

CANONICAL CORRELATION

By:

Dr. Akash Asthana

Assistant Professor,

Dept. of Statistics,

University of Lucknow, Lucknow

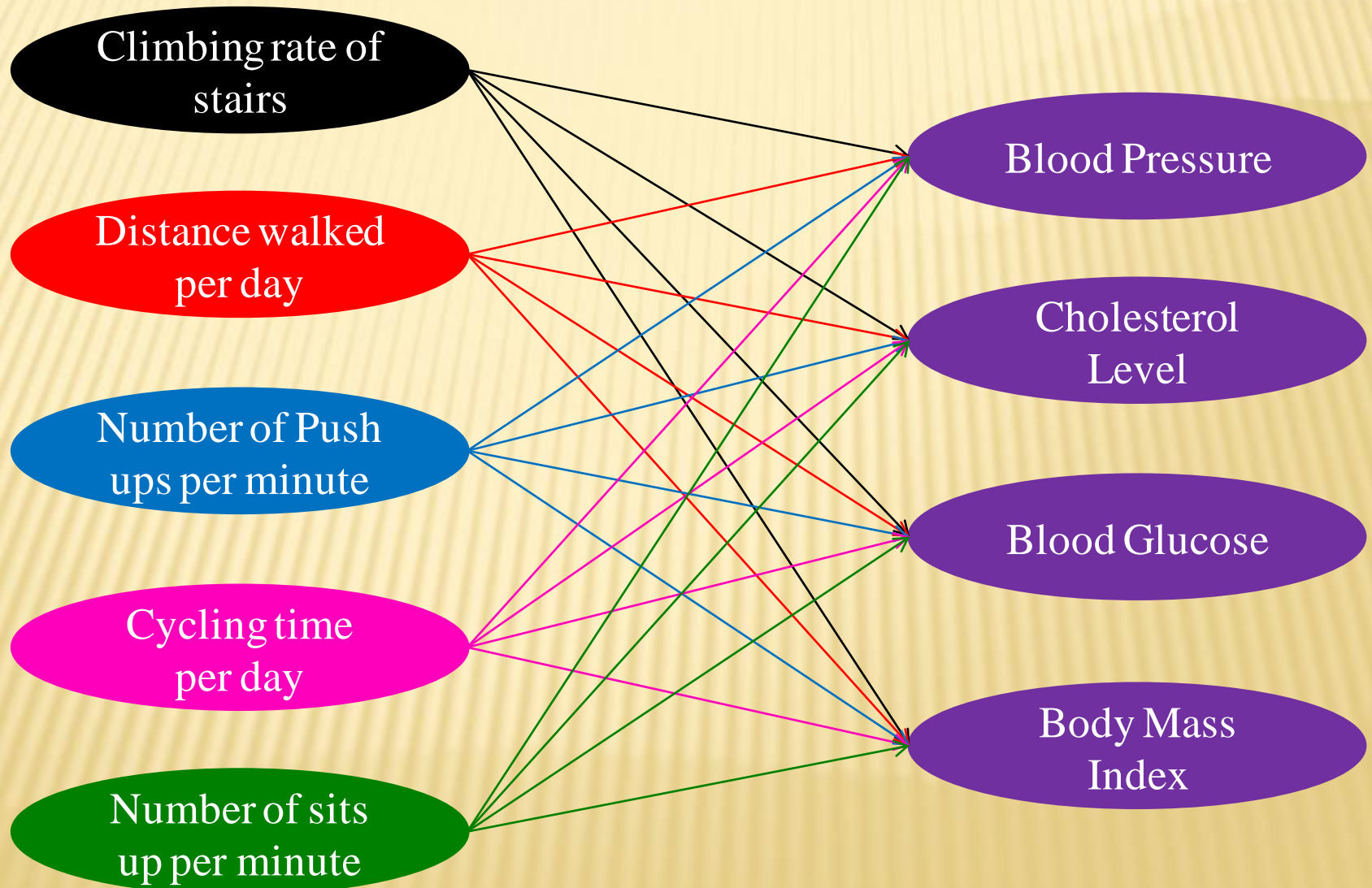
CANONICAL CORRELATION

- ✘ For study the relation between two variables the correlation coefficient is used.
- ✘ But sometimes instead of two variables the two sets of variables are available and one is interested in the relation between these two sets of variables.
- ✘ To study the relation instead of individual variables the linear combinations of these variables are used.
- ✘ These linear combinations are constructed in such a manner that the one linear combination of one set of variables is highly correlated with one linear combination of other set of variable and independent of all the other linear combinations.
- ✘ Also the linear combinations obtained from same set of variables should be independent of each other.
- ✘ These linear combinations of variables are called as canonical variates and correlation between the linear combinations of two sets is called as canonical correlation.

CANONICAL CORRELATION (EXAMPLE)

- ✘ Consider, as an example, variables related to exercise and health.
- ✘ In one set we can take variables associated with exercise, observations such as the climbing rate on a stairs, running speed, weight lifting, the number of push-ups per minute, number of sit ups per minute etc.
- ✘ In other set of variables we can consider the measures of overall health, such as blood pressure, cholesterol levels, glucose levels, body mass index, etc.
- ✘ Two types of variables are measured and the relationships between the exercise variables and the health variables are of interest.
- ✘ It is clear that the each variable in first affects to each variable in the second set.

CANONICAL CORRELATION (EXAMPLE)



CANONICAL CORRELATION (EXAMPLE)

- ✘ As a second example consider variables measured on environmental health and environmental toxins.
- ✘ A number of environmental health variables such as frequencies of sensitive species, species diversity, total biomass, productivity of the environment, etc. may be considered in one set.
- ✘ In second set of variables on environmental toxins can be considered, such as the concentrations of heavy metals, pesticides, dioxin, etc.

CANONICAL CORRELATION

- ✘ For obtaining the canonical variates let's consider two different set of variables $X^{(1)}$ and $X^{(2)}$. In other words we can partition vector X of p -variables into two parts $X^{(1)}$ having p_1 variables and $X^{(2)}$ having p_2 variables.
- ✘ Without loss of generality we can take $E[X] = 0$ and $V(X) = \Sigma$.
- ✘ Further the variance-covariance matrix Σ can be partitioned as:

$$\Sigma = \begin{bmatrix} V(X_1) & Cov(X_1, X_2) \\ Cov(X_2, X_1) & V(X_2) \end{bmatrix} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

- ✘ Now the main interest is to study the relation between $X^{(1)}$ and $X^{(2)}$.
- ✘ For the purpose we can consider two linear combinations, one of the elements of $X^{(1)}$ and other as linear combination of the elements of $X^{(2)}$.

CANONICAL CORRELATION

- ✘ Let these linear combinations are $U = \alpha'X^{(1)}$ and $V = \gamma'X^{(2)}$ such that $V(U) = 1$ and $V(V) = 1$.
- ✘ $E[U] = E[\alpha'X^{(1)}] = \alpha'E[X^{(1)}] = 0$; $E[V] = E[\gamma'X^{(2)}] = \gamma'E[X^{(2)}] = 0$
- ✘ $V(U) = V(\alpha'X^{(1)}) = \alpha'V(X^{(1)})\alpha = \alpha'\Sigma_{11}\alpha = 1$ (1)
- ✘ $V(V) = V(\gamma'X^{(2)}) = \gamma'V(X^{(2)})\gamma = \gamma'\Sigma_{22}\gamma = 1$ (2)
- ✘ Now under these conditions the correlation between U and V can be defined by the covariance between U and V.
- ✘ The covariance between U and V is given by:
$$\begin{aligned}Cov(U, V) &= Cov(\alpha'X^{(1)}, \gamma'X^{(2)}) \\ &= \alpha'Cov(X^{(1)}, X^{(2)})\gamma \\ &= \alpha'\Sigma_{12}\gamma\end{aligned}$$
(3)
- ✘ For obtaining the canonical variates (3) is to be maximized under the conditions (1) and (2).

CANONICAL CORRELATION

- ✘ For maximizing (3) under conditions (1) and (2) the Lagrange's method of multiplier is used. The function φ can be defined as:

$$\varphi = \alpha' \Sigma_{12} \gamma - \frac{1}{2} \lambda (\alpha' \Sigma_{11} \alpha - 1) - \frac{1}{2} \mu (\gamma' \Sigma_{22} \gamma - 1)$$

- ✘ Where λ and μ are the Lagrange's multipliers (scalars).

- ✘ For Maximum $\frac{\partial \varphi}{\partial \alpha'} = 0$ and $\frac{\partial \varphi}{\partial \gamma'} = 0$

- ✘ From it we will get:

$$\Sigma_{12} \gamma - \lambda \Sigma_{11} \alpha = 0 \quad (4)$$

$$\Sigma_{12}' \alpha - \mu \Sigma_{22} \gamma = 0 \quad (5)$$

- ✘ Pre multiplying equation (4) by α' and equation (5) by γ' we get

$$\alpha' \Sigma_{12} \gamma - \lambda \alpha' \Sigma_{11} \alpha = 0 \quad (6)$$

$$\gamma' \Sigma_{12}' \alpha - \mu \gamma' \Sigma_{22} \gamma = 0 \quad (7)$$

- ✘ As $\alpha' \Sigma_{11} \alpha = 1$ and $\gamma' \Sigma_{22} \gamma = 1$, we get:

$\lambda = \mu = \alpha' \Sigma_{12} \gamma$; putting in equation (5) we get:

CANONICAL CORRELATION

$$\Sigma_{12}\gamma - \lambda\Sigma_{11}\alpha = 0 \quad (8)$$

$$\Sigma_{21}\alpha - \mu\Sigma_{22}\gamma = 0 \quad (9)$$

- ✘ In matrix form we can write equation (8) and equation (9) as:

$$\begin{bmatrix} -\lambda\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\lambda\Sigma_{22} \end{bmatrix} \begin{bmatrix} \alpha \\ \gamma \end{bmatrix} = \mathbf{0} \quad (10)$$

- ✘ Equation (10) will have a solution if:

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

- ✘ Equation (11) is a polynomial in λ of degree p . Therefore it will have at most p distinct solutions. Let these solutions be $\lambda_1, \lambda_2, \dots, \lambda_p$ such that $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$.
- ✘ Let the solutions of (10) be denoted by $\alpha^{(1)}$ and $\gamma^{(1)}$. Then the first canonical variates will be $U_1 = \alpha^{(1)'}\mathbf{X}^{(1)}$ and $V_1 = \gamma^{(1)'}\mathbf{X}^{(2)}$.

CANONICAL CORRELATION

✘ Let up to rth satge ($r < p$) we get the canonical variates as
 $\{U_1 = \alpha^{(1)'} X^{(1)}, V_1 = \gamma^{(1)'} X^{(2)}\}; \{U_2 = \alpha^{(2)'} X^{(1)}, V_2 = \gamma^{(2)'} X^{(2)}\}; \dots;$
 $\{U_r = \alpha^{(r)'} X^{(1)}, V_r = \gamma^{(r)'} X^{(2)}\}$

✘ Then for $(r+1)$ th canonical variates we can define the linear combinations: $U = \alpha X^{(1)}, V = \gamma X^{(2)}$ such that:

$$V(U) = \alpha' \Sigma_{11} \alpha = 1 \quad (12)$$

$$V(V) = \gamma' \Sigma_{22} \gamma = 1 \quad (13)$$

$$\text{Cov}(U, U_i) = \alpha' \Sigma_{11} \alpha^{(i)} = 0 \quad \forall i = 1, 2, \dots, r \quad (14)$$

$$\text{Cov}(V, V_i) = \gamma' \Sigma_{22} \gamma^{(i)} = 0 \quad \forall i = 1, 2, \dots, r \quad (15)$$

$$\text{Cov}(U, V_i) = \alpha' \Sigma_{12} \gamma^{(i)} = \lambda_i \alpha' \Sigma_{11} \alpha^{(i)} = 0 \quad \forall i = 1, 2, \dots, r \quad (16)$$

$$\text{Cov}(U_i, V) = \alpha^{(i)'} \Sigma_{12} \gamma = \lambda_i \gamma^{(i)'} \Sigma_{22} \gamma = 0 \quad \forall i = 1, 2, \dots, r \quad (17)$$

$$\text{Cov}(U, V) = \alpha' \Sigma_{12} \gamma \quad (18)$$

✘ Then we can define the function $\varphi_{(r+1)}$ as:

$$\varphi_{(r+1)} = \alpha' \Sigma_{12} \gamma - \frac{1}{2} \lambda (\alpha' \Sigma_{11} \alpha - 1) - \frac{1}{2} \mu (\gamma' \Sigma_{22} \gamma - 1) + \sum_{i=1}^r v_i \alpha' \Sigma_{11} \alpha^{(i)} + \sum_{i=1}^r \theta_i \gamma' \Sigma_{22} \gamma^{(i)}$$

CANONICAL CORRELATION

✘ For maximum correlation among (r+1)th pair of canonical variates: $\frac{\partial \varphi_{(r+1)}}{\partial \alpha'} = 0$ and $\frac{\partial \varphi_{(r+1)}}{\partial \gamma'} = 0$

✘ It will give rise to following equations:

$$\Sigma_{12}\gamma - \lambda\Sigma_{11}\alpha + \sum_{i=1}^r v_i \Sigma_{11}\alpha^{(i)} = 0 \quad (19)$$

$$\Sigma_{21}\alpha - \mu\Sigma_{22}\gamma + \sum_{i=1}^r \theta_i \Sigma_{22}\gamma^{(i)} = 0 \quad (20)$$

✘ Premultiplying Eqⁿ (19) by $\alpha^{(i)}$ ' and Eqⁿ (20) by $\gamma^{(i)}$ ' and using Eqⁿ from (14) to (17) we get:

$$v_i \alpha^{(i)'} \Sigma_{11} \alpha^{(i)} = v_i = 0$$

$$\theta_i \gamma^{(i)'} \Sigma_{22} \gamma^{(i)} = \theta_i = 0$$

✘ Putting in equation (19) and (20) we get:

CANONICAL CORRELATION

$$\Sigma_{12}\gamma - \lambda\Sigma_{11}\alpha = 0 \quad (21)$$

$$\Sigma_{21}\alpha - \mu\Sigma_{22}\gamma = 0 \quad (22)$$

- ✘ Which are similar to equation (4) and (5). Therefore we can say that α and γ will be the solution of equation (10) and λ be solution of equation (11).

$$\begin{bmatrix} -\lambda\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\lambda\Sigma_{22} \end{bmatrix} \begin{bmatrix} \alpha \\ \gamma \end{bmatrix} = \mathbf{0}$$

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

- ✘ Hence in general we can say that the pairs of canonical covariate can be obtained by solving equation (10) and the values of λ are the solutions of the equation (11) arranged in descending order. The r th pair of canonical variates are given by $U_r = \alpha^{(r)'}X^{(1)}$, $V_r = \gamma^{(i)'}X^{(2)}$

STEPS FOR OBTAINING CANONICAL VARIATES

- ✘ Let Y and Z be two sets of Variables for which canonical variates are required to be obtained. Then:
1. Define the vector $X' = (Y' Z')$.
 2. Obtain the correlation matrix for X . If Y and Z be the sets of variables such that $E(Y) = 0$ and $E(Z) = 0$ then Variance-covariance matrix of X can be treated as correlation matrix. If Population variance is unknown then under the assumption that $X \sim N(0, \Sigma)$ we can use the MLE of Σ .
 3. Solve the equation
$$\begin{vmatrix} -\lambda\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\lambda\Sigma_{22} \end{vmatrix} = 0$$
 and obtain the value of λ .
 4. Arrange the values of λ in descending order.
 5. For each value of λ solve the equation
$$\begin{bmatrix} -\lambda\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\lambda\Sigma_{22} \end{bmatrix} \begin{bmatrix} \alpha \\ \gamma \end{bmatrix} = 0$$
 6. Now define $U_r = \alpha^{(r)'}X^{(1)}$, $V_r = \gamma^{(i)'}X^{(2)}$. it will give the r^{th} set of canonical covariates and λ_r be the canonical correlation among them.

REFERENCES

1. Anderson TW, An introduction to Multivariate Statistical Analysis, 3rd Edition, John Wiley & Sons Inc., New Jersey.
2. Lesson 13 for Course of Penn State University STAT505 available on website <https://online.stat.psu.edu/stat505/book/export/html/682>.
3. Malhotra NK, Birks DF, Marketing Research an Applied Approach, 4th Edition, Prentice Hall, New Delhi.
4. Johnson RA, Wichern DW, Applied Multivariate Statistical Analysis, 3rd Edition, Prentice Hall, New Delhi.
5. Morrison DF, Multivariate Statistical Methods, 2nd Edition, McGraw Hill Publication, India.