Stat 571
11/12/13
1

Canonical Correlation

Start with 2 sets of variables

$X_1, \dots, X_p$ and $Y_1, \dots, Y_q$

Let $W = a_1 X_1 + \dots + a_p X_p = \vec{a}' \vec{X}$

and $Z = b_1 Y_1 + \dots + b_q Y_q = \vec{b}' \vec{Y}$

Goal: find $\vec{a}$ and $\vec{b}$ that maximize the correlation between W and Z.

Let $\begin{bmatrix} \vec{X} \\ \hline \vec{Y} \end{bmatrix} = \begin{bmatrix} X_1 \\ \vdots \\ X_p \\ \hline Y_1 \\ \vdots \\ Y_q \end{bmatrix}_{(p+q)\times 1}$

2

Let $\Sigma = \left[\begin{array}{c|c} \overset{p\times p}{\Sigma_{XX}} & \overset{p\times q}{\Sigma_{XY}} \\ \hline \underset{q\times p}{\Sigma_{YX}} & \underset{q\times q}{\Sigma_{YY}} \end{array}\right]$

$$\text{Corr}(W,Z) = \frac{\text{Cov}(W,Z)}{\sqrt{V(W)\,V(Z)}}$$

$$\text{Cov}(W,Z) = \text{Cov}(\vec{a}'\vec{X}, \vec{b}'\vec{Y}) = \vec{a}'\,\text{Cov}(\vec{X},\vec{Y})\,\vec{b}$$

$$= \vec{a}'\,\Sigma_{XY}\,\vec{b}$$

③

$$V(w) = V(\vec{a}'\vec{X}) = \vec{a}'\,Cov(\vec{X})\,\vec{a}$$
$$= \vec{a}'\,\Sigma_{xx}\,\vec{a}$$

$$V(z) = V(\vec{b}'\vec{Y}) = \vec{b}'\,Cov(\vec{Y})\,\vec{b}$$
$$= \vec{b}'\,\Sigma_{YY}\,\vec{b}$$

$$Corr(w,z) = \frac{\vec{a}'\,\Sigma_{XY}\,\vec{b}}{\sqrt{(\vec{a}'\Sigma_{xx}\vec{a})(\vec{b}'\Sigma_{YY}\vec{b})}}$$

Without loss of generality, we can choose $\vec{a}$ and $\vec{b}$ such that $\vec{a}'\Sigma_{xx}\vec{a} = \vec{b}'\Sigma_{YY}\vec{b} = 1$

④

Restatement of our goal:

Maximize $r = \vec{a}'\Sigma_{XY}\vec{b}$, subject to the constraints $\vec{a}'\Sigma_{XX}\vec{a} = \vec{b}'\Sigma_{YY}\vec{b} = 1$

Use LaGrange multipliers:

Let $F(\vec{a}, \vec{b}, \delta_1, \delta_2)$

$$= \vec{a}'\Sigma_{XY}\vec{b} - \delta_1(\vec{a}'\Sigma_{XX}\vec{a} - 1) - \delta_2(\vec{b}'\Sigma_{YY}\vec{b} - 1)$$

5

1. $\dfrac{\partial F}{\partial \vec{a}} = \Sigma_{XY}\vec{b} - \delta_1 2\Sigma_{XX}\vec{a} \overset{set}{=} \vec{0}$

2. $\dfrac{\partial F}{\partial \vec{b}} = \Sigma_{YX}\vec{a} - \delta_2 2\Sigma_{YY}\vec{b} \overset{set}{=} \vec{0}$

3. $\dfrac{\partial F}{\partial \delta_1} = -(\vec{a}'\Sigma_{XX}\vec{a} - 1) \overset{set}{=} 0$

4. $\dfrac{\partial F}{\partial \delta_2} = -(\vec{b}'\Sigma_{YY}\vec{b} - 1) \overset{set}{=} 0$

6

1. $\Sigma_{XY}\vec{b} = 2\delta_1\Sigma_{XX}\vec{a}$

$\vec{a}'\Sigma_{XY}\vec{b} = 2\delta_1\vec{a}'\Sigma_{XX}\vec{a} = 2\delta_1$

2. $\Sigma_{YX}\vec{a} = 2\delta_2\Sigma_{YY}\vec{b}$

$\vec{a}'\Sigma_{XY} = 2\delta_2\vec{b}'\Sigma_{YY}$

$\vec{a}'\Sigma_{XY}\vec{b} = 2\delta_2\vec{b}'\Sigma_{YY}\vec{b} = 2\delta_2$

$\therefore \delta_1 = \delta_2 = \frac{1}{2}\vec{a}'\Sigma_{XY}\vec{b}$

Back to 1:

$$\Sigma_{XY}\vec{b} = 2\delta_1 \Sigma_{XX}\vec{a}$$

$$\Sigma_{XX}^{-1}\Sigma_{XY}\vec{b} = 2\delta_1 \underbrace{\Sigma_{XX}^{-1}\Sigma_{XX}}_{I}\vec{a}$$

$$\vec{a} = \frac{1}{2\delta_1}\Sigma_{XX}^{-1}\Sigma_{XY}\vec{b}$$

Substitute into 2:

$$\Sigma_{YX}\left(\frac{1}{2\delta_1}\Sigma_{XX}^{-1}\Sigma_{XY}\vec{b}\right) = 2\delta_2 \Sigma_{YY}\vec{b}$$

$$\Sigma_{YY}^{-1}\Sigma_{YX}\Sigma_{XX}^{-1}\Sigma_{XY}\vec{b} = \underbrace{4\delta_1\delta_2}_{4\delta_1^2}\underbrace{\Sigma_{YY}^{-1}\Sigma_{YY}}_{I}\vec{b}$$

$$\Big(\underbrace{\Sigma_{YY}^{-1}\Sigma_{YX}\Sigma_{XX}^{-1}\Sigma_{XY}}_{M} - \underbrace{4\delta_1^2}_{\lambda} I\Big)\vec{b} = \vec{0}$$

$$(M - \lambda I)\vec{b} = \vec{0}$$

That is, $\vec{b}$ is the eigenvector of M, corresponding to the eigenvalue λ.

9

But $\lambda = 4\delta_1^2 = 4\left(\frac{1}{2}\vec{a}'\Sigma_{XY}\vec{b}\right)^2$
$$= \left(\vec{a}'\Sigma_{XY}\vec{b}\right)^2$$
$$= r^2$$

Since our goal was to maximize r,
take λ to be λ_1, the largest eigenvalue
of M.

Continuation of the process: Once W_1 and Z_1 are found, find W_2 & Z_2 that have the highest possible correlation, subject to being uncorrelated with W_1 & Z_1

10

That is, maximize
$$r = \frac{\vec{a}_2'\Sigma_{XY}\vec{b}_2}{\sqrt{(\vec{a}_2'\Sigma_{XX}\vec{a}_2)(\vec{b}_2'\Sigma_{YY}\vec{b}_2)}}$$

Subject to: $\vec{a}_2'\Sigma_{XX}\vec{a}_2 = \vec{b}_2'\Sigma_{YY}\vec{b}_2 = 1$ and

$$\vec{a}_2'\Sigma_{XX}\vec{a}_1 = \vec{a}_2'\Sigma_{XY}\vec{b}_1 = \vec{b}_2'\Sigma_{YY}\vec{b}_1 = \vec{b}_2'\Sigma_{YX}\vec{a}_1 = 0$$

Solution is again found by LaGrange multipliers, and it says to select λ_2, the 2nd-largest eigenvalue of M, and its eigenvector $\vec{b}_2$.

(11)

$H_0: \lambda_k = \lambda_{k+1} = \dots = \lambda_s = 0 \quad (s = \min(p,q))$

__Bartlett's Test__ let $\Lambda_k = \prod_{i=k}^{s} (1-\lambda_i)$

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

Then $V_k \sim \chi^2$ $df = (p-k+1)(q-k+1)$

__Rao's Test__ let $F = \frac{v_2}{v_1}\left[\frac{1-\Lambda_k^{1/b}}{\Lambda_k^{1/b}}\right]$,

where $a = n-1-\frac{1}{2}(p+q+1)$,

(12)

$$b = \sqrt{\frac{(p-k+1)^2(q-k+1)^2-4}{(p-k+1)^2+(q-k+1)^2-5}},$$

$v_1 = (p-k+1)(q-k+1)$, $v_2 = ab - \frac{1}{2}v_1 + 1$

$$F \sim F_{v_1, v_2}$$

SAS: PROC CANCORR (PSPP)

Minitab: non existent

SPSS: CANCORR Statement (Syntax file)

R : CANCORR?
CCA

HW#5 p.573 #10.18 (13)
due 11/19