B A B II

TEORI DASAR

2.1. Model regresi linier dengan matrik

2.1.1. Model umum regresi

Analisa regresi adalah suatu metoda untuk menganalisa data yang terdiri atas dua atau lebih variabel. Analisa mengenai hubungan antara dua variabel tersebut membutuhkan data yang terdiri dua kelompok hasil pengamatan, sehingga menghasilkan pasangan pengamatan sebanyak n yang dinyatakan dalam pasangan terurut (x_i, y_i) dimana $i = 1, 2, \ldots, n$. Sebagai contoh variabel X mungkin merupakan jumlah pupuk dalam kg/ha, sedangkan variabel Y merupakan hasil produksi dalam ton/ha. Jika pasangan pengamatan (x_i, y_i) digambarkan diatas kertas skala hitung maka akan diperoleh serangkaian titik-titik koordinat yang menghubungkan kedua hasil pengamatan yang dinamakan diagram pencar (scater diagram).

Model umum regresi linier adalah $y = x\beta + \varepsilon$ (2.1.1)

dimana,

y adalah variabel tak bebas (respon)

x adalah variabel bebas

β adalah parameter yang akan ditaksir

ε adalah kesalahan random

Secara umum, y dapat dihubungkan dengan p variabel regresor, sehingga didapat model,

$$y = \beta_o + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p + \varepsilon \qquad (2.1.2)$$

disebut model regresi linier berganda dengan p

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p + \varepsilon \qquad (2.1.2)$$

disebut model regresi linier berganda dengan p regresor. Parameter β_j, $j=0,1,2,\ldots,p$ disebut koefisien regresi, yang menggambarkan perubahan y perunit perubahan x_j, maka semua nilai variabel x_i $(i\neq j)$ dianggap konstan.

2.1.2. Estimasi parameter

Metode kuadrat terkecil digunakan untuk mengestimasi koefisien regresi pada (2.1.2).. Dianggap bahwa pengamatan $n>p$ dipenuhi, dan diambil y_i adalah variabel respon ke-i dan x_{ij} adalah pengamatan ke-i. Data digambarkan pada tabel 2.1. sebagai berikut :

Tabel 2.1 Data untuk regresi linier berganda

Pengamatan i	y	x_1	x_2	...	x_p
1	y_1	x_{11}	x_{12}	...	x_{1p}
2	y_2	x_{21}	x_{22}	...	x_{2p}
⋮	⋮	⋮	⋮		⋮
n	y_n	x_{n1}	x_{n2}	...	x_{np}

Diasumsikan ε_i pada model adalah $E(\varepsilon_i)=0$, juga ada asumsi tentang homokedastisitas yaitu setiap kesalahan random mempunyai variansi yang sama, $Var(\varepsilon_i)=E(\varepsilon_i^2)=\sigma^2$ untuk semua-i dan tidak berkorelasi artinya antara kesalahan random yang satu dengan yang lain saling bebas atau $Kov(\varepsilon_i,\varepsilon_j)=0$, untuk $i\neq j$.

Ditulis suatu model sesuai (2.1.2) adalah

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \varepsilon_i$$

$$= \beta_0 + \sum_{j=1}^{p} \beta_j x_{ij} + \varepsilon_i \qquad , i=1,2, \dots, n \qquad (2.1.3)$$

sehingga :

$$\varepsilon_i = y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_{ij} \quad , \text{ maka } \quad \varepsilon_i^2 = (y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_{ij})^2$$

fungsi kuadrat terkecilnya adalah :

$$S(\beta_0, \beta_1, \dots, \beta_p) = \sum_{i=1}^{n} \varepsilon_i^2$$

$$= \sum_{i=1}^{n} (y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_{ij})^2 \qquad (2.1.4)$$

Fungsi S dimimimumkan pada $\beta_0, \beta_1, \dots, \beta_p$

Estimator kuadrat terkecil dari $\beta_0, \beta_1, \dots, \beta_p$ harus memenuhi :

$$\left. \frac{\partial S}{\partial \beta_c} \right|_{\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p} = -2 \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \sum_{j=1}^{p} \hat{\beta}_j x_{ij}) = 0$$

(2.1.4a)

$$\left. \frac{\partial S}{\partial \beta_j} \right|_{\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p} = -2 \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \sum_{j=1}^{p} \hat{\beta}_j x_{ij}) x_{ij} = 0$$

$j = 1, 2, \dots, p$ (2.1.4b)

Persamaan normal kuadrat terkecil :

$$n \hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} x_{i1} + \hat{\beta}_2 \sum_{i=1}^{n} x_{i2} + \dots + \hat{\beta}_p \sum_{i=1}^{n} x_{ip} = \sum_{i=1}^{n} y_i$$

$$\hat{\beta}_0 \sum_{i=1}^{n} x_{i1} + \hat{\beta}_1 \sum_{i=1}^{n} x_{i1}^2 + \hat{\beta}_2 \sum_{i=1}^{n} x_{i1} x_{i2} + \dots + \hat{\beta}_p \sum_{i=1}^{n} x_{i1} x_{ip} = \sum_{i=1}^{n} x_{i1} y_i$$

$$\vdots \qquad \vdots \qquad \qquad \vdots \qquad \qquad \qquad \vdots \qquad \qquad \vdots$$

$$\hat{\beta}_0 \sum_{i=1}^{n} x_{ip} + \hat{\beta}_1 \sum_{i=1}^{n} x_{ip} x_{i1} + \hat{\beta}_2 \sum_{i=1}^{n} x_{ip} x_{i2} + \dots + \hat{\beta}_p \sum_{i=1}^{n} x_{ip}^2 = \sum_{i=1}^{n} x_{ip} y_i \qquad (2.1.5)$$

Solusi untuk menentukan koefisien-koefisien persamaan regresi dari $k=p+1$ persamaan-persamaan normal diatas adalah merupakan taksiran kuadrat terkecil untuk $\beta_0, \beta_1, \ldots, \beta_p$ yaitu $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_p$

2.1.3. Analisa regresi dengan matrik

Untuk menyelasaikan model regresi ganda dapat dengan notasi matrik. Karena dapat ditunjukkan secara tepat model, data dan hasilnya.

Dari persamaan (2.1.5), maka model pengamatan (2.1.1) dapat ditulis dalam notasi matrik : $Y = X\beta + \varepsilon$

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$$

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}, \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

dengan :

Y vektor berukuran (nx1) dari pengamatan

X matrik berukuran (nx(p+1)) variabel bebas

β vektor berukuran ((p+1)x1) koefisien regresi

ε vektor berukuran (nx1) kesalahan random

untuk mendapatkan suatu vektor estimator-estimator kuadrat terkecil, $\hat{\beta}$ dengan meminimumkan

$$S(\beta) = \sum_{i=1}^{n} \varepsilon_i^2 = \varepsilon^T \varepsilon$$
$$= (Y - X\beta)^T (Y - X\beta)$$
$$= Y^T Y - \beta^T X^T Y - Y^T X\beta + \beta^T X^T X\beta$$
$$= Y^T Y - 2\beta^T X^T Y + \beta^T X^T X\beta \qquad (2.1.6)$$

karena $\beta^T X^T Y$ adalah matrik berukuran (*1x1*) atau suatu skalar dan tranposenya $(\beta^T X^T Y)^T = Y^T X\beta$ adalah juga suatu skalar yang sama.

Estimator kwadrat terkecil harus memenuhi :

$$\left.\frac{\partial S}{\partial \beta}\right|_{\hat{\beta}} = -2X^T Y + 2X^T X\beta = 0$$

sehingga $X^T X \hat{\beta} = X^T Y$ (2.1.7)

persamaan (2.1.7) adalah persamaan normal kuadrat terkecil yang identik dengan (2.1.5). Sehingga dengan menggandakan kedua sisi dari (2.1.7) dengan invers dari $X^T X$ maka estimator kuadrat terkecil dari β adalah :

$$\hat{\beta} = (X^T X)^{-1} X^T Y \qquad (2.1.8)$$

dan ini menunjukkan bahwa $(X^T X)^{-1}$ ada. Matrik $(X^T X)^{-1}$ akan selalu ada jika variabel-variabel bebasnya adalah bebas linier, yaitu jika tidak ada suatu kolom dalam X yang merupakan kombinasi linier dari kolom-kolom yang lain.

Sehingga lebih jelasnya (2.1.7) dapat ditulis :

$$\begin{bmatrix} n & \Sigma x_{i1} & \Sigma x_{i2} & \dots \; \Sigma x_{ip} \\ \Sigma x_{i1} & \Sigma x_{i1}^2 & \Sigma x_{i1} x_{i2} & \dots \; \Sigma x_{i1} x_{ip} \\ \vdots & \vdots & \vdots & \vdots \\ \Sigma x_{ip} & \Sigma x_{ip} x_{i1} & \Sigma x_{i1} x_{i2} & \dots \; \Sigma x_{ip}^2 \end{bmatrix} \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_p \end{bmatrix} = \begin{bmatrix} \Sigma \; y_i \\ \Sigma \; x_{i1} \; y_i \\ \vdots \\ \Sigma \; x_{ip} \; y_i \end{bmatrix}$$

untuk $k=p+1$, terlihat bahwa $X^T X$ adalah matrik simetri berukuran (kxk) dan $X^T Y$ adalah vektor kolom berukuran (kx1). Struktur khusus dari matrik $X^T X$, elemen-elemen diagonal dari matrik $X^T X$ adalah jumlah kuadrat dari elemen-elemen dalam kolom X, dan elemen-elemen diluar diagonal adalah jumlah cross-produk dari kolom-kolom X.

Model regresi yang sesuai untuk variabel bebas $x_i^T = [\; 1 \quad x_1 \quad x_2 \; \dots \; x_p \;]$ adalah

$$\begin{aligned} \hat{y} &= x_i^T \hat{\beta} \\ &= \hat{\beta}_0 + \sum_{j=1}^{p} \beta_j \; x_j \end{aligned}$$

Vektor yang sesuai untuk nilai $\hat{y}_i$ dengan nilai pengamatan y_i adalah :

$$\begin{aligned} \hat{Y} &= X \hat{\beta} \\ &= X \, (X^T X)^{-1} X^T \; Y \\ &= P \; Y \end{aligned} \qquad (2.1.9)$$

$P = X(X^T X)^{-1} X^T$ adalah matrik berukuran (nxn) disebut matrik prediksi, karena matrik tersebut memetakan vektor nilai pengamatan y_i menjadi vektor nilai yang cocok $\hat{y}_i$.

Residual dituliskan dengan notasi matrik, merupakan selisih antara nilai pengamatan y_i dengan nilai yang cocok $\hat{y}_i$. Dinotasikan e_i, dimana $e_i = y_i - \hat{y}_i$ sehingga *n* residual dapat ditulis dalam notasi matrik :

$$\begin{aligned} e &= Y - \hat{Y} \\ &= Y - X\hat{\beta} \\ &= Y - X(X^T X)^{-1} X^T Y \\ &= Y - PY \\ &= (I - P)\,Y \end{aligned}$$

maka (I - P) disebut matrik residual sebab mengaplikasikan pada Y menghasilkan residual biasa.

2.2. Matrik Prediksi

Model regresi linier umum :

$$Y = X\beta + \varepsilon, \tag{2.2.1}$$

dengan asumsi pada bagian 2.1.4, dan dari bagian 2.1.5 bahwa matriks $P = X(X^T X)^{-1} X^T$ (2.2.2) menentukan beberapa standart hasil kuadrat terkecil.

Matrik P mempunyai peran penting pada analisa regresi linier dan teknik analisa multivariat yang lain. Dan memiliki sifat-sifat penting, yang digunakan untuk memperoleh beberapa hasil pada pembahasan tentang pendeteksian data berpengaruh. Disebut matrik Prediksi sebab tranformasi matriks itu, dengan mengaplikasikan pada Y, menghasilkan nilai terprediksi $\hat{Y}$.

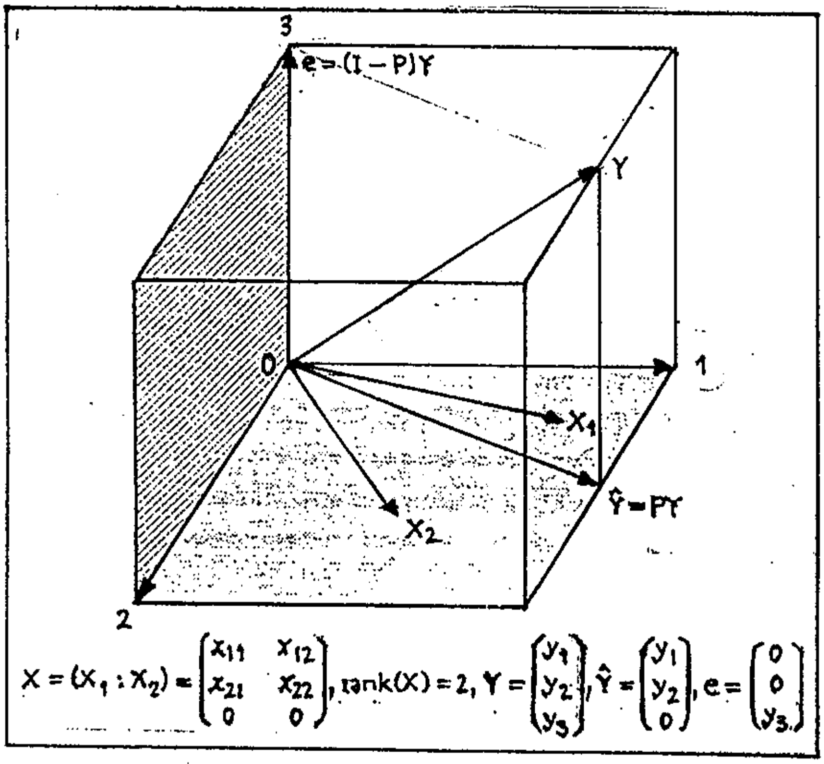
2.1.1 Peranan P dan (I-P) pada regresi linier.

Diambil $\mathcal{R}$ sebagai ruang euclidean dimensi-n dan $\mathcal{R}_X$ sebagai subruang spanned dengan kolom X berdimensi p. Matrik prediksi P adalah simetri dan idempoten.

Kemudian Y diproyeksikan secara ortogonal (garis tegak) pada $\mathcal{R}_X$. Alasan ini kadang-kadang disebut matriks proyeksi atau proyektor orthogonal pada $\mathcal{R}_X$ (bayangan dari P). Matrik residual (I-P) juga simetri dan idempoten. Ini merupakan proyeksi orthogonal pada bayangan (I-P), suatu subruang orthogonal untuk $\mathcal{R}_X$. Maka P dan (I-P) adalah proyektor orthogonal.

Contoh 2.1.

Suatu kasus Y diregresikan, melalui titik asal, pada dua prediktor, X_1 dan X_2, dengan $n=3$. Pada gambar 2.1 kita melihat bahwa vektor prediksi nilai $\hat{Y}$= PY adalah proyeksi orthogonal (garis tegak) dari Y pada subruang spanned dimensi-2 oleh X_1 dan X_2. Dapat juga dilihat pada gambar 2.1, bahwa vektor residual e=(I-P)Y adalah proyeksi orthogonal dari Y pada penambahan subruang dimensi satu yaitu orthogonal pada X_1 dan X_2. Juga dapat dilihat e tegak lurus pada $\hat{Y}$ ($\hat{Y}$ berada pada subruang spanned oleh $(X_1 : X_2)$ dan e berada pada komplemen orthogonalnya).



gambar 2.1

contoh $\hat{Y}$ adalah proyeksi orthogonal dari Y pada ruang spanned oleh X_1 dan X_2

P proyeksi orthogonal vektor dimensi-n pada subruang spanned dimensi-p melalui kolom X, P memproyeksikan X pada dirinya sendiri, maka PX=X. Sehingga (I-P)X=0. Maka, vektor residual e adalah :

$$e = (I-P)Y = (I-P)(X\beta+\varepsilon) = (I-P)\varepsilon \qquad (3.5)$$

maka hubungan antaran e dan ε tergantung hanya pada P.

Diambil p_{ij} sebagai elemen ke-ij dari P. Maka elemen diagonal dari P, adalah

$$p_{ii} = x_i^T(X^TX)^{-1}x_i, \quad i=1,2,\ldots,n \qquad (2.2.3)$$

Elemen ke-ij dari P ialah

$$p_{ij}=x_i^T(X^TX)^{-1}x_j, \qquad i,j=1,2,\ldots,n. \qquad (2.2.4)$$

2.2.1. Sifat-sifat matrik prediksi

Sifat 2.1

Matrik P dan (I-P) adalah simetri dan idempoten

bukti :

P matrik simetri

$$\begin{aligned} P^T &= (X(X^TX)^{-1}X^T)^T \\ &= X((X^TX)^{-1})^TX^T \\ &= X(X^TX)^{-1}X^T \\ &= P \end{aligned}$$

P matrik idempoten

$$\begin{aligned} P.P &= (X(X^TX)^{-1}X^T)\ (X(X^TX)^{-1}X^T) \\ &= (X(X^TX)^{-1}(X^TX)(X^TX)^{-1}X^T) \\ &= (X\ I\ (X^TX)^{-1}X^T) \\ &= (X\ (X^TX)^{-1}X^T) \\ &= P \end{aligned}$$

(I-P) matrik simetri

$$\begin{aligned} (I-P)^T &= (I^T - P^T) \\ &= (I - P) \end{aligned}$$

(I-P) matrik idempoten

$$(I-P)(I-P) = I.I - I.P - P.I + P.P$$
$$= I - P - P + P$$
$$= I - P$$

sifat 2.2

Matrik X berukuran (nxk), maka tr(P)=rank(P)=k

bukti :

X^TX berukuran (nxk) dimana $n>k$, dengan asumsi ke-2 maka rank(X)=k, maka X^TX adalah matrik definit positif berukuran (kxk), dan karena eigenvalue dari matrik definit positif seluruhnya adalah positif maka X^TXmerupakan matrik nonsingular, jadi rank(X^TX)=k, demikian juga rank$(X^TX)^{-1}$=k, karena perkalian suatu matrik dengan suatu matrik yang nonsingular tidak merubah ranknya, maka

$$\text{rank}(P)=\text{rank}(X(X^TX)^{-1}X)$$
$$=\text{rank}(X^TX)=k$$

Matrik P adalah simetri dan idempoten. Berdasar sifat matrik idempoten maka eigenvalue dari matrik P adalah 0 atau 1. Karena P matrik simetris maka rank(P) sama dengan jumlah eigenvaluenya yang tidak sama dengan 0, maka jumlah eigenvalue matrik P sama dengan rank(P)=k, karena P adalah matrik simetris berukuran n, dengan eigenvalue λ_i (i=1,2,...,n) maka $\text{tr}(P)=\sum_{i=1}^{n}\lambda_i=\text{rank}(P)$

Lemma 2.1.

Invers suatu matrik terpartisi.

misalkan A matrik partisi sebagai berikut:

$$A=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

(a) Jika A dan A_{11} adalah nonsingular, maka

$$A^{-1} = \begin{bmatrix} A_{11}^{-1}+A_{11}^{-1}A_{12}MA_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}M \\ -MA_{21}A_{11}^{-1} & M \end{bmatrix} \quad (2.2.5a)$$

dimana $M = (A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1}$

(b) Jika A dan A_{22} adalah nonsingular, maka

$$A^{-1}=\begin{bmatrix} N & -N\,A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{12}N & A_{22}^{-1}+A_{22}^{-1}A_{21}NA_{12}A_{22}^{-1} \end{bmatrix} \quad (2.2.5b)$$

dimana $N = (A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1}$

bukti :

Dibuktikan dengan menunjukkan bahwa $A^{-1}A = I$

(a). $$AA^{-1}=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} A_{11}^{-1}+A_{11}^{-1}A_{12}MA_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}M \\ -MA_{21}A_{11}^{-1} & M \end{bmatrix}$$

$$=\begin{bmatrix} A_{11}(A_{11}^{-1}+A_{11}^{-1}A_{12}MA_{21}A_{11}^{-1})+A_{12}(-MA_{21}A_{11}^{-1}) & A_{11}(-A_{11}^{-1}A_{12}M)+A_{12}M \\ A_{21}(A_{11}^{-1}+A_{11}^{-1}A_{12}MA_{21}A_{11}^{-1})+A_{22}(-MA_{21}A_{11}^{-1}) & A_{21}(-A_{11}^{-1}A_{12}M)+A_{22}M \end{bmatrix}$$

$$=\begin{bmatrix} I+A_{12}MA_{21}A_{11}^{-1}-A_{12}MA_{21}A_{11}^{-1} & -A_{12}M+A_{12}M \\ A_{21}\ A_{11}^{-1}+A_{21}A_{11}^{-1}A_{12}MA_{21}A_{11}^{-1}-A_{22}MA_{21}A_{11}^{-1} & -A_{21}A_{11}^{-1}A_{12}M+A_{22}M \end{bmatrix}$$

$$=\begin{bmatrix} I & 0 \\ (I-(A_{22}-A_{21}A_{11}^{-1}A_{12})M)A_{21}A_{11}^{-1} & (A_{22}-A_{21}A_{11}^{-1}A_{12})M \end{bmatrix}$$

$$=\begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

(b). $$AA^{-1}=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} N & -NA_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}N & A_{22}^{-1}+A_{22}^{-1}A_{21}NA_{12}A_{22}^{-1} \end{bmatrix}$$

$$=\begin{bmatrix} A_{11}N-A_{12}A_{22}^{-1}A_{21}N & -A_{11}NA_{12}A_{22}^{-1}+A_{12}(A_{22}^{-1}+A_{22}^{-1}A_{21}NA_{12}A_{22}^{-1}) \\ A_{21}N-A_{22}A_{22}^{-1}A_{21}N & -A_{21}NA_{12}A_{22}^{-1}+A_{22}(A_{22}^{-1}+A_{22}^{-1}A_{21}NA_{12}A_{22}^{-1}) \end{bmatrix}$$

$$=\begin{bmatrix} (A_{11}-A_{12}A_{22}^{-1}A_{21})N & (-(A_{11}-A_{12}A_{22}^{-1}A_{12})N+I)A_{12}A_{22}^{-1}) \\ A_{21}N-A_{21}N & -A_{21}NA_{12}A_{22}^{-1}+I+A_{21}NA_{12}A_{22}^{-1} \end{bmatrix}$$

$$=\begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

Sifat 2.3.

Diketahui $X=(X_1:X_2)$, dimana

X_1 matrik berukuran (nxr) dengan $\text{rank}(X_1)=r$

X_2 matrik berukuran $(nx(r-1))$ dengan $\text{rank}(X_2)=k-r$

diketahui $P_1=X_1(X_1^TX_1)^{-1}X_1^T$ matrik prediksi untuk P_1

$W=(I-P_1)X_2$ adalah proyeksi dari X_2 pada komplemen orthogonal X_1

dan $P_2=W(W^TW)^{-1}W^T$ adalah matrik prediksi untuk W.

maka P dapat ditulis sebagai :

$$X(X^TX)^{-1}X^T=X_1(X_1^TX_1)^{-1}X_1^T+(I-P_1)X_2\{X_2^T(I-P_1)X_2^T\}^{-1}X_2^T(I-P_1) \quad (2.2.6a)$$

atau

$$P = P_1 + P_2 \tag{2.2.6b}$$

bukti

P dapat ditulis dalam bentuk

$$P=(X_1 \;\; X_2)\begin{bmatrix} X_1^TX_1 & X_1^TX_2 \\ X_2^TX_1 & X_2^TX_2 \end{bmatrix}^{-1}\begin{bmatrix} X_1^T \\ X_2^T \end{bmatrix} \tag{2.2.7a}$$

Dengan mengunakan (2.2.7a)

untuk menghitung invers dari (X^TX) bentuk dipartisikan, didapat :

$$(X^TX)^{-1}=\begin{bmatrix} (X_1^TX_1)^{-1}+(X_1^TX_1)^{-1}X_1^TX_2MX_2^TX_1(X_1^TX)^{-1} & -(X_1^TX_1)^{-1}X_1^TX_2M \\ -MX_2^TX_1(X_1^TX_1)^{-1} & M \end{bmatrix}$$

(2.2.7b)

$$\begin{aligned} \text{dimana } M &= (X_2^TX_2-X_2^TX_1(X_1^TX_1)^{-1}X_1^TX_2)^{-1} \\ &= (X_2^T(I-X_1(X_1^TX_1)^{-1}X_1^T)X_2)^{-1} \\ &= (X_2^T(I-P_1)X_2)^{-1} \end{aligned}$$

Dengan mensubstitusikan (2.2.7b) kedalam (2.2.6a), didapat

$$\begin{aligned} P&=P_1+P_1X_2MX_2^TP_1-P_2X_2MX_2^T-X_2MX_2^TP_1+X_2MX_2^T \\ &=P_1+(I-P_1)X_2MX_2^T(I-P) \\ &=P_1+(I-P_2)X_2(X_2^T(I-P)X_2)^{-1}X_2^T(I-P) \\ &=P_1+P_2 \end{aligned}$$

sifat 2.3 diatas menunjukkan bahwa matrik prediksi P dapat dikomposisikan menjadi dua (atau lebih) matrik prediksi.

Sifat 2.4.

Untuk $i=1,2,...,n$ dan $j=1,2,...,n$ maka

a. $0 \le p_{ii} \le 1$, untuk semua i

b. $-0.5 \le p_{ij} \le 0.5$, untuk semua $j \ne i$

bukti :

dengan sifat 2.1, elemen diagonal ke-i dari P dapat ditulis $p_{ii} = \sum_{j=1}^{n} p_{ij}^2 = p_{ii}^2 + \sum_{j \ne i} p_{ij}^2$ (2.2.8a)

mengikuti persamaan tersebut maka $0 \le p_{ii} \le 1$, untuk semua i maka (a) terbukti

(2.2.8a) dapat juga ditulis sebagai

$$p_{ii} = p_{ii}^2 + p_{ij}^2 + \sum_{r \ne i,j} p_{ir}^2 \qquad (2.2.8b)$$

oleh karena $0 \le p_{ii} \le 1$ maka didapat $-0.5 \le p_{ij} \le 0.5$

maka (b) terbukti.

sifat 2.5.

Untuk $i=1,2,...,n$ dan $j=1,2,...,n$ maka

a. jika $p_{ii}=0$ atau 1, maka $p_{ij}=0$

b. $$p_{ii} + \frac{e_i^2}{e^T e} \le 1$$

bukti :

Dari persamaan (2.2.8a) jelas bahwa $p_{ii}=0$ atau 1, untuk semua $j \ne i$ maka $p_{ij}=0$

maka (a) terbukti

hal itu menunjukkan bahwa jika p_{ii} besar (mendekati 1) atau kecil (mendekati 0), maka p_{ii} kecil untuk semua $j \ne i$.

Didefinisikan $Z=(X:Y)$, $P_x=X(X^TX)^{-1}X^T$ dan $P_z=Z(Z^TZ)^{-1}Z^T$. Berdasarkan (3.8a) didapat

$$P_z = P_x + \frac{(I-P_x)YY^T(I-P_x)}{Y^T(I-P_x)Y}$$

$$= P_x + \frac{ee^T}{e^Te}$$

maka elemen-elemen diagonal $P_z \leq 1$, maka (b) terbukti. hal itu menunjukkan bahwa untuk p_{ii} lebih besar, residual biasa ke-i (e_i), lebih kecil.

2.2.3. Penghapusan suatu data pengamatan

Lemma 2.2.

Untuk A dan D matrik nonsingular masing-masing berderajat k dan m, dan B matrik berukuran (kxm) dan C(kxm) maka dibuktikan bahwa inversnya ada.

$$(A + B\ D\ C^T)^{-1}=A^{-1} - A^{-1}B(D^{-1}+ C^TA^{-1}B)^{-1}C^TA^{-1} \qquad (2.2.9)$$

bukti :

dengan menunjukkan bahwa produk ruas kiri dan ruas kanan adalah matrik identitas.

$$(A + B\ D\ C^T)\{A^{-1} - A^{-1}B(D^{-1}+ C^TA^{-1}B)^{-1}C^TA^{-1}\}$$
$$=I-B(D^{-1}+C^TA^{-1}B)^{-1}C^TA^{-1}+BDC^TA^{-1}-BDC^TA^{-1}B(D^{-1}+C^TA^{-1}B)^{-1}C^TA^{-1}$$
$$=B\{I-(D^{-1}+C^TA^{-1}B)^{-1}C^TA^{-1}B+DC^TA^{-1}B-DC^TA^{-1}B(D^{-1}+C^TA^{-1}B)^{-1}C^TA^{-1}B\}B^{-1}$$
$$=B(I+DC^TA^{-1}B)\{I-(D^{-1}+C^TA^{-1}B)^{-1}C^TA^{-1}B\}B^{-1}$$
$$=BD(D^{-1}+C^TA^{-1}B)\ \{D-(D^{-1}+C^TA^{-1}B)^{-1}C^TA^{-1}BD\}D^{-1}B^{-1}$$
$$=BD\{(D^{-1}+C^TA^{-1}B)D - C^TA^{-1}BD\}D^{-1}B^{-1}$$
$$=BD(I+C^TA^{-1}BD - C^TA^{-1}BD)D^{-1}B^{-1}$$
$$=BDID^{-1}B^{-1} = I$$

Pada lemma 2.2 disubstitusi $B=x_i$, $C^T=x_i^T$, $D=1$ dan $A=(X^T_{(i)}X_{(i)})$, dimana $X_{(i)}$ adalah notasi matrik X dengan baris ke-*i* dihapus. Persamaan 3.11 didapat

$$(X^TX)^{-1} = (X^T_{(i)}X_{(i)} + x_ix_i^T)^{-1}$$

$$= (X^T_{(i)}X_{(i)})^{-1} + \frac{(X^T_{(i)}X_{(i)})^{-1}x_ix_i^T\ (X^T_{(i)}X_{(i)})^{-1}}{1-x_i^T(X^T_{(i)}X_{(i)})^{-1}x_i}$$

Untuk $A=X^TX$, $B=x_i$, $C^T=x_i^T$ dan $D=-1$,

$$(X^T_{(i)}X_{(i)})^{-1}=(X^TX - x_ix_i^T)^{-1}$$

$$=(X^TX)^{-1}+\frac{(X^TX)^{-1}x_ix_i^T\ (X^TX)^{-1}}{1-x_i^T(X^TX)^{-1}x_i}$$

untuk *n* pengamatan, dinotasikan X, dan penghapusan satu pengamatan ke-*i*, x_i^T.

Untuk $\beta_{(i)}=(X_{(i)}X^T_{(i)})\quad X^{-1}_{(i)}Y^T_{(i)}$

adalah estimasi dari β untuk pengamatan ke-*i* dihapus, didapat

$$\beta_{(i)}=\left[(X^TX)^{-1}+\frac{(X^TX)^{-1}x_ix_i^T\ (X^TX)^{-1}}{1-x_i^T(X^TX)^{-1}x_i}\right]X^T_{(i)}Y_{(i)}$$

$$=\left[(X^TX)^{-1}+\frac{(X^TX)^{-1}x_ix_i^T\ (X^TX)^{-1}}{1-p_{ii}}\right]X^T_{(i)}Y_{(i)}\quad (2.2.10a)$$

disamping itu juga diperoleh bahwa :

$$\hat{\beta} = (X^T X)^{-1} X^T Y$$

$$= (X^T X)^{-1} (X^T_{(i)} : x_i) \begin{bmatrix} Y_{(i)} \\ y_i \end{bmatrix}$$

$$= (X^T X)^{-1} \; X^T_{(i)} Y_{(i)} + (X^T X)^{-1} \; x_i y_i$$

jadi bisa ditulis

$$(X^T X)^{-1} \; X^T_{(i)} Y_{(i)} = \hat{\beta} - (X^T X)^{-1} \; x_i y_i \qquad (2.2.10b)$$

dengan nensubstitusikan (2.2.10b) ke (2.2.10a) dan menyederhanakannya diperoleh

$$\hat{\beta}_{(i)} = (X^T X)^{-1} \; X^T_{(i)} Y_{(i)} + \frac{(X^T X)^{-1} x_i x_i^T \; (X^T X)^{-1}}{1 - p_{ii}} \; X^T_{(i)} Y_{(i)}$$

$$= \hat{\beta} - (X^T X)^{-1} \; x_i y_i + \frac{(X^T X)^{-1} x_i x_i^T \; (X^T X)^{-1}}{1 - p_{ii}} \; X^T_{(i)} Y_{(i)}$$

$$\hat{\beta} - \hat{\beta}_{(i)} = (X^T X)^{-1} \; x_i y_i - \frac{(X^T X)^{-1} x_i x_i^T \; (X^T X)^{-1}}{1 - p_{ii}} \; X^T_{(i)} Y_{(i)}$$

$$= (X^T X)^{-1} \; x_i \left[y_i - \frac{x_i^T \; (X^T X)^{-1}}{1 - p_{ii}} \; X^T_{(i)} Y_{(i)} \right]$$

$$= (X^T X)^{-1} x_i \left[y_i - \frac{x_i^T \; \{\hat{\beta} - (X^T X)^{-1} \; x_i y_i\}}{1 - p_{ii}} \right]$$

$$= \frac{(X^T X)^{-1} x_i}{1 - p_{ii}} \left[y_i - x_i^T (X^T X)^{-1} x_i y_i - x_i^T \; \hat{\beta} + x_i^T (X^T X)^{-1} x_i y_i \right]$$

$$= (X^T X)^{-1} x_i \left[\frac{y_i - x_i^T \hat{\beta}}{1 - p_{ii}} \right]$$

$$= (X^T X)^{-1} x_i \frac{e_i}{1 - p_{ii}} \qquad (2.2.11)$$

2.2.4. Eigenvalue matrik P

Sifat 2.6.

Eigenvalue dari P adalah 0 atau 1

bukti :

secara jelas dari kenyataan bahwa eigenvalue dari matrik idempoten adalah 0 atau 1

sifat 2.7.

Terdapat $(n-k)$ eigenvalue dari P sama dengan 0, maka sisanya k eigenvalue sama dengan 1.

bukti

diambil λ_i, $i=1,2,...,n$ sebagai eigenvalue dari P. maka dengan sifat 2.2 didapat $\sum_{i=1}^{n} \lambda_i = tr(P)=k$

tetapi dengan sifat 2.6, $\lambda_i=0$ atau 1 untuk semua i. Selanjutnya, k eigenvalue dari P sama dengan 1 dan sisanya $(n-k)$ eigenvalue sama dengan 0.

Lemma 2.3.

Determinan matrik terpartisi. misalkan matrik S terpartisi sebagai : $S = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$

(a). Jika A nonsingular, maka $\det(S)=\det(A)\det(D-CA^{-1}B)$

(b). Jika D nonsingular, maka $\det(S)=\det(D)\det(A-BD^{-1}C)$

bukti :

(a). $\det(S)=\det(AD-BC)$

$=\det(A(D-CA^{-1}B)$

$=\det(A)\det(D-CA^{-1}B)$

(b). $\det(S)=\det(AD-BC)$

$=\det((A-BD^{-1}C)D)$

$=\det(A-B^{-1}C)\det(D)$

$=\det(D)\det(A-B^{-1}C)$

Lemma 2.4.

misalkan B dan C adalah matrik berukuran (kxm). Jika A matrik nonsingular berukuran (kxk), maka:

$\det(A-BC^T) = \det(A)\ \det(I-C^TA^{-1}B)$

bukti :

dengan lemma 2.3(a), didapat

$$\det\begin{pmatrix} A & B \\ C^T & I \end{pmatrix} = \det(A)\det(I-C^TA^{-1}B)$$

dengan lemma 2.3(b), didapat

$$\det\begin{pmatrix} A & B \\ C^T & I \end{pmatrix} = \det(A-BC^T)$$

maka $\det(A-BC^T) = \det(A)\ \det(I-C^TA^{-1}B)$ ■

2.2.5. Contoh numerik

Contoh 2.2.

Suatu garis lurus pada himpunan data terdiri dari 5 titik, empat pada x=1 dan satu pada x=4. sehingga didapat

$$X^T = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 4 \end{bmatrix} \qquad X^T X = \begin{bmatrix} 5 & 8 \\ 8 & 20 \end{bmatrix}$$

$$(X^T X)^{-1} = \frac{1}{36} \begin{bmatrix} 20 & -8 \\ -8 & 5 \end{bmatrix} \qquad \text{dan}$$

$$P = X(X^T X)^{-1} X^T = \begin{bmatrix} 0.25 & 0.25 & 0.25 & 0.25 & 0 \\ 0.25 & 0.25 & 0.25 & 0.25 & 0 \\ 0.25 & 0.25 & 0.25 & 0.25 & 0 \\ 0.25 & 0.25 & 0.25 & 0.25 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

dengan sifat 2.2, didapat trace(P)=rank(P)=2 dan

$$\sum_{i=1}^{n} \sum_{j=1}^{n} P_{ij}^2 = 2$$

dari $P_{5,5}=1$, dengan sifat 2.5(a), maka semua $P_{5,j}=P_{j,5}=0$, untuk semua j=1,2,3,4 juga dari $P_{5,5}=1$.

Matrik terpartisi P dapat diambil sebagai :

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{bmatrix}$$

dengan

$$P_{11} = \begin{bmatrix} 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 \end{bmatrix} \qquad P_{12} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad P_{22} = 1$$

2.3. Asumsi-asumsi pada kuadrat terkecil

1. Asumsi linieritas

Asumsi ini secara implisit pada model (2.1.1) yang berarti setiap nilai respon terobservasi y_i. ditulis sebagai fungsi linier baris ke-i dari dari X yaitu x_i^T,

$$y_i = x_i^T \beta + \varepsilon_i, \qquad i = 1,2,\ldots,n \qquad (2.3.1)$$

2. Asumsi perhitungan

Untuk mendapatkan perkiraan unik dari β, syarat perlu $(X^TX)^{-1}$ ada, atau ekuivalen rank(X)=k

3. Asumsi distribusi

Analisa statistik berdasar pada kuadrat terkecil (contoh test-t, test-F) memenuhi:

a. X diukur tanpa error

b. ε_i tidak tergantung pada x_i^T, $i=1,2,\ldots,n$

c. $\varepsilon \sim N_n(0,\sigma^2 I)$

4. Asumsi implisit

Semua pengamatan dapat dipercaya dan seharusnya mempunyai aturan yang sama dalam menentukan hasil kuadrat terkecil dan pengaruh kesimpulan.

Sebelum menentukan kesimpulan dari analisa, penting untuk mengecek validnya asumsi tersebut.

2.4. Hasil estimasi standart kuadrat terkecil

Jika asumsi terpenuhi, teori kuadrat terkecil memberikan hasil yang baik, yaitu :

1. Sifat-sifat $\hat{\beta}$ vektor berukuran (kx1) adalah :

a. $E(\hat{\beta}) = \beta$ (2.4.1a)

bukti :

$$E(\hat{\beta}) = E[(X^T X)^{-1} X^T Y]$$
$$= E[(X^T X)^{-1} X^T (X\beta + \varepsilon)]$$
$$= E[(X^T X)^{-1} X^T X\beta + (X^T X)^{-1} X^T \varepsilon)]$$
$$= \beta$$

dari $E(\varepsilon)=0$ dan $(X^T X)^{-1} X^T X = I$

maka $\hat{\beta}$ adalah estimator tidak bias dari β

b. $\hat{\beta}$ adalah estimator tidak bias terbaik (BLUE), dari β, artinya diantara kelas estimator linier tidak bias, $\hat{\beta}$ mempunyai variansi terkecil.

Variansi dari $\hat{\beta}$ adalah $Var(\hat{\beta}) = \sigma^2 (X^T X)^{-1}$ (2.4.1b)

bukti:

$$Var[\hat{\beta}] = Var[(X^T X)^{-1} X^T Y]$$
$$= (X^T X)^{-1} X^T \; Var[Y] \; X(X^T X)^{-1}$$
$$= \sigma^2 \; ((X^T X)^{-1} X^T X (X^T X)^{-1}$$
$$= \sigma^2 \; (X^T X)^{-1}$$

c. $\hat{\beta} \sim N_k(\beta, \sigma^2 (X^T X)^{-1})$ (2.4.1c)

$N_k(\mu, \Sigma)$ adalah distribusi normal multivariat dengan rata-rata μ suatu vektor berukuran (kx1), dan varian Σ suatu matrik berukuran (kxk).

2. Nilai vektor prediksi $\hat{Y}$, berukuran (nx1) adalah :

$$\hat{Y} = X\hat{\beta} = (X^T X)^{-1} X^T Y = PY, \quad \text{dimana } P=(X^T X)^{-1} X^T$$

mempunyai sifat-sifat :

a. $E(\hat{Y})=X\beta$ (2.4.2a)

b. $Var(\hat{Y})=\sigma^2 P$ (2.4.2b)

c. $\hat{Y} \sim N_n(X\beta,\sigma^2 P)$ (2.4.2c)

bukti :

a. dari $\hat{Y}=X(X^T X)^{-1} X^T Y$

$$\begin{aligned} E[\hat{Y}] &= E[X(X^T X)^{-1} X^T Y] \\ &= X(X^T X)^{-1} X^T E[Y] \\ &= X(X^T X)^{-1} X^T X\beta \\ &= X\beta \end{aligned}$$

b. $$\begin{aligned} Var[\hat{Y}] &= Var[X(X^T X)^{-1} X^T Y] \\ &= X(X^T X)^{-1} X^T Var[Y] X(X^T X)^{-1} X^T \\ &= \sigma^2 X(X^T X)^{-1} X^T X(X^T X)^{-1} X^T \\ &= \sigma^2 X(X^T X)^{-1} X^T = \sigma^2 P \end{aligned}$$

c. Sehingga $\hat{Y} \sim N_n(X\beta,\sigma^2 P)$

3. Vektor residual biasa berukuran nx1

$$e = Y - \hat{Y} = Y - PY = (I-P)Y$$

mempunyai sifat-sifat:

a. $E(e) = 0$ (2.4.3a)

b. $Var(e) = \sigma^2 (I-P)$ (2.4.3b)

c. $e \sim N_n(0,\sigma^2(I-P))$ (2.4.3c) bukti:

a. dari $e = Y - \hat{Y} = Y - PY = Y - X(X^T X)^{-1} X^T Y$

maka $E(e) = E(Y) - E(X(X^T X)^{-1} X^T Y)$

$= X\beta - X(X^T X)^{-1} X^T X\beta$

$= X\beta - X\beta = 0$

b. $Var[e] = Var[(I-P)Y]$

$= (I-P)\ Var[Y]\ (I-P)$

$= \sigma^2\ (I-P)$

c. sehingga $e \sim N_n(0, \sigma^2(I-P))$

4. Estimator tidak bias dari σ^2 adalah

$$\hat{\sigma}^2 = \frac{e^T e}{n - k} = \frac{Y^T (I-P) Y}{n - k} \qquad (2.4.4)$$

dimana $e^T e$ adalah jumlah kuadrat residual.