Situasi dimana Variansi Sampel Umum adalah Nol

Variansi akan nol dalam beberapa situasi. Variansi yang nol ini mengindikasikan exterem degeneracy dalam pengertian sekurang-kurangnya satu baris dari matriks deviasi tsb, dapat ditunjukkan sebagai kombinasi linear oleh baris yang lainnya. Sebagaimana telah kita tunjukkan secara geometri, dalam kasus ini dimana satu dari vktor selisih –untuk contoh,$e_i' = [x_{i1} - \bar{x}_i, \dots, x_{in} - \bar{x}_i]$ - lies in the hypewr plane generate by $e_1, \dots, e_{i-1}, e_{i+1}, e_p$.

$$\begin{bmatrix} y'_1 - \bar{x}_1 1' \\ \vdots \\ y'_p - \bar{x}_p 1' \end{bmatrix} = \begin{bmatrix} x_{11} - \bar{x}_1 & \dots & x_{1n} - \bar{x}_1 \\ \vdots & \ddots & \vdots \\ x_{p1} - \bar{x}_p & \dots & x_{pn} - \bar{x}_p \end{bmatrix}$$

$$= \mathrm{X} - \bar{\mathrm{x}}1'$$

Hasil 3.2 Variansi adalah nol jika dan hanya jika sekurang-kurangnya satu vector selisih lies in oleh semua kombinasi linear yang lainnya; yaitu, ketika baris dari matriks selisih dalam (3-18) adalah tak bebas linear.

Bukti. Jika baris dari matriks selisih (X - $\bar{x}1'$) adalah tak bebas linear, ada kombinasi linear dari baris tersebut atau, ekuivalen, dengan kolom dari (X - $\bar{x}1'$)$'$ sedemikian sehingga

$$0 = \ell\colon kolom1(\mathrm{X} - \bar{x}1')' + \dots + \ell_p kolomp(\mathrm{X} - \bar{x}1')'$$

$$= (\mathrm{X} - \bar{x}1')'\ell \text{ untuk } \ell \neq 0$$

Tapi sebagai pengujian, $(n-1)\mathrm{S}\,\ell = (\mathrm{X} - \bar{x}1') \quad (\mathrm{X} - \bar{x}1')'$ dan

$$(n-1)\mathrm{S}\,\ell = (\mathrm{X} - \bar{x}1') \quad (\mathrm{X} - \bar{x}1')'\ell = 0$$

Jadi untuk ℓ yang sama cocok untuk tak bebas linear, $\ell_1 kolom\ 1$(S)$+ \dots +$ $\ell_p\, kolom\ p$(S)=S $\ell = 0$, di dalam kolom dari S. Dengan hasil pada pasal 2A.9,$|\mathrm{S}| = 0$.

Dalam penjelasan lain, jika $|\mathrm{S}| = 0$ ada sebuah kombinasi linear $\mathrm{S}\ell$ kolom dari S sehingga $\mathrm{S}\ell = 0$. Yaitu, $0 = (n-1)\mathrm{S}\,\ell = (\mathrm{X} - \bar{x}1') \quad (\mathrm{X} - \bar{x}1')'\ell$. Kalikan dengan hasil dari ℓ'

$$0 = \ell'(\mathrm{X} - \bar{x}1') \quad (\mathrm{X} - \bar{x}1')'\ell = L^2_{(\mathrm{X} - \bar{x}1')'\ell}$$

Dan, untuk akhirya sama dengan nol, kita harus memiliki $(\mathrm{X} - \bar{x}1')'\ell = 0$. Lalu kolom dari $(\mathrm{X} - \bar{x}1')'$, ekuivalen, dengan baris dari $(\mathrm{X} - \bar{x}1')$ adalah tak bebas linear.

Contoh 3.8

Tunjukkan bahwa $|\mathrm{S}| = 0$ untuk

$$X = \begin{bmatrix} 1 & 4 & 4 \\ 2 & 1 & 0 \\ 5 & 6 & 4 \end{bmatrix}$$

Disini $\bar{x}' = [3 \quad 1 \quad 5]$ jadi

$$X - \bar{x}1' = \begin{bmatrix} 1-3 & 4-3 & 4-3 \\ 2-1 & 1-1 & 0-1 \\ 5-5 & 6-5 & 4-5 \end{bmatrix} = \begin{bmatrix} -2 & 1 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix}$$

Vektor residualnya $e'_1 = [-2 \quad 1 \quad 1]$, $e'_2 = [1 \quad 0 \quad -1]$, $e'_3 = [0 \quad 1 \quad -1]$. Karena $e_3 = e_1 + 2e_2$, terdapat baris degenerasi. Ini artinya bahwa salah satu beris vector residual, sebagai contoh e_3, bisa ditempatkan diantara baris yang lainnya. Akibatnya, $|S| = 0$

$$S = \begin{bmatrix} 3 & -\frac{3}{2} & 0 \\ -\frac{3}{2} & 1 & \frac{1}{2} \\ 0 & \frac{1}{2} & 1 \end{bmatrix}$$

Lalu

$$|S| = 3\begin{vmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{vmatrix}(-1)^2 + \left(-\frac{3}{2}\right)\begin{vmatrix} -\frac{3}{2} & \frac{1}{2} \\ 0 & 1 \end{vmatrix}(-1)^3 + (0)\begin{vmatrix} -\frac{3}{2} & 1 \\ 0 & \frac{1}{2} \end{vmatrix}(-1)^4$$

$$= 3\left(1 - \frac{1}{4}\right) + \left(\frac{3}{2}\right)\left(-\frac{3}{2} - 0\right) + 0 = \frac{9}{4} - \frac{9}{4} = 0$$

Hasil 3.3 Jika $n \leq p$, yaitu, ukuran sampel $\leq$ banyaknya variable, maka $|S| = 0$ untuk semua sampel.

Variansi Umum yang Dinotasikan dengan $|\mathbf{R}|$ dan Interpretasi Geometri.

Variansi sampel umum dari variable standar = $|R|$

$$\frac{(y'_i - \bar{x}_i 1')}{\sqrt{s_{ii}}} = \begin{bmatrix} \frac{x_{i1} - \bar{x}_i}{\sqrt{s_{ii}}} \\ \vdots \\ \frac{x_{in} - \bar{x}_i}{\sqrt{s_{ii}}} \end{bmatrix}, \qquad i = 1, 2, \dots, p$$

Nilai $|S|$ dan $|R|$ dihubungkan dengan persamaan

$$|S| = (s_{11} s_{22} \dots s_{pp})|R|$$

Contoh 3.9

Misalkan

$$S = \begin{bmatrix} 4 & 3 & 1 \\ 3 & 9 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

Lalu kita dapatkan $s_{11} = 4, s_{22} = 9, s_{33} = 1$.Selain itu

$$R = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 & \frac{2}{3} \\ \frac{1}{2} & \frac{2}{3} & 1 \end{bmatrix}$$

$$|S| = 4\begin{vmatrix} 9 & 2 \\ 2 & 1 \end{vmatrix}(-1)^2 + 3\begin{vmatrix} 3 & 2 \\ 1 & 1 \end{vmatrix}(-1)^3 + 1\begin{vmatrix} 3 & 9 \\ 1 & 2 \end{vmatrix}(-1)^4$$

$$= 4(9-4) - 3(3-2) + 1(6-9) = 14$$

$$|R| = 1\begin{vmatrix} 1 & \frac{2}{3} \\ \frac{2}{3} & 1 \end{vmatrix}(-1)^2 + \frac{1}{2}\begin{vmatrix} \frac{1}{2} & \frac{2}{3} \\ \frac{1}{2} & 1 \end{vmatrix}(-1)^3 + \frac{1}{2}\begin{vmatrix} \frac{1}{2} & 1 \\ \frac{1}{2} & \frac{2}{3} \end{vmatrix}(-1)^4$$

$$= \left(1 - \frac{4}{9}\right) - \left(\frac{1}{2}\right)\left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{2}\right)\left(\frac{1}{3} - \frac{1}{2}\right) = \frac{7}{18}$$

Dari informasi di atas, kita lihat bahwa

$$14 = |S| = (s_{11}s_{22} \dots s_{pp})|R| = (4)(9)(1)\left(\frac{7}{18}\right) = 14$$

Variansi Umum Lainnya

Kita definisikan variansi sampel total sebagai penjumlahan dari elemem-elemen yang terletak pada diagonal dari varians-kovarians sampel matriks S. Lalu

Total variansi sampel = $s_{11} + s_{22} + \cdots + s_{pp}$

Contoh 3.10

Dari contoh 3.8

$$S = \begin{bmatrix} 3 & -\frac{3}{2} & 0 \\ -\frac{3}{2} & 1 & \frac{1}{2} \\ 0 & \frac{1}{2} & 1 \end{bmatrix}$$

Jadi total variansi sampelnya = $s_{11} + s_{22} + s_{33} = 3 + 1 + 1 = 5$

Mean Sampel,Kovarians, dan Korelasi sebagai Operasi Matriks.

Kita telah mengembangkan gambaran geometri dari data matriks X dan mendapat statistic deskriptif $\bar{x}$ dan S. Dalam penambahannya, tidaklah mungkin untuk menghubungkan secara aljabar perhiutungan dari $\bar{x}$ dan S secara langsung kepada X menggunakan operasi matriks. Hasilnya, yang mana menggambarkan hubungan antara $\bar{x}$, S, dan kumpulan data lengkap X secara ringkas, adalah mudah memrogramkan ke dalam computer.

Kita memiliki bahwa $\bar{x}_i = \frac{x_{i1}1 + x_{i2}1 + \cdots + x_{in}1}{n} = \frac{y'_i 1}{n}$. Oleh karena itu,

$$\bar{x} = \begin{bmatrix} \bar{x}_1 \\ \vdots \\ \bar{x}_p \end{bmatrix} = \begin{bmatrix} \frac{y'_1 1}{n} \\ \vdots \\ \frac{y'_p 1}{n} \end{bmatrix} = \frac{1}{n} \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{p1} & \cdots & x_{pn} \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

$$\bar{x} = \frac{1}{n} X1$$

Itu adalah, $\bar{x}$ dihitung dari data matriks dengan menempatkan perkalian dengan vector 1 dan lalu mengalikan hasilnya dengan konstanta $1/n$.

Selanjutnya, kita membuat matriks $p \times n$ dengan maksud mengalikan kedua ruas dengan $1'$; yaitu,

$$\bar{x}1' = \frac{1}{n} X11' = \begin{bmatrix} \bar{x}_{11} & \cdots & \bar{x}_{11} \\ \vdots & \ddots & \vdots \\ \bar{x}_{p1} & \cdots & \bar{x}_p \end{bmatrix}$$

Kurangi matriks X dengan hasil di atas

$$X - \frac{1}{n} X11' = \begin{bmatrix} x_{11} - \bar{x}_1 & \cdots & x_{1n} - \bar{x}_1 \\ \vdots & \ddots & \vdots \\ x_{p1} - \bar{x}_p & \cdots & x_{pn} - \bar{x}_p \end{bmatrix}$$

Sekarang matriks $(n-1)S$ terwakili dengan mengalikan matriks di atas dengan transpose nya, atau

$$(n-1)S = \begin{bmatrix} x_{11}-\bar{x}_1 & \cdots & x_{1n}-\bar{x}_1 \\ \vdots & \ddots & \vdots \\ x_{p1}-\bar{x}_p & \cdots & x_{pn}-\bar{x}_p \end{bmatrix} \times \begin{bmatrix} x_{11}-\bar{x}_1 & \cdots & x_{p1}-\bar{x}_p \\ \vdots & \ddots & \vdots \\ x_{1n}-\bar{x}_n & \cdots & x_{pn}-\bar{x}_p \end{bmatrix}$$

$$= \left(X - \frac{1}{n}X11'\right)\left(X - \frac{1}{n}X11'\right)' = X\left(I - \frac{1}{n}11'\right)X'$$

karena

$$\left(I - \frac{1}{n}11'\right)\left(I - \frac{1}{n}11'\right)' = I - \frac{1}{n}11' - \frac{1}{n}11' + \frac{1}{n^2}11'11' = I - \frac{1}{n}11'$$

Untuk meringkaskan, lambing matriks yang menghubungkan $\bar{x}$ dan S dengan kumpulan data X adalah

$$\bar{x} = \frac{1}{n}X1$$

$$S = \frac{1}{n-1}X\left(I - \frac{1}{n}11'\right)X'$$

Hasil untuk S_n adalah sama kecuali $1/n$ digantikan $1/(n-1)$ sebagai factor awal.

Hubungan di dalam persamaan matriks di atas menunjukkan dengan jelas bagaimana operasi didalam matriks pada data matriks X menuju ke $\bar{x}$ dan S.

Sekali S dihitung, ini dapat dihubungkan dengan korelasi sampel matriks R. Hasilnya dapat juga “dibalikkan” agar hubungannya R ke S. Kita pertama-tama mendefinisikan matriks $p \times p$ sebagai standar deviasi matriks $D^{1/2}$ lalu hitung invers nya, $\left(D^{1/2}\right)^{-1} = D^{-1/2}$.

Misalkan

$$D^{1/2} = \begin{bmatrix} \sqrt{s_{11}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sqrt{s_{pp}} \end{bmatrix}$$

Lalu

$$D^{-1/2} = \begin{bmatrix} \frac{1}{\sqrt{s_{11}}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{\sqrt{s_{pp}}} \end{bmatrix}$$

Karena

$$S = \begin{bmatrix} s_{11} & \cdots & s_{1p} \\ \vdots & \ddots & \vdots \\ s_{1p} & \cdots & s_{pp} \end{bmatrix}$$

Dan

$$R = \begin{bmatrix} \frac{s_{11}}{\sqrt{s_{11}}\sqrt{s_{11}}} & \cdots & \frac{s_{1p}}{\sqrt{s_{11}}\sqrt{s_{pp}}} \\ \vdots & \ddots & \vdots \\ \frac{s_{1p}}{\sqrt{s_{11}}\sqrt{s_{pp}}} & \cdots & \frac{s_{pp}}{\sqrt{s_{pp}}\sqrt{s_{pp}}} \end{bmatrix} = \begin{bmatrix} 1 & \cdots & r_{1p} \\ \vdots & \ddots & \vdots \\ r_{1p} & \cdots & 1 \end{bmatrix}$$

Kita memiliki

$$R = D^{-1/2} S D^{-1/2}$$

Kalikan kedua ruas dengan $D^{1/2}$ dan karena $D^{-1/2}D^{1/2} = D^{1/2}D^{-1/2} = I$ menghasilkan

$$S = D^{1/2} R D^{1/2}$$

Oleh karena itu R bias kita dapatkan dari informasi dalam S, dimana S dapat diperoleh dari $D^{1/2}$ dan R.