BAB II

TEORI PENUNJANG

2.1 Aljabar Matriks

Matriks merupakan himpunan skalar yang disusun secara empat persegi panjang menurut baris-baris dan kolom-kolom. Untuk membahas matriks perlu kiranya mengingat definisi-definisi berikut ini :

<u>Definisi 2.1</u>

Suatu matriks $A = [a_{ij}]$ berukuran mxn. Transpose dari A bernotasi A^T adalah matriks berukuran nxm yang didapatkan dari matriks A dengan menuliskan baris ke-i dari matriks A sebagai kolom ke-i dari A^T, atau dapat ditulis $A^T = [a_{ji}]$.

Beberapa sifat matriks transpose :

1. $(A + B)^T = A^T + B^T$
2. $(A^T)^T = A$
3. $\lambda\ (A^T) = (\lambda A)^T$
4. $(AB)^T = B^T A^T$

<u>Definisi 2.2</u>

Matriks A dikatakan simetris apabila transpose matriks A sama dengan matriks A atau matriks A simetris bila $A = A^T$.

<u>Definisi 2.3</u>

Sebuah matriks bujur sangkar berukuran nxn dikatakan mempunyai invers bila ada suatu matriks B, sehingga $AB = BA = I_n$. Matriks B disebut invers dari matriks

A, ditulis A^{-1}, merupakan matriks bujur sangkar berukuran nxn.

Theorema 2.1

Jika A matriks non singular maka $A^{-1} = \frac{1}{\det A}$ adj A

Bukti :

Ambil A adalah matriks berukuran nxn atau berordo n,

Ambil A adj A = (b_{ij}), dengan i = 1,2,...,n dan j = 1,2,...,n. Adjoint dari matriks A merupakan transpose dari matriks kofaktor A, yaitu :

$$\text{Adj } (A) = K^T = \begin{bmatrix} K_{11} & K_{21} & \cdots & K_{n1} \\ K_{12} & K_{22} & \cdots & K_{n2} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ K_{1n} & K_{2n} & \cdots & K_{nn} \end{bmatrix}$$

Sedangkan kofaktor dituliskan dengan $K_{ij} = (-1)^{i+j} |M_{ij}|$ dimana $|M_{ij}|$ merupakan determinan minor matriks yang diperoleh dari matriks A berukuran nxn dihilangkan baris ke-i dan kolom ke-j atau merupakan determinan dari matriks yang berukuran (n-1)x(n-1).

Determinan matriks A = $|A| = a_{i1}K_{i1} + a_{i2}K_{i2} + \ldots + a_{in}K_{in}$

$$= \sum_{t=1}^{n} a_{it} K_{it}$$

$[b_{ij}]$ = A adj A

$$= A \begin{bmatrix} K_{11} & K_{21} & \cdots & K_{n1} \\ . & . & \cdots & . \\ . & . & \cdots & \\ K_{1n} & K_{2n} & \cdots & K_{nn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ . & . & \cdots & . \\ . & . & \cdots & . \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} K_{11} & K_{21} & \cdots & K_{n1} \\ K_{12} & K_{22} & \cdots & K_{n2} \\ . & . & \cdots & . \\ . & . & \cdots & . \\ K_{1n} & K_{2n} & \cdots & K_{nn} \end{bmatrix}$$

$$= \begin{bmatrix} |A| & 0 & \cdots & 0 \\ 0 & |A| & \cdots & 0 \\ . & . & \cdots & . \\ . & . & \cdots & . \\ 0 & 0 & \cdots & |A| \end{bmatrix}$$

$$= |A| \; I$$

jadi A adj A = (det A) I

$$A^{-1} = \frac{1}{\det A} \text{ adj } A$$

contoh

Carilah A^{-1} dari matriks $A = \begin{bmatrix} 4 & 3 & 2 \\ 2 & 1 & 1 \\ 1 & 2 & 2 \end{bmatrix}$

Penyelesaian :

$$M_{11} = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix},\ M_{12} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix},\ M_{13} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

$$K_{11} = (-1)^{1+1}\det\ (M_{11}) = (-1)^{2}\ (\ 2 - 2\ ) = 0$$

$$K_{12} = (-1)^{1+2}\det\ (M_{12}) = (-1)^{3}\ (\ 4 - 1\ ) = -3$$

$$K_{13} = (-1)^{1+3}\det\ (M_{13}) = (-1)^{4}\ (\ 4 - 1\ ) = 3$$

$$M_{21} = \begin{bmatrix} 3 & 2 \\ 2 & 2 \end{bmatrix},\ M_{22} = \begin{bmatrix} 4 & 2 \\ 1 & 2 \end{bmatrix},\ M_{23} = \begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix}$$

$$K_{21} = (-1)^{2+1}\det\ (M_{21}) = (-1)^{3}\ (\ 6 - 4\ ) = -2$$

$$K_{22} = (-1)^{2+2}\det\ (M_{22}) = (-1)^{4}\ (\ 8 - 2\ ) = 6$$

$$K_{23} = (-1)^{2+3}\det\ (M_{23}) = (-1)^{5}\ (\ 8 - 3\ ) = -5$$

$$M_{31} = \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix},\ M_{32} = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix},\ M_{33} = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}$$

$$K_{31} = (-1)^{3+1}\det\ (M_{31}) = (-1)^{4}\ (\ 3 - 2\ ) = 1$$

$$K_{32} = (-1)^{3+2}\det\ (M_{32}) = (-1)^{5}\ (\ 4 - 4\ ) = 0$$

$$K_{33} = (-1)^{3+3}\det\ (M_{33}) = (-1)^{6}\ (\ 4 - 6\ ) = -2$$

$$\begin{aligned} \det\ (A) &= a_{11}\ K_{11} + a_{12}\ K_{12} + a_{13}\ K_{13} \\ &= 4\ (0) + 3\ (-3) + 2\ (3) \\ &= -9 + 6 \\ &= -3 \end{aligned}$$

$$\text{Adj (A)} = \begin{bmatrix} K_{11} & K_{21} & K_{31} \\ K_{12} & K_{22} & K_{32} \\ K_{13} & K_{23} & K_{33} \end{bmatrix} = \begin{bmatrix} 0 & -2 & 1 \\ -3 & 6 & 0 \\ 3 & -5 & -2 \end{bmatrix}$$

jadi :

$$A^{-1} = \frac{1}{\det A} \text{ adj } A = \frac{1}{-3} \begin{bmatrix} 0 & -2 & 1 \\ -3 & 6 & 0 \\ 3 & -5 & -2 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 2/3 & -1/3 \\ 1 & -2 & 0 \\ -1 & 5/3 & 2/3 \end{bmatrix}$$

Definisi 2.4

Bentuk kuadratik dari variabel $x_1, x_2, \ldots, x_n$ dapat ditulis sebagai :

$$X^T A X = [\, x_1, x_2, \ldots, x_n \,] \, A \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ x_n \end{bmatrix}$$

$$= a_{11}\, x_1^2 + a_{22}\, x_2^2 + \ldots + a_{nn}\, x_n^2$$

$$+ \sum_{i \neq j} a_{ij} x_i \, x_j$$

dengan A adalah matriks simetris

Definisi 2.5

Bentuk kuadratik $X^T A X$ disebut definit positif jika $X^T A X > 0$ untuk setiap $X \neq 0$, sedangkan matriks A disebut mat r iks definit positif $X^T A X$ adalah bentuk kuadratik definit positif.

contoh :

$$[\,x_1 \quad x_2\,] \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2\,x_1^2 - 2\,x_1 x_2 + 2\,x_2^2$$

Definisi 2.6

Himpunan m buah vektor $\{ u_1, u_2, \ldots, u_m \}$ dikatakan bebas linier bila terdapat skalar-skalar λ_1, λ_2, ..., λ_m sedemikian sehingga $\lambda_1 u_1 + \lambda_2 u_2 + \ldots + \lambda_m u_m = 0$ hanya terpenuhi oleh $\lambda_1 = \lambda_2 = \ldots = \lambda_m = 0$

contoh :

Apakah [2,3] dan [1,3] bebas linier ?

Penyelesaian :

Diketahui [2,3] dan [1,3], jika terdapat λ_1 , λ_2 maka persamaan dapat dibentuk : λ_1 [2,3] + λ_2 [1,3] = 0

atau $2\,\lambda_1 + \lambda_2 = 0$

$3\,\lambda_1 + 3\,\lambda_2 = 0$

jika dua persamaan tersebut diselesaikan maka akan didapat $\lambda_1 = \lambda_2 = 0$; jadi bebas linier.

Definisi 2.7

A suatu matriks bujur sangkar, λ skalar yang memenuhi persamaan $A\,v = \lambda\,v$, untuk suatu vektor kolom $v \neq 0$, maka dikatakan λ adalah suatu akar karakterisrtik (eigen value) dari A, dan v yang mempunyai persamaan $A\ v = \lambda\ v$ disebut vektor karakteristik (eigen vektor) yang bersangkutan dengan λ .

contoh :

Hitunglah akar karakteristik dan vektor karakteristik dari matriks $A = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$

Penyelesaian :

Misalkan λ skalar dan $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ vektor yang memenuhi :

$\begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \lambda \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$, merupakan persamaan karakteristik :

$$\begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 1-\lambda & 2 \\ 3 & 2-\lambda \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \text{...........(2.1.1)}$$

sehingga determinan dari $\begin{bmatrix} 1-\lambda & 2 \\ 3 & 2-\lambda \end{bmatrix}$ adalah :

$(1 - \lambda)\ (2 - \lambda) - 6 = 0$

$\lambda^2 - 3\lambda - 4 = 0$

$\lambda_1 = 4 \ , \ \lambda_2 = -1$

Vektor karakteristik dicari dengan memasukkan harga λ kedalam persamaan (2.1.1) yaitu :

Untuk $\lambda_1 = 4$

$$\begin{bmatrix} 1-4 & 2 \\ 3 & 2-4 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

atau $\quad -3\,v_1 + 2\,v_2 = 0$

$\quad 3\,v_1 - 2\,v_2 = 0$

ambil persamaan $-3v_1 + 2v_2 = 0$

jika $v_1 = 2\mu$ maka $v_2 = 3\mu$, jadi $v = \mu \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ merupakan vektor-vektor karakteristik yang bersangkutan untuk λ_1.

Untuk $\lambda_2 = -1$

$$\begin{bmatrix} 1+1 & 2 \\ 3 & 2+1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

atau $2v_1 + 2v_2 = 0$

$3v_1 + 3v_2 = 0$

ambil persamaan $2v_1 + 2v_2 = 0$

jika $v_1 = \mu$ maka $v_2 = -\mu$, jadi $v = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ merupakan vektor-vektor karakteristik yang bersangkutan untuk λ_1.

2.2 Matriks Diagonal dan Diagonalisasi Matriks

2.2.1 Matriks Diagonal

Matriks diagonal adalah matriks yang elemen-elemen diluar diagonal utama bernilai nol. Apabila matriks D bersifat diagonal, maka dapat ditulis sebagai $D = \mathrm{diag}\,(d_{11}, d_{22}, \ldots, d_{nn})$ di mana d untuk $i = 1,2,\ldots,n$ merupakan nilai elemen diagonal utama ke-i.

2.2.2 Diagonalisai Matriks Bujur Sangkar

Suatu cara untuk mengubah matriks bujur sangkar menjadi matriks diagonal melalui hubungan ekuivalensi dan kongruensi.

Definisi 2.8

Matriks bujur sangkar A dikatakan dapat

didiagonalisasi jika terdapat matriks P yang non singular sehingga $P^{-1}A\ P$ merupakan matriks diagonal maka matriks P disebut mendiagonalisasi matriks A.

Theorema 2.2

Jika A adalah matriks bujur sangkar berukuran nxn, maka pernyataan-pernyataan berikut ekivalen satu sama lain :

(i) Matriks A dapat didiagonalisasi

(ii) Matriks A mempunyai n vektor eigen bebas linier

bukti :

(i) ⇒ (ii)

Karena matriks A diketahui dapat didiagonalisasi, maka terdapat matriks non singular P, yaitu :

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix}$$

menurut definisi 2.8 diperoleh $P^{-1}A\ P$ merupakan matriks diagonal, misalkan $P^{-1}A\ P = D$ dengan :

$$D = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

jelas D merupakan matriks diagonal.

Karena P matriks non singular maka matriks P mempunyai invers yaitu $P^{-1}P = P\ P^{-1} = I$ sehingga dapat diperoleh persamaan :

$$P^{-1}A\,P = D$$
$$P\,P^{-1}A\,P = P\,D$$
$$A\,P = P\,D$$

dengan

$$P\,D = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1\,p_{11} & \lambda_2\,p_{12} & \cdots & \lambda_1\,p_{1n} \\ \lambda_1\,p_{21} & \lambda_2\,p_{22} & \cdots & \lambda_2\,p_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \lambda_1\,p_{n1} & \lambda_2\,p_{n2} & \cdots & \lambda_n\,p_{nn} \end{bmatrix} \quad \ldots(2.2.2)$$

Jika dimisalkan $\tilde{P}_1$,$\tilde{P}_2$,. . . . ., $\tilde{P}_n$ menyatakan vektor-vektor kolom dari matriks P, yaitu :

$$\tilde{P}_1 = \begin{bmatrix} p_{11} \\ p_{12} \\ \cdot \\ \cdot \\ p_{n1} \end{bmatrix},\ \tilde{P}_2 = \begin{bmatrix} p_{12} \\ p_{22} \\ \cdot \\ \cdot \\ p_{n2} \end{bmatrix},\ \ldots\ldots,\ \tilde{P}_n = \begin{bmatrix} p_{1n} \\ p_{2n} \\ \cdot \\ \cdot \\ p_{nn} \end{bmatrix}$$

maka dari persamaan (2.2.2) kolom-kolom P D yang berurutan adalah $\lambda_1\,\tilde{P}_1$, $\lambda_2\,\tilde{P}_2$, . . . , $\lambda_n\,\tilde{P}_n$ yang juga merupakan kolom-kolom A P yang berurutan. Akan tetapi untuk matriks A P atau matriks perkalian antara matriks A dengan matriks P akan diperoleh kolom-kolom yang berurutan adalah A $\tilde{P}_1$,

$A\ \tilde{P}_2$, ... , $A\ \tilde{P}_n$. Sehingga akan diperoleh :

$$A\ \tilde{P}_1 = \lambda_1 \tilde{P}_1\ ,\ A\ \tilde{P}_2 = \lambda_2 \tilde{P}_2\ ,\ ....\ ,\ A\ \tilde{P}_n = \lambda_n \tilde{P}_n \quad ...(2.2.3)$$

karena P matriks non singular mengakibatkan vektor-vektor kolomnya tidak nol.

Sehingga dalam persamaan (2.2.3) λ merupakan akar karakteristik (eigenvalue) dari A dan $\tilde{P}_1$,$\tilde{P}_2$, . . . ,$\tilde{P}_n$ adalah vektor karakteristik (eigenvektor) yang bersesuaian. Dari matriks P yang non singular dapat diperoleh $\tilde{P}_1$, $\tilde{P}_2$, ... , $\tilde{P}_n$ bebas linier. Jadi matriks A mempumyai n vektor karakteristik yang bebas linier.

(ii) $\Rightarrow$ (i)

Diketahui bahwa matriks A mempunyai n vektor eigen yang bebas linier yaitu $\tilde{P}_1$, $\tilde{P}_2$, ... , $\tilde{P}_n$ dengan nilai eigen yang bersesuaian adalah λ_1, λ_2, ... , λ_n , sedangkan $\tilde{P}_1$, $\tilde{P}_2$, ... , $\tilde{P}_n$ merupakan vektor-vektor kolom dari matriks P sehingga matriks P non singular. Dan juga diketahui $A\ \tilde{P}_1 = \lambda_1 \tilde{P}_1$, $A\ \tilde{P}_2 = \lambda_2 \tilde{P}_2$, ... , $A\ \tilde{P}_n = \lambda_n \tilde{P}_n$, terlihat A P = P D . Dari A P = P D diperoleh hubungan-hubungan :

$$A\ P = P\ D$$

$$P^{-1} A\ P = P^{-1} P\ D$$

$$P^{-1} A\ P = D$$

dengan D adalah matriks diagonal yang mempunyai nilai karakteristik λ_1, λ_2, ... , λ_n pada diagonal utama, sehingga $P^{-1}A\ P$ juga merupakan matriks diagonal. Dari definisi (2.8) maka P mendiagonalisasi matriks A atau matriks A dapat didiagonalisai.

2.3 Diferensiasi Terhadap Vektor dan Matriks

Misalkan f(x) adalah suatu fungsi kontinyu dari elemen-elemen vektor $x^T = [x_1, x_2, \ldots, x_p]$ dimana derivatif parsial pertama dan keduanya $\dfrac{\partial f(x)}{\partial x_i}$, $\dfrac{\partial^2 f(x)}{\partial x_i \partial x_j}$

didefinisikan vektor operator derivatif parsial sebagai :

$$\frac{\partial}{\partial x} = \begin{bmatrix} \dfrac{\partial}{\partial x_1} \\ \cdot \\ \cdot \\ \cdot \\ \dfrac{\partial}{\partial x_p} \end{bmatrix} \quad \ldots\ldots\ldots\ldots(2.3.1)$$

Penerapan operator ini untuk hasil fungsi f(x) berupa vektor-vektor parsial, yaitu :

$$\frac{\partial f(x)}{\partial x} = \begin{bmatrix} \dfrac{\partial f(x)}{\partial x_1} \\ \cdot \\ \cdot \\ \cdot \\ \dfrac{\partial f(x)}{\partial x_p} \end{bmatrix} \quad \ldots\ldots\ldots\ldots(2.3.2)$$

Beberapa contoh fungsi dan derivatifnya :

1. Jika f(x) adalah konstanta untuk semua x, maka

$$\frac{\partial f(x)}{\partial x} = \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} \quad \ldots\ldots\ldots\ldots\ldots\ldots(2.3.3)$$

2. Matriks derivatif parsial kedua dari fungsi p variabel disebut matriks Hess :

$$H = \frac{\partial^2 f(x)}{\partial x_i \; \partial x_j} = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_i^2} & \cdots\cdots\cdots & \frac{\partial^2 f(x)}{\partial x_i \partial x_p} \\ . \quad \frac{\partial^2 f(x)}{\partial x_{i+1}^2} & \cdots & . \\ . \quad\quad . & & . \; . \; . \\ \frac{\partial^2 f(x)}{\partial x_i \, \partial x_p} \quad . & \cdots\cdots & \frac{\partial^2 f(x)}{\partial^2 x_p} \end{bmatrix}$$

Matriks Hess lebih diutamakan simetris jika kondisi daerahnya kontinyu dan terdapat derivatif pertama dan kedua yang dipenuhi oleh f(x).

2.4 Model Regresi dan Estimasi Parameter

Seringkali dijumpai dalam suatu penelitian untuk menyelesaikan permasalahan digunakan analisis regresi, beberapa permasalahan regresi dapat mencakup lebih dari satu variabel bebas. Model-model regresi yang menggunakan lebih dari satu variabel bebas disebut model regresi berganda. Model regresi berganda merupakan salah satu teknik statistik yang telah digunakan secara luas diberbagai permasalahan.

2.4.1 Model Regresi Berganda

Model umum regresi linier dengan sesatan Normal adalah :

$$y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_{p-1} x_{i,p-1} + \varepsilon_i \quad \text{......(2.4.1)}$$

dengan :

$\varepsilon_i \sim \text{NID}(0, \sigma^2)$

$\beta_0, \beta_1, \dots, \beta_{p-1}$ parameter-parameter

sehingga variabel tak bebas atau respon y yang dihubungkan dengan k variabel bebas berbentuk :

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + \varepsilon \quad \text{.........(2.4.2)}$$

disebut model regresi linier berganda dengan k variabel bebas.

2.4.2 Estimasi Parameter

Metode kuadrat terkecil dapat digunakan untuk memperkirakan koefisien regresi pada persamaan (2.4.2). Misalkan terdapat n obsevasi dengan n > k dan misalkan x_{ij} menyatakan observasi ke-i pada tingkat variabel ke-j sehingga jika kita asumsikan bahwa error ε dalam model mempunyai $E(\varepsilon) = 0$, $V(\varepsilon) = \sigma^2$ dan $\{\varepsilon_i\}$ adalah variabel-variabel random yang tak berkorelasi. Maka persamaan dapat berbentuk :

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i \quad \text{...(2.4.3)}$$

$$= \beta_0 + \sum_{j=1}^{k} \beta_j x_{ij} + \varepsilon_i$$

Jika persamaan (2.4.3) disajikan dalam bentuk matriks maka akan berbentuk :

$$\tilde{Y} = \tilde{X}\,\tilde{\beta} + \tilde{\varepsilon}$$

dengan :

$$\tilde{Y} = \begin{bmatrix} y_1 \\ y2 \\ . \\ . \\ y_n \end{bmatrix} \quad \tilde{X} = \begin{bmatrix} 1 & x_{11} & x_{12} \dots\dots & x_{1k} \\ 1 & x_{21} & x_{22} \dots\dots & x_{2k} \\ . & . & . \dots\dots & . \\ . & . & . \dots\dots & . \\ 1 & x_{n1} & x_{n2} \dots\dots & x_{nk} \end{bmatrix}, \; \tilde{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ . \\ . \\ \beta_k \end{bmatrix}, \tilde{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ \varepsilon_n \end{bmatrix}$$

$\tilde{Y}$ adalah vektor berukuran nx1 dari observasi-observasi

$\tilde{X}$ adalah matriks berukuran nx(k+1)

$\tilde{\beta}$ adalah vektor berukuran (k+1)x1 dari parameter

$\tilde{\varepsilon}$ adalah vektor error random berukuran nx1

Untuk mendapatkan estimator vektor kuadrat terkecil $\hat{\beta}$ yang minimum digunakan jumlah kuadrat galatnya,yaitu :

$$\begin{aligned} S &= \sum_{i=1}^{n} \varepsilon = \varepsilon^T \varepsilon \\ &= (Y - X\beta)^T (Y - X\beta) \\ &= Y^TY - \beta^TX^TY - Y^TX\beta + \beta^TX^TX\beta \\ &= Y^TY - 2\beta^TX^TY + \beta^TX^TX\beta \quad \dots\dots\dots\dots\dots(2.4.4) \end{aligned}$$

karena $\beta^TX\,Y$ adalah sebuah matriks skalar (1x1) dan $(\beta^TX^TY)^T = Y^TX\beta$ adalah skalar yang sama.

Penaksir-penaksir kuadrat terkecil harus memenuhi :

$$\left.\frac{\partial S}{\partial \beta}\right|_{\hat{\beta}} = -2X^TY + 2X^TX\hat{\beta} = 0 \quad \dots\dots.(2.4.5)$$

Untuk mendapatkan penaksir yang benar-benar minimum haruslah :

$$\left.\frac{\partial^2 S}{\partial \beta^2}\right|_{\hat{\beta}} > 0$$

dari persamaan (2.4.5) dapat diperoleh :

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

dengan $X \neq 0$ maka $2\ X^T X$ definit positif atau dengan perkataan lain $2\ X^T X > 0$ sehingga taksiran β minimum.

Penyederhanaan dari persamaan (2.4.5) adalah :

$$X^T\ X\ \hat{\beta} = X^T\ Y \qquad \text{.............(2.4.6)}$$

persamaan (2.4.6) merupakan persamaan normal kuadrat terkecil. Untuk menyelesaikan persamaan normal ini adalah dengan cara mengalikan kedua ruas persamaan (2.4.6) dengan $(X^T\ X)^{-1}$, maka diperoleh penaksir kuadrat terkecil untuk β adalah :

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

asalkan $(\ X^T\ X\ )$ non singular.

Penulisan kembali persamaan (2.4.6) dalam persamaan matriks adalah sebagai berikut :

$$\begin{bmatrix} n & \Sigma x_{i1} & \Sigma x_{i2} & \cdots & \Sigma x_{ik} \\ \Sigma x_{i1} & \Sigma x_{i1}^2 & . & \cdots & \Sigma x_{i1}x_{ik} \\ . & . & . & \cdots & . \\ . & . & . & \cdots & . \\ \Sigma x_{ik} & \Sigma x_{i1}x_{ik} & & \cdots & \Sigma x_{ik}^2 \end{bmatrix} \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ . \\ . \\ \hat{\beta}_k \end{bmatrix} = \begin{bmatrix} \Sigma y_i \\ \Sigma x_{i1} y_i \\ . \\ . \\ \Sigma x_{ik} y_i \end{bmatrix}$$

sehingga perkiraan model regresinya adalah :

$$\hat{Y} = X \hat{\beta}$$

dan bentuk skalarnya adalah

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \dots + \hat{\beta}_k x_{ik} .$$

2.5 Fungsi Kepadatan Peluang

Definisi 2.9

Untuk peubah acak $(X_1, X_2, \dots, X_k)$ yang kontinyu, fungsi $F_{X_1, X_2, \dots, X_k}(x_1, x_2, \dots, x_k) =$

$$\int_{-\infty}^{x_k} \dots \int_{-\infty}^{x_1} f_{X_1, X_2, \dots, X_k}(x_1, x_2, \dots, x_k) \, dx_1 \dots dx_k$$

dengan $f_{X_1, X_2, \dots, X_k}(x_1, x_2, \dots, x_k) \geq 0$

$f_{X_1, X_2, \dots, X_k}(x_1, x_2, \dots, x_k)$ disebut fungsi kepadatan peluang sedangkan $F_{X_1, X_2, \ldots, X_k}(x_1, x_2, \ldots, x_k)$

disebut fungsi sebaran kumulatif.

Untuk peubah acak yang kontinyu, fungsi sebaran kumulatif adalah $F_X(x) = \int_{-\infty}^{x} f_X(x)\, dx$.

Sehingga fungsi kepadatan peluang mempunyai dua ketentuan yaitu :

(i) $f_{X_1, X_2, \ldots, X_k}(x_1, x_2, \ldots, x_k) \geq 0$

(ii) $\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} f_{X_1, \ldots, X_k}(x_1, x_2, \ldots, x_k)\, dx_1 \ldots dx_k = 1$

Fungsi kepadatan peluang tersebut biasa digunakan untuk menentukan nilai peluang. Misalnya untuk peubah acak X yang kontinyu dengan fungsi kepadatan peluang $f_X(x)$ dan $a < X < b$ maka nilai peluangnya adalah $P[\, a < X < b\,] = \int_a^b f_X(x)\, dx$.

Secara umum untuk $X \in R$ dengan fungsi kepadatan peluang $f_X(x)$ maka nilai nilai peluangnya akan diperoleh $P[\, X \in R\,] = \int \ldots \int_R f_X(x)\, dx$.

contoh :

Terdapat fungsi $f(x,y) = K(x + y)$, $0 < x < 1$, $0 < y < 1$ jika K merupakan konstanta yang positif, maka $f(x,y) \geq 0$

$$\int_0^1 \int_0^1 K\, f(x,y)\, dx\, dy = \int_0^1 \int_0^1 K(x + y)\, dx\, dy$$

$$= K \int_0^1 \int_0^1 (x + y)\, dx\, dy$$

$$= K \int_0^1 (1/2 + y)\, dy$$

$$= K (1/2 + 1/2)$$

$$= K$$

Untuk K = 1 fungsi $f (x,y) = x + y$ merupakan fungsi kepadatan peluang yaitu :

$$\int_0^1 \dots \int_0^1 K\, f (x,y)\, dx\, dy = K$$

$$= 1$$

2.6 Distribusi Keluarga Eksponensial

Definisi 2.10

Suatu Variabel random Y berdistribusi Normal, jika fungsi kepadatan peluangnya (p.d.f) berbentuk :

$$f(y) = \frac{1}{\sqrt{(2 \pi \sigma^2)}} \exp [- (y - \mu)^2 / 2 \sigma^2]$$

$$- \infty < y < \infty$$

Parameter rata-rata μ dan parameter variansi σ^2 harus memenuhi kondisi-kondisi $-\infty < \mu < \infty$ dan $\sigma > 0$.

Definisi 2.11

Suatu distribusi merupakan anggota keluarga eksponensial k parameter bila fungsi kepadatan peluangnya tersebut dapat dinyatakan sebagai :

$$f(x|\theta) = h(x).c(\theta) \exp [\sum_{i=1}^{n} a_i(\theta)\ t_i(x)]$$

dengan :

$h(x) \geq 0$ merupakan fungsi non negatif dari x

$t_i(x)$ adalah fungsi berharga riil dari x tidak tergantung pada θ

$c(\theta) \geq 0$ merupakan fungsi non negatif dari θ

$a_i(\theta)$ merupakan fungsi berharga riil dari θ tidak tergantung pada x

contoh :

$f(x ; \theta) = \theta\ e^{-\theta x}$ termasuk keluarga eksponensial

yaitu : $f (x ; \theta) = \theta\ e^{-\theta x}$

$= \theta \exp [- \theta x]$

dengan $c(\theta) = \theta$, $h(x) = 1$, $a_i(\theta) = - \theta$, $t_i(x) = x$.

Sedangkan untuk model linier generalisasi dengan peubah acak y_i bentuk keluarga eksponensialnya dapat ditulis sebagai :

$$f(y_i | \theta_i, \emptyset, w_i) = \text{eks} \{ \frac{y_i\theta_i - b(\theta_i)}{\emptyset} w_i + c(y_i, \emptyset, w_i) \}$$

dimana :

θ_i adalah parameter kanonik atau parameter asli yang nilainya bergantung pada observasi (i)

$\emptyset$ adalah parameter penyebaran (dispersi) yang nilainya tidak bergantung pada observasi (i), sehingga nilai dari parameter dapat berupa nilai yang konstan

$b(\theta_i)$ merupakan fungsi berharga riil dari parameter θ_i

$c (y_i, \emptyset, w_i)$ adalah fungsi berharga riil dari y_i yang tidak bergantung pada parameter θ akan tetapi tergantung pada $\emptyset$

w_i adalah bobot dengan $w_i = 1$ untuk data tak berkelompok (i = 1,2, ...,n).

contoh :

Untuk distribusi Normal pada definisi 2.10 dengan menyatakan :

$$\frac{1}{\sqrt{(2\pi\sigma^2)}} = \exp \ln (2\pi\sigma^2)^{-1/2}$$

maka fungsi kepadatan peluang dapat dinyatakan :

$$f(y) = \exp \ln (2\pi\sigma^2)^{-1/2} \exp [-(y-\mu)^2 / 2\sigma^2]$$

$$= e^{-1/2 \ln (2\pi\sigma^2)} \cdot e^{-\frac{1}{2\sigma^2}(y^2 - 2\mu y + \mu^2)}$$

$$= e^{-1/2 \ln (2\pi\sigma^2) - \frac{1}{2\sigma^2} y^2 + \frac{y\mu}{\sigma^2} - \frac{1}{2}\mu^2}$$

$$= \exp [\frac{y\mu - \frac{\mu^2}{2}}{\sigma^2} - \frac{1}{2} [\frac{y^2}{\sigma^2} + \ln (2\pi\sigma^2)]]$$

Bentuk ini merupakan bentuk fungsi kepadatan peluang distribusi keluarga eksponensial.

2.7 Metode Maksimum Likelihood

Definisi 2.12

Misalkan (Y_1 , Y_2 , ..., Y_n) merupakan vektor random dengan fungsi fungsi kepadatan peluang (p. d. f) $f_{\tilde{\theta}}(y_1, y_2, ..., y_n)$ ($\tilde{\theta} \in \Omega$; $\tilde{\theta}$:

vektor parameter, Ω : ruang parameter).

Fungsi $L(\tilde{\theta}\ ; y_1, y_2, \ldots, y_n) = f(\tilde{\theta}\ ; y_1, y_2, \ldots, y_n)$ dipandang sebagai fungsi dari $\tilde{\theta}$ disebut fungsi likelihood.

Jika $Y_1, Y_2, \ldots, Y_n$ berdistribusi saling bebas (iid) dengan fungsi kepadatan peluang $f_{\tilde{\theta}}(y_i)$, maka fungsi likelihoodnya adalah :

$$L(\tilde{\theta}\ ; y_1, y_2, \ldots, y_n) = \prod_{i=1}^{n} f_{\tilde{\theta}}(y_i).$$

Prinsip metode maksimum likelihood adalah mencari suatu fungsi dari $y_1, y_2, \ldots, y_n$ sedemikian sehingga, jika $\tilde{\theta}$ diganti dengan $\hat{\tilde{\theta}}(y_1, y_2, \ldots y_n)$ fungsi likelihood L adalah suatu maksimum yaitu :

$$L[\hat{\tilde{\theta}}(y_1, y_2, \ldots, y_n)\ ; y_1, y_2, \ldots, y_n] \geq L[\tilde{\theta}\ ; y_1, y_2, \ldots, y_n].$$

untuk setiap $\tilde{\theta} \in \Omega$.

Sedangkan taksiran maksimum likelihood dari $\tilde{\theta}$ yaitu $\hat{\tilde{\theta}}$ dapat diperoleh dari fungsi likelihood dengan membuat turunan dari fungsi likelihood $L(\tilde{\theta})$ sama dengan nol , yaitu :

$$\frac{\partial L(\tilde{\theta})}{\partial \tilde{\theta}} = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2.7.1)$$

Dalam persamaan (2.7.1) terlihat bahwa dalam fungsi likelihood tersebut hanya terdapat sebuah fungsi parameter θ yang tidak diketahui, estimator maksimum likelihood θ adalah nilai θ yang memaksimumkan fungsi likelihood $L(\theta)$.

Metode maksimum likelihood dapat digunakan untuk memperkirakan dalam keadaan dimana beberapa parameter tidak diketahui, misalnya parameter yang tidak diketahui adalah θ_1, θ_2, ... , θ_k. Dalam keadaan seperti ini fungsi likelihood adalah sebuah fungsi k parameter tidak diketahui dan estimator maksimum likelihood $\hat{\theta}_1$, $\hat{\theta}_2$, ... , $\hat{\theta}_k$ akan diperoleh dengan menyamakan turunan parsial pertama sama dengan nol, yaitu :

$$\frac{\partial L(\theta_1, \theta_2, \ldots, \theta_k)}{\partial \theta_i} = 0$$

dan harga-harga $\hat{\theta}_1$, $\hat{\theta}_2$, ... , $\hat{\theta}_k$ merupakan calon estimator maksimum likelihood yang mungkin, untuk memperoleh estimator $\hat{\theta}_1$, $\hat{\theta}_2$, ... , $\hat{\theta}_k$ yang benar-benar mamaksimumkan fungsi likelihood $L(\tilde{\theta} \mid \tilde{y})$ harus ditunjukan bahwa :

$$\left. \frac{\partial^2 L(\tilde{\theta} \mid \tilde{y})}{\partial \theta_i^2} \right|_{\theta_i = \hat{\theta}_i} < 0$$

Karena fungsi logaritma naik tegas pada (0 ,∞) ini berarti bahwa $L(\tilde{\theta} \mid \tilde{y})$ dan $\ln L(\tilde{\theta} \mid \tilde{y})$ akan mempunyai ektsrim yang sama, sehingga akan lebih mudah digunakan deferensiasi dari logaritma alamnya (ln) yaitu :

$$\frac{\partial \ln L(\tilde{\theta} \mid \tilde{y})}{\partial \tilde{\theta}_i} = 0$$

yang merupakan persamaan maksimum likelihood.