 1)\right] \ln(0.5082) = 44.674$$

$df = pq = 2 \times 3 = 6$

$\chi^2_{0.05;6} = 12.592$

Karena statistik uji (44.674) > dari $\chi^2_{0.05;6}$ (12.592) **maka Tolak H0. → ada hubungan multivariat antara variabel X dan Y.**

Uji Hipotesis (Parsial) → korelasi kanonik pertama

Hipotesis:

$H_0 = \rho_1 = 0$

$H_1 = \rho_1 \neq 0$

Statistik uji:

$$\Lambda_1 = \prod_{i=1}^{k}(1-\rho_1^2) = \left(1-(0.688^2)\right) = 0.526656$$

Gunakan aproksimasi Bartlett:

$$\chi^2 = -\left[n-1-\frac{1}{2}(p+q+1)\right]\ln \Lambda_1$$

$$\chi^2 = -\left[70-1-\frac{1}{2}(2+3+1)\right]\ln(0.526656) = 41.037$$

$df = (p-1)(q-1) = (2-1)(3-1) = 2$

$\chi^2_{0.05;2} = 5.591$

Karena statistik uji (41.037) > dari $\chi^2_{0.05;6}$ (5.591) **maka Tolak H0.**

Sehingga, korelasi kanonik pertama signifikan.

Uji Hipotesis (Parsial) → korelasi kanonik kedua

Hipotesis:

$H_0 = \rho_2 = 0$

$H_1 = \rho_2 \neq 0$

Statistik uji:

$$\Lambda_1 = \prod_{i=1}^{k}(1-\rho_2^2) = \left(1-(0.187^2)\right) = 0.965031$$

Gunakan aproksimasi Bartlett:

$$\chi^2 = -\left[n-1-\frac{1}{2}(p+q+1)\right]\ln \Lambda_1$$

$$\chi^2 = -\left[70-1-\frac{1}{2}(2+3+1)\right]\ln(0.965031) = 2.349274$$

$df = (p-1)(q-1) = (2-1)(3-1) = 2$

$\chi^2_{0.05;2} = 5.591$

Karena statistik uji (2.349274) < dari $\chi^2_{0.05;6}$ (5.591) **maka Terima H0.**

Sehingga, korelasi kanonik kedua tidak signifikan.

b) Membentuk variat kanonik (koefisien kanonik baku) — menggunakan variabel yang sudah distandarkan

Mencari vector eigen

$\lambda_1 = 0.473$

$$(M - \lambda I)v = 0$$

$$\begin{bmatrix} 0.37 - 0.473 & 0.3 \\ 0.12 & 0.14 - 0.473 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = 0$$

$$\begin{bmatrix} -0.103 & 0.3 \\ 0.12 & -0.333 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = 0$$

Maka diperoleh:

$-0.103x + 0.3y = 0$, jika $x = 1$ maka $y = 0.343$

Lalu normalisasi agar $var(U_1) = 1$:

$\sqrt{a'R_{11}a} = 1.29 \rightarrow a^* = \begin{bmatrix} \frac{1}{1.29} \\ \frac{0.343}{1.29} \end{bmatrix} = \begin{bmatrix} 0.775 \\ 0.266 \end{bmatrix}$

Sehingga, $\boldsymbol{U_1 = 0.775X_1 + 0.266X_2}$

Hubungan antara a_1 dan b_1

$$b_1 = \frac{1}{\rho_1} R_{YY}^{-1} R_{YX} a^*$$

$$b_1 = \frac{1}{0.688}\begin{bmatrix} 1.17 & -0.39 & -0.11 \\ -0.39 & 1.27 & -0.36 \\ -0.11 & -0.36 & 1.15 \end{bmatrix}\begin{bmatrix} 0.26 & 0.33 \\ 0.67 & 0.59 \\ 0.34 & 0.34 \end{bmatrix}\begin{bmatrix} 0.775 \\ 0.266 \end{bmatrix} = \frac{1}{0.688}\begin{bmatrix} 0.049 \\ 0.619 \\ 0.132 \end{bmatrix} = \begin{bmatrix} 0.072 \\ 0.899 \\ 0.192 \end{bmatrix}$$

Maka diperoleh,

$$\boldsymbol{V_1 = 0.072Y_1 + 0.899Y_2 + 0.192Y_3}$$

c) Tabel koefisien & korelasi (loadings)

Loading untuk U_1

$$Loading\ (U_1) = R_{11}a^*$$

$$R_{11}a^* = \begin{bmatrix} 1 & 0.8 \\ 0.8 & 1 \end{bmatrix} \begin{bmatrix} 0.775 \\ 0.266 \end{bmatrix} = \begin{bmatrix} 0.988 \\ 0.886 \end{bmatrix}$$

Loading untuk V_1

$$Loading\ (V_1) = \frac{R_{22}b}{\sqrt{b'R_{22}b}}$$

$$R_{22}b = \begin{bmatrix} 1 & 0.37 & 0.21 \\ 0.37 & 1 & 0.35 \\ 0.21 & 0.35 & 1 \end{bmatrix} \begin{bmatrix} 0.072 \\ 0.899 \\ 0.192 \end{bmatrix} = \begin{bmatrix} 0.445 \\ 0.993 \\ 0.522 \end{bmatrix}$$

$$\sqrt{b'R_{22}b} = [0.072 \quad 0.899 \quad 0.192] \begin{bmatrix} 0.445 \\ 0.993 \\ 0.522 \end{bmatrix} = \sqrt{1.025} = 1.012$$

Maka

$$Loading\ (V_1) = \frac{R_{22}b}{\sqrt{b'R_{22}b}} = \begin{bmatrix} 0.440 \\ 0.981 \\ 0.515 \end{bmatrix}$$

Tabel 1 Koefisien & Korelasi (Loadings)

Variabel	Koefisien Kanonik	Korelasi (Loading)
Set X (Criterion Set)		
X_1 – Dining	**0.775**	**0.986**
X_2 – Movies	**0.266**	**0.886**
Set Y (Predictor Set)		
Y_1 – Age	**0.072**	**0.440**
Y_2 – Income	**0.899**	**0.981**
Y_3 – Education	**0.192**	**0.515**

Interpretasi:

- Variat kanonik U_1 menggambarkan aktivitas hiburan, terutama Dining dan Movies.
- Variat kanonik V_1 menggambarkan status sosial ekonomi, didominasi oleh pendapatan.
- Korelasi antara kedua variat adalah 0.69, menunjukkan hubungan yang kuat.

- Artinya, orang dengan pendapatan & pendidikan lebih tinggi cenderung lebih sering mengonsumsi hiburan (makan di luar & menonton film).

Code R

```
# Definisikan Matriks
R11 <- matrix(c(1, 0.8,
            0.8, 1), nrow=2, byrow=TRUE)

R12 <- matrix(c(0.26, 0.67, 0.34,
            0.33, 0.59, 0.34), nrow=2, byrow=TRUE)

R22 <- matrix(c(1, 0.37, 0.21,
            0.37, 1, 0.35,
            0.21, 0.35, 1), nrow=3, byrow=TRUE)

# ==========================================================
# (a) Hitung matriks kanonik M = R11^-1 R12 R22^-1 R21
# ==========================================================
R21 <- t(R12)
M <- solve(R11) %*% R12 %*% solve(R22) %*% R21
M

# ==========================================================
# (b) Hitung nilai eigen dan korelasi kanonik
# ==========================================================
eigen_M <- eigen(M)
lambda <- eigen_M$values
lambda

rho <- sqrt(lambda)
rho

# ==========================================================
# (c) Uji signifikansi Wilks' Lambda (simultan)
# ==========================================================
n <- 70
p <- 2
q <- 3
```

```
wilks_lambda <- prod(1 - rho^2)
m <- n - 1 - 0.5 * (p + q + 1)
chi_sq <- -m * log(wilks_lambda)
df <- p * q
p_value <- 1 - pchisq(chi_sq, df)

cat("\nWilks' Lambda =", round(wilks_lambda, 4),
    "\nChi-square =", round(chi_sq, 3),
    "\ndf =", df,
    "\np-value =", round(p_value, 6), "\n")

if (p_value < 0.05) {
  cat("→ Tolak H0: hubungan kanonik signifikan\n")
} else {
  cat("→ Gagal tolak H0: hubungan tidak signifikan\n")
}

#==============================================================
#UJI HIPOTESIS PARSIAL
#==============================================================
# Fungsi untuk uji Bartlett sekuensial (parsial)
bartlett_seq_test <- function(rho, n, p, q) {
  m_all <- length(rho)           # jumlah akar (<= min(p,q))
  results <- data.frame(
    k = integer(),
    Lambda_k = numeric(),
    chi_sq = numeric(),
    df = integer(),
    p_value = numeric(),
    signif = logical()
  )

  # konstanta Bartlett
  C <- (n - 1 - 0.5*(p + q + 1))

  for (k in 1:m_all) {
    # Lambda_k = product_{i=k..m_all} (1 - rho_i^2)
    Lambda_k <- prod(1 - rho[k:m_all]^2)
    df_k <- (p - k + 1) * (q - k + 1)
    # perlakuan numerik: jaga Lambda_k > 0
```

```
    if (Lambda_k <= 0) Lambda_k <- .Machine$double.eps
    chi_sq_k <- - C * log(Lambda_k)
    pval_k <- 1 - pchisq(chi_sq_k, df_k)
    signif_k <- pval_k < 0.05
    results <- rbind(results, data.frame(
      k = k,
      Lambda_k = Lambda_k,
      chi_sq = chi_sq_k,
      df = df_k,
      p_value = pval_k,
      signif = signif_k
    ))
  }
  return(results)
}

# Jalankan uji sekuensial
seq_results <- bartlett_seq_test(rho = rho, n = n, p = p, q = q)

# Tampilkan hasil dengan format rapi
print("Hasil uji Bartlett sekuensial (parsial):")
print(seq_results)

# ============================================================
# (d) Hitung vektor eigen (koefisien kanonik a1)
# ============================================================
a1_raw <- eigen_M$vectors[,1]   # eigenvector untuk λ1
a1_raw

# Normalisasi agar var(U1)=1, yaitu a'R11a = 1
a1 <- a1_raw / sqrt(t(a1_raw) %*% R11 %*% a1_raw)
a1

# ============================================================
# (e) Hitung koefisien b1 di sisi Y
# b1 = (1/rho1) * R22^-1 * R21 * a1
# ============================================================
b1 <- (1/rho[1]) * solve(R22) %*% R21 %*% a1
b1
```

```
# ==========================================================
# (f) Hitung loading U1 dan V1
# ==========================================================
loading_U1 <- R11 %*% a1
loading_V1 <- R22 %*% b1 / as.numeric(t(b1) %*% R22 %*% b1)

loading_U1
loading_V1
```

9. Latihan

Andrews and Herzberg [1] give data obtained from a study of a comparison of nondiabetic and diabetic patients. Three primary variables,

$$X_1^{(1)} = \text{glucose intolerance}$$
$$X_2^{(1)} = \text{insulin response to oral glucose}$$
$$X_3^{(1)} = \text{insulin resistance}$$

and two secondary variables,

$$X_1^{(2)} = \text{relative weight}$$
$$X_2^{(2)} = \text{fasting plasma glucose}$$

were measured. The data for $n = 46$ nondiabetic patients yield the covariance matrix

$$\mathbf{S} = \left[\begin{array}{c|c} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \hline \mathbf{S}_{21} & \mathbf{S}_{22} \end{array}\right] = \left[\begin{array}{ccc|cc} 1100 & 390.7 & 102.4 & 0.187 & 20.23 \\ 390.7 & 2376 & 1137 & -0.414 & -17.96 \\ 102.4 & 1137 & 2130 & 4.189 & -14.840 \\ \hline 0.187 & -0.414 & 4.189 & 0.016 & 0.216 \\ 20.23 & -17.96 & -14.840 & 0.216 & 70.56 \end{array}\right]$$

(a) Determine the sample canonical correlations, and test the hypothesis $H_0: \boldsymbol{\Sigma}_{12} = \mathbf{0}$ (or, equivalently, $\boldsymbol{\rho}_{12} = \mathbf{0}$) at the $\alpha = .05$ level. If H_0 is rejected, test for the significance ($\alpha = .05$) of the first canonical correlation.

(b) Using standardized variables, construct the canonical variates corresponding to the "significant" canonical correlation(s).

(c) Using the results in Parts a and b, prepare a table showing the canonical variate coefficients (for "significant" canonical correlations) and the sample correlations of the canonical variates with their component variables.

Code Soal Latihan (untuk checking hitungan manual)

```
# Definiskan Matriks
R11 <- matrix(c(1100, 390.7, 102.4,
            390.7, 2376, 1137,
            102.4, 1137, 2130), nrow=3, byrow=TRUE)

R12 <- matrix(c(0.187, 20.23,
            -0.414, -17.96,
            4.189, -14.840), nrow=3, byrow=TRUE)

R22 <- matrix(c(0.016, 0.216,
            0.216, 70.56), nrow=2, byrow=TRUE)

# ============================================================
# (a) Hitung matriks kanonik M = R11^-1 R12 R22^-1 R21
# ============================================================
R21 <- t(R12)
M <- solve(R11) %*% R12 %*% solve(R22) %*% R21
M

# ============================================================
# (b) Hitung nilai eigen dan korelasi kanonik
# ============================================================
eigen_M <- eigen(M)
lambda <- eigen_M$values
lambda

rho <- sqrt(lambda)
rho

# ============================================================
# (c) Uji signifikansi Wilks' Lambda (simultan)
# ============================================================
n <- 50
p <- 3
q <- 2
wilks_lambda <- prod(1 - rho^2)
m <- n - 1 - 0.5 * (p + q + 1)
chi_sq <- -m * log(wilks_lambda)
```

```
df <- p * q
p_value <- 1 - pchisq(chi_sq, df)

cat("\nWilks' Lambda =", round(wilks_lambda, 4),
    "\nChi-square =", round(chi_sq, 3),
    "\ndf =", df,
    "\np-value =", round(p_value, 6), "\n")

if (p_value < 0.05) {
  cat("→ Tolak H0: hubungan kanonik signifikan\n")
} else {
  cat("→ Gagal tolak H0: hubungan tidak signifikan\n")
}

#==============================================================
#UJI HIPOTESIS PARSIAL
#==============================================================
# Fungsi untuk uji Bartlett sekuensial (parsial)
bartlett_seq_test <- function(rho, n, p, q) {
  m_all <- length(rho)            # jumlah akar (<= min(p,q))
  results <- data.frame(
    k = integer(),
    Lambda_k = numeric(),
    chi_sq = numeric(),
    df = integer(),
    p_value = numeric(),
    signif = logical()
  )

  # konstanta Bartlett
  C <- (n - 1 - 0.5*(p + q + 1))

  for (k in 1:m_all) {
    # Lambda_k = product_{i=k..m_all} (1 - rho_i^2)
    Lambda_k <- prod(1 - rho[k:m_all]^2)
    df_k <- (p - k + 1) * (q - k + 1)
    # perlakuan numerik: jaga Lambda_k > 0
    if (Lambda_k <= 0) Lambda_k <- .Machine$double.eps
    chi_sq_k <- - C * log(Lambda_k)
    pval_k <- 1 - pchisq(chi_sq_k, df_k)
```

```
    signif_k <- pval_k < 0.05
    results <- rbind(results, data.frame(
      k = k,
      Lambda_k = Lambda_k,
      chi_sq = chi_sq_k,
      df = df_k,
      p_value = pval_k,
      signif = signif_k
    ))
  }
  return(results)
}

# Jalankan uji sekuensial
seq_results <- bartlett_seq_test(rho = rho, n = n, p = p, q = q)

# Tampilkan hasil dengan format rapi
print("Hasil uji Bartlett sekuensial (parsial):")
print(seq_results)

# ============================================================
# (d) Hitung vektor eigen (koefisien kanonik a1)
# ============================================================
a1_raw <- eigen_M$vectors[,1]   # eigenvector untuk λ1
a1_raw

# Normalisasi agar var(U1)=1, yaitu a'R11a = 1
a1 <- a1_raw / sqrt(t(a1_raw) %*% R11 %*% a1_raw)
a1

# ============================================================
# (e) Hitung koefisien b1 di sisi Y
# b1 = (1/rho1) * R22^-1 * R21 * a1
# ============================================================
b1 <- (1/rho[1]) * solve(R22) %*% R21 %*% a1
b1

# ============================================================
# (f) Hitung loading U1 dan V1
# ============================================================
```

```
loading_U1 <- R11 %*% a1
loading_V1 <- R22 %*% b1 / as.numeric(t(b1) %*% R22 %*% b1)

loading_U1

loading_V1
```