BAB 10

Menginterpretasikan Populasi Variabel Kanonik

Variabel kanonik secara umumnya artifisal. Jika variabel awal $X^{(1)}$ dan $X^{(2)}$ digunakan, koefisien kanonik a dan b mempunyai unit proporsi dari himpunan $X^{(1)}$ dan $X^{(2)}$. Jika variabel awal yang distandardisasikan mempunyai rata-rata nol dan unit varians, maka koefisien kanonik tidak mempunyai unit dari pengukuran, dan pasti diinterpretasikan ke dalam bentuk variabel yang distandarkan.

10.2.1 Mengidentifikasi Varibel Kanonik

Walaupun variabel kanonik artifisal, variabel kanonik dapat diidentifikasi dalam bentuk variabel pokok. Identifikasi sering dibantu dengan menghitung korelasi antara variabel kanonik dan variabel awal.

Misalkan A = [$a_1, a_2, ..., a_p$]' dan B = [$b_1, b_2, ..., b_p$]', sehingga vektor dari variabel kanonik adalah $\underset{(px1)}{U} = AX^{(1)} \text{ dan } \underset{(qx1)}{V} = BX^{(2)}$ **(2-10)**

dimana kita awalnya tertarik di variabel kanonik pertama *p* di **V**. Maka,

$$Cov(U, X^{(1)}) = Cov(AX^{(1)}, X^{(1)}) = A\sum\nolimits_{11} \qquad \textbf{(2-11)}$$

Karena $Var(U_i) = 1, Corr(U_i, X_k^{(1)})$ diperoleh dengan membagi $Cov(U, X^{(1)})$ oleh $\sqrt{\text{var}(X_k^{(1)})} = \sigma_{kk}^{1/2}$. Secara ekuivalen,

$$Corr(U_i, X_k^{(1)}) = Cov(U_i, \sigma_{kk}^{-1/2} X_k^{(1)}).$$

Pendahuluan (p x p) diagonal matriks $V_{11}^{-1/2}$ elemen diagonal ke-k $\sigma_{kk}^{-1/2}$ dalam bentuk

matriks, $\underset{(pxp)}{\rho_{U,X^{(1)}}} = Corr(U, X^{(1)}) = Cov(U, V_{11}^{-1/2} X^{(1)}) = Cov(AX^{(1)}, V_{11}^{-1/2} X^{(1)}) = A\sum\nolimits_{11} V_{11}^{-1/2}$

Perhitungan yang sama untuk bagian $(U, X^{(2)}), (V, X^{(1)})$ dan $(V, X^{(2)})$ menghasilkan

$$\underset{(pxp)}{\rho_{U,X^{(1)}}} = A\sum_{11} V_{11}^{-1/2}, \qquad \underset{qxq}{\rho_{U,X^{(2)}}} = B\sum_{22} V_{22}^{-1/2},$$
$$\underset{(pxq)}{\rho_{U,X^{(2)}}} = A\sum_{12} V_{22}^{-1/2}, \qquad \underset{(qxp)}{\rho_{U,X^{(1)}}} = B\sum_{21} V_{11}^{-1/2}, \qquad \textbf{(2-12)}$$

dimana $V_{22}^{-1/2}$ adalah matriks diagonal (q x q) dengan elemen ke-i $\sqrt{\text{var}(X_i^{(2)})}$.

Variabel kanonik diturunkan dari variabel standard terkadang diinterpretasikan dengan menghitung korelasi.

$$\rho_{U,Z^{(1)}} = A_Z \rho_{11} \qquad \rho_{V,Z^{(2)}} = B_Z \rho_{22}$$
$$\rho_{U,Z^{(2)}} = A_Z \rho_{12} \qquad \rho_{V,Z^{(1)}} = B_Z \rho_{21} \qquad \textbf{(2-13)}$$

dimana $\underset{(pxp)}{A_Z}$ dan $\underset{(qxq)}{B_Z}$ adalah matriks yang barisnya memuat koefisien kanonik untuk himpunan $Z^{(1)}$ dan $Z^{(2)}$ secara berurut. Korelasi pada matriks yang ditunjukkan (2-13) mempunyai nilai numerik sama dengan yang dimunculkan (2-12), yakni $\rho_{U,X^{(1)}} = \rho_{U,Z^{(1)}}$ dan seterusnya. Mengikuti ini,

$\rho_{U,X^{(1)}} = A\sum_{11} V_{11}^{-1/2} = AV_{11}^{1/2}V_{11}^{-1/2}\sum_{11} V_{11}^{-1/2} = A_Z \rho_{11} = \rho_{U,Z^{(1)}}$ korelasi tidak dipengaruhi oleh standaridisasi.

10.2.2 Korelasi Kanonik Sebagai Generalisasi Dari Koefisien Korelasi Lainnya

Pertama-tama, koefisien korelasi menyamaratakan korelasi antara dua variabel. Ketika $X^{(1)}$ dan $X^{(2)}$ masing-masing terdiri dari variabel tunggal, sehingga p = q = 1, $|Corr(X_1^{(1)}, X_1^{2})| = |Corr(aX_1^{(1)}, bX_1^{2})|$ untuk semua a, b. Oleh karena itu variasi kanonik $U_1 = X_1^{(1)}$ dan $V_1 = X_1^{(2)}$ memiliki korelasi $\rho_1^* |Corr(X_1^{(1)}, X_1^{2})|$ ketika $X^{(1)}$ dan $X^{(2)}$

memiliki komponen lebih, kondisi $a' = [0,...,0,1,0,...,0]$ dengan 1 pada posisi ke-i dan $b' = [0,...,0,1,0,...,0]$ dengan 1 pada posisi ke-i menghasilkan,

$$\left|Corr(X_i^{(1)}, X_k^{2})\right| = \left|Corr(a'X_1^{(1)}, b'X_1^{2})\right| \le \max_{a,b} Corr(a'X_1^{(1)}, b'X_1^{2}) = \rho_1^* \qquad \textbf{(2-14)}$$

yaitu bahwa korelasi kanonik yang pertama lebih besar dari harga mutlak semua elemen dalam $\rho_{12} = V_{11}^{-1/2}\sum_{12}V_{22}^{-1/2}$.

Kedua, perkalian koefisien korelasi $\rho_{1(X^{(2)})}$ adalah persoalan khusus dari korelasi kanonik ketika $X^{(1)}$ memiliki elemen tunggal $X_2^{(1)}$ (p=1), menimbulkan

$$\rho_{1(X^{(2)})} = \max_b \left|corr X_1^{(1)}, b'X^{(2)}\right| = \rho_1^* \quad \text{, untuk p=1} \qquad \textbf{(2-15)}$$

Ketika p > 1, ρ_1^* lebih besar dari setiap korelasi perkalian $X_i^{(1)}$ dengan $X^{(2)}$ atau korelasi perkalian $X_1^{(2)}$ dengan $X^{(1)}$. Akibatnya,

$$\rho_{U_k(X^{(2)})} = \max_b Corr(U_k, b'X^{(2)}) = Corr(U_k, V_k) = \rho_k^*,\; k = 1,2,\ldots,p \qquad \textbf{(2-16)}$$

yaitu korelasi kanonik juga bermacam-macam koefisien korelasi dari U_k dengan $X^{(2)}$ atau atau bermacam-macam koefisien korelasi dari V_k dengan $X^{(1)}$.

Karena penerapan dari bermacam-macam koefisien korelasi, korelasi kanonik ke-k kuadrat, ρ_k^{*2}, adalah sebanding dengan varians dari variasi kanonik U_k yang dijelaskan oleh himpunan $X^{(2)}$ dan juga sebanding dengan varians dari variasi kanonik V_k yang dijelaskan oleh himpunan $X^{(1)}$. Oleh karena itu, ρ_k^{*2} seringkali dinamakan varians bersama antara dua himpunan $X^{(1)}$ dan $X^{(2)}$. Untuk nilai yang terbesar, ρ_k^{*2} sering disebut varian, kadang-kadang dianggap sebagai kadang-kadang dikenal sebagi sebuah pengukur dari set yang berlebihan. 7*overlap* (tumpang tindih).

10.2.3 r Variabel kanonik pertama sebagai sebuah ringkasan dari perubahan

Perubahan koordinat dari $X^{(1)}$ ke $U = AX^{(1)}$ dan dari $X^{(2)}$ ke $V = BX^{(2)}$ dilakukan untuk memaksimalkan $Corr(U_1, V_1)$ dengan berturut-turut

$Corr(U_i, V_i)$ dimana (U_i, V_i) memiliki korelasi nol dengan pasangan (U_i, V_i), (U_2, V_2), ..., (U_{i-1}, V_{i-1}). Korelasi antara himpunan $X^{(1)}$ dan $X^{(2)}$ telah dimasukkan kedalam pasangan variabel kanonik.

Dengan membuat model, vektor koefisien a_i, b_i dipilih untuk memaksimumkan korelasi, tidak perlu menampilkan variabel penaksir himpunan bagian dari kovarian $\sum_{11}$ dan $\sum_{22}$. Ketika beberapa pasangan dari variabel kanonik yang pertama memberikan kesimpulan yang kecil dari variabilitas dalam $\sum_{11}$ dan $\sum_{22}$, maka tidaklah jelas bagaimana korelasi kanonik dapat diinterpretasikan.

10.2.4 Interpretasi Geometrik dari Analisis Korelasi Kanonik Populasi

Interpretasi geometrik dari prosedur pemilihan variabel kanonik memberrikan pengetahuan yang berharga kedalam sifat analisis korelasi kanonik. Transformasi

$$U = AX^{(1)}$$

dari $X^{(1)}$ ke U memberikan $Cov(U) = A\sum_{11} A' = I$.

Dari 2.1 dan $A = E^{'}\sum_{11}^{-1/2} = E^{'}P_1 \Lambda_1^{-1/2} P_1^{'}$ dimana $E^{'}$ adalah matriks orrthogonal dengan baris $e_i^{'}$ dan $\sum_{11} = P_1 \Lambda_1 P_1^{'}$. Sekarang $P_1^{'}X^{(1)}$ adalah himpunan dari komponen utama yang berasal dari $X^{(1)}$ saja. Matriks $\Lambda_1^{-1/2} P_1^{'}X^{(1)}$ memiliki ke-i baris $\frac{1}{\sqrt{\lambda_i}} P_i^{'}X^{(1)}$, yang komponen utama ke-i nya ditetapkan memiliki varians I. Yaitu $Cov(\Lambda_1^{-1/2} P_1^{'}X^{(1)}) = \Lambda_1^{-1/2} P_1^{'} \sum_{11} P_1 \Lambda_1^{-1/2} = \Lambda_1^{-1/2} P_1^{'} P_1 \Lambda_1 P_1^{'} P_1 \Lambda_1^{-1/2} = \Lambda_1^{-1/2} \Lambda_1 \Lambda_1^{-1/2} = 1.$

Akibatnya, U = $AX^{(1)}$ = $E^{'}P_1 \Lambda_1^{-1/2} P_1^{'}X^{(1)}$ dapat diinterpretasikan sebagai:

1. Transformasi dari $X^{(1)}$ ke komponen utama standar yang tidak berkorelasi,
2. Rotasi orrthogonal P_1 yang ditentukan oleh $\sum_{11}$, dan

3. Rotasi *E'* yang ditentukan dari matriks kovarian penuh ∑.

   Interpretasi serupa berlaku untuk $V = BX'$.