BAB II

MODEL REGRESI GANDA DENGAN MATRIKS

Dalam analisa regresi dibedakan dua jenis variabel yaitu variabel bebas atau variabel prediktor dan variabel tak bebas atau variabel respon. Penentuan variabel bebas dan yang tak bebas dalam beberapa hal tidak mudah dapat dilaksanakan. Studi yang cermat, diskusi yang seksama, berbagai pertimbangan, kewajaran masalah yang dihadapi dan pengalaman akan membantu memudahkan penentuan. Variabel yang mudah didapat atau tersedia sering dapat digolongkan kedalam variabel bebas, sedangkan variabel yang terjadi karena variabel bebas itu merupakan variabel tak bebas. Untuk keperluan analisis, variabel bebas dinyatakan dengan X_1, X_2 ..., X_k ($k \geq 1$) sedangkan variabel tak bebas dinyatakan dengan Y.

Misalnya, untuk fenomena yang meliputi hasil panen padi dan volume pupuk yang digunakan, diambil variabel tak bebas atau respon Y = hasil panen. Untuk tiga variabel yang meliputi pertumbuhan bakteri, macam zat perantara tempat bakteri hidup dan waktu, dapat diambil respon Y = pertumbuhan bakteri, prediktor X_1 = macam zat perantara dan prediktor X_2 = waktu. Tetapi untuk dua variabel tentang berat dan tinggi badan, salah satu bisa dipilih sebagai variabel bebas.

Karena banyaknya data pengamatan yang terjadi lebih dari dua variabel, maka secara umum data hasil pengamatan Y terjadi sebagai akibat dari variabel-variabel bebas X_1, X_2, ..., X_{p-1} yang biasa disebut persamaan regresi linier ganda. Sehingga model umum regresi linier ganda dengan sesatan normal adalah

$$Y_i = \beta_0 + \beta_1 X_{i1} + \dots + \beta_{p-1} X_{i,p-1} + \epsilon_i.$$

dengan :

β_0, β_1, ... β_{p-1} parameter-parameter.

$\epsilon_i \sim NID(0, \sigma^2)$

2.1 Model Umum Regresi Linier Ganda Dengan Matriks

Model umum regresi linier ganda dengan matriks yaitu

$$\underset{n\times 1}{\underset{\sim}{Y}} = \underset{n\times p}{\underset{\sim}{X}} \; \underset{p\times}{\underset{\sim}{\beta}} + \underset{n\times 1}{\underset{\sim}{\epsilon}}$$

Untuk menyatakan model ini dengan menggunakan matriks didefinisikan matriks-matriks sebagai berikut :

$$\underset{nx1}{\underset{\sim}{Y}} = \begin{bmatrix} Y_1 \\ Y_2 \\ . \\ . \\ . \\ Y_n \end{bmatrix} \qquad \underset{nxp}{\underset{\sim}{X}} = \begin{bmatrix} 1 & X_{11} & X_{12} & X_{1,p-1} \\ 1 & X_{21} & X_{22} & X_{2,p-1} \\ 1 & X_{31} & X_{32} & X_{3,p-1} \\ . & & & \\ . & & & \\ . & & & \\ 1 & X_{n1} & X_{n2} & X_{n,p-1} \end{bmatrix}$$

$$\underset{px1}{\underset{\sim}{\beta}} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ . \\ . \\ . \\ \beta_{p-1} \end{bmatrix} \qquad \underset{nx1}{\underset{\sim}{\epsilon}} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ . \\ . \\ . \\ \epsilon_n \end{bmatrix}$$

dengan :

$\underset{\sim}{Y}$ vektor observasi

$\underset{\sim}{\beta}$ vektor parameter

$\underset{\sim}{\epsilon} \sim NID\ (0, \sigma^2 \underset{\sim}{I})$

Dengan anggapan normalitas ini, didapat

$\underset{\sim}{Y} \sim NID\ (\ E(\underset{\sim}{Y}),\ \sigma^2 \underset{\sim}{I}\ )$

dengan :

$E(\underset{\sim}{Y}) = \underset{\sim}{X}\ \underset{\sim}{\beta}$

$var(\underset{\sim}{Y}) = \sigma^2 \underset{\sim}{I}$, matriks variansi-covariansi dari $\underset{\sim}{Y}$.

2.2 Estimator-estimator Kuadrat terkecil

Estimator-estimator dari koefisien regresi $b_0, b_1, \ldots, b_{p-1}$, ditulis sebagai vektor yaitu :

$$\underset{px1}{\underset{\sim}{b}} = \begin{bmatrix} b_0 \\ b_1 \\ \cdot \\ \cdot \\ \cdot \\ b_{p-1} \end{bmatrix}$$

Persamaan normal untuk model ini,

$$(\underset{\sim}{X}'\underset{\sim}{X})\underset{\sim}{b} = \underset{\sim}{X}'\underset{\sim}{Y}$$

dan estimator-estimator kuadrat terkecil didapat dari penyelesaian persamaan normal

$$\underset{\sim}{b} = (\underset{\sim}{X}'\underset{\sim}{X})^{-1} (\underset{\sim}{X}'\underset{\sim}{Y})$$

Estimator-estimator ini, tak bias dengan variansi minimum.

2.3 Analisis Variansi

Setelah didapat estimator $\underset{\sim}{b}$, selanjutnya dapat dihitung harga-harga

$$\hat{Y}_i = b_0 + b_1X_{i1} + \ldots + b_{p-1}\, X_{ip-1},\ i = 1, 2, \ldots, n.$$

Dengan matriks harga-harga terhitung ini ditulis sebagai :

$$\underset{nx1}{\underset{\sim}{\hat{Y}}} = \underset{nxp}{\underset{\sim}{X}} \quad \underset{px1}{\underset{\sim}{b}}$$

$$\underset{n\times 1}{\hat{\underset{\sim}{Y}}} = \begin{bmatrix} \hat{Y}_1 \\ \hat{Y}_2 \\ \cdot \\ \cdot \\ \cdot \\ \hat{Y}_n \end{bmatrix} \qquad \underset{n\times p}{\underset{\sim}{X}} = \begin{bmatrix} 1 & X_{11} & X_{12} & X_{1,p-1} \\ 1 & X_{21} & X_{22} & X_{2,p-1} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 1 & X_{n1} & X_{n2} & X_{n,p-1} \end{bmatrix}$$

$$\underset{p\times 1}{\underset{\sim}{b}} = \begin{bmatrix} b_0 \\ b_1 \\ \cdot \\ \cdot \\ \cdot \\ b_p \end{bmatrix}$$

suku sisa $e_i = Y_i - \hat{Y}_i$, $i = 1, 2, \ldots, n$

ditulis dengan matriks

$$\underset{n\times 1}{\underset{\sim}{e}} = \underset{n\times 1}{\underset{\sim}{Y}} - \underset{n\times 1}{\hat{\underset{\sim}{Y}}}$$

Seperti pada model regresi linier sederhana harga-harga jumlah kuadrat dihitung dengan rumus-rumus sebagai berikut :

$$JKT = \underset{\sim}{Y}'\underset{\sim}{Y} - n\,\bar{Y}^2$$

$$JKR = b'\underset{\sim}{X}'\underset{\sim}{Y} - n\,\bar{Y}^2$$

$$JKS = \underset{\sim}{e}'\underset{\sim}{e} = \underset{\sim}{Y}'\underset{\sim}{Y} - b'\underset{\sim}{X}'\underset{\sim}{Y}$$

Tabel anava model regresi linier umum

Sumber variasi	d.b	Jumlah kuadrat (JK)	Rata-rata kuadrat (RK)
Regresi	p-1	$JKR = \underset{\sim}{b}'\underset{\sim}{X}'\underset{\sim}{Y} - n\bar{Y}^2$	$RKR = JKR/p-1$
Sesatan	n-p	$JKS = \underset{\sim}{Y}'\underset{\sim}{Y} - \underset{\sim}{b}'\underset{\sim}{X}'\underset{\sim}{Y}$	$RKS = JKS/n-p$
Total	n-1	$JK = \underset{\sim}{Y}'\underset{\sim}{Y} - n\bar{Y}^2$	

Estimator kuadrat terkecil $\underset{\sim}{b}$ merupakan estimator tak bias $E(\underset{\sim}{b}) = \underset{\sim}{\beta}$. Matriks variansi-covariansi dari b :

$$Var(\underset{\sim}{b}) = \begin{bmatrix} Var\ (b_0) & Cov\ (b_0, b_1) & \ldots & Cov\ (b_0, b_{p-1}) \\ \vdots & \vdots & & \vdots \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & & Var\ (b_{p-1}) \end{bmatrix}$$

$$Var\ (\underset{\sim}{b}) = \sigma^2 (\underset{\sim}{X}'\underset{\sim}{X})^{-1}$$

Estimator dari Var (b) :

$$s^2(\underset{\sim}{b}) = \begin{bmatrix} s^2(b_0) & s(b_0, b_1) & \ldots & s(b_0, b_{p-1}) \\ & s^2(b_1) & \ldots & s(b_1, b_{p-1}) \\ & & & \vdots \\ & & & s^2(b_{p-1}) \end{bmatrix}$$

$$s^2(\underset{\sim}{b}) = RKS\ (\underset{\sim}{X}'\underset{\sim}{X})^{-1}$$

2.4 Contoh :

Data berikut merupakan hasil dari 13 ulangan dengan tiga variabel independen (X_1,X_2,X_3) dan variabel respons Y. Akan dicari estimator dari persamaan regresi ganda untuk data observasi sebagai berikut :

Tabel data observasi

Y	X_1	X_2	X_3
25,5	1,74	5,30	10,8
31,2	6,32	5,42	9,4
25,9	6,22	8,41	7,2
38,4	10,52	4,63	8,5
18,4	1,19	11,60	9,4
26,7	1,22	5,85	9,9
26,4	4,10	6,62	8,0
25,9	6,32	8,72	9,1
32,0	4,08	4,42	8,7
25,2	4,15	7,60	9,2
39,7	10,15	4,83	9,4
35,7	1,72	3,12	7,6
26,5	1,70	5,30	8,2

Dari data dihitung harga-harga :

$n = 13$ $\Sigma Y_i = 377,5$ $\Sigma Y^2{}_i = 11.400,15$

$\Sigma X_{i1} = 59,43$ $\Sigma X_{i2} = 81,82$ $\Sigma X_{i3} = 115,40$

$\Sigma X^2{}_{i1} = 394,7255$ $\Sigma X^2{}_{i2} = 576,7264$ $\Sigma X^2{}_{i3} = 1035,96$

$\Sigma X_{i1}Y_i = 1877,567$ $\Sigma X_{i2}Y_i = 2246,661$

$\Sigma X_{i3}Y_i = 3337,78$

$\Sigma X_{i1}X_{i2} = 360,6621$ $\Sigma X_{i1}X_{i3} = 522,078$ $\Sigma X_{i2}X_{i3} = 728,31$

persamaan normal :

$$\underset{\sim}{X}'\underset{\sim}{X}\underset{\sim}{b} = \underset{\sim}{X}'\underset{\sim}{Y}$$

$$\begin{bmatrix} n & \Sigma X_{i1} & \Sigma X_{i2} & \Sigma X_{i3} \\ \Sigma X_{i1} & \Sigma X^2_{i1} & \Sigma X_{i1}X_{i2} & \Sigma X_{i1}X_{i3} \\ \Sigma X_{i2} & \Sigma X_{i1}X_{i2} & \Sigma X^2_{i2} & \Sigma X_{i2}X_{i3} \\ \Sigma X_{i3} & \Sigma X_{i1}X_{i3} & \Sigma X_{i2}X_{i3} & \Sigma X^2_{i3} \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} \Sigma Y_i \\ \Sigma X_{i1}Y_i \\ \Sigma X_{i2}Y_i \\ \Sigma X_{i3}Y_i \end{bmatrix}$$

$$\begin{bmatrix} 13 & 59,43 & 81,82 & 115,40 \\ 59,43 & 394,7255 & 360,6621 & 522,078 \\ 81,82 & 360,6621 & 576,7264 & 728,31 \\ 115,40 & 522,0780 & 728,3100 & 1035,96 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 377,5 \\ 1877,567 \\ 2246,661 \\ 3337,78 \end{bmatrix}$$

$(\underset{\sim}{X}'\underset{\sim}{X})^{-1}$ dihitung dengan menggunakan komputer atau dengan rumus :

$$\underset{\substack{\sim \\ nxn}}{A} = [a_{ij}], \quad \underset{\sim}{A}^{-1} = \frac{\text{adj } A}{|\ A\ |}$$

dengan :

$$\text{adj } \underset{\sim}{A} = \begin{bmatrix} \alpha_{11} & \alpha_{21} & \cdots & \alpha_{n1} \\ \alpha_{12} & \alpha_{22} & & \alpha_{n2} \\ \vdots & \vdots & & \vdots \\ \alpha_{1n} & \alpha_{2n} & & \alpha_{nn} \end{bmatrix}$$

α_{ij} = kofaktor dari a_{ij}

$|A|$ = determinan dari A

penyelesaian dari persamaan normal, adalah

$$\underset{\sim}{b} = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\ \underset{\sim}{X}'\underset{\sim}{Y}$$

$$\begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 8,0648 & -0,0862 & -0,0942 & -0,7905 \\ -0,0862 & 0,0085 & 0,0017 & 0,0027 \\ -0,0942 & 0,0017 & 0,0166 & -0,0021 \\ -0,7905 & 0,0037 & -0,0021 & 0,0886 \end{bmatrix} \begin{bmatrix} 377,5 \\ 1877,567 \\ 2246,661 \\ 3337,780 \end{bmatrix}$$

$$= \begin{bmatrix} 39,1574 \\ 1,0161 \\ -1,8616 \\ -0,3433 \end{bmatrix}$$

selanjutnya dihitung harga jumlah kuadrat

$$JK = \underset{\sim}{Y}'\underset{\sim}{Y} = (\Sigma Y^2{}_i) \Rightarrow JKT = \underset{\sim}{Y}'\underset{\sim}{Y} - n\bar{Y}^2 = \underset{\sim}{Y}'\underset{\sim}{Y} - \frac{(\Sigma Y)^2}{n}$$

$$= 11.400,15 - 13\ \frac{377,5}{13} = 438,13$$

$$JKR = \underset{\sim}{b}'\underset{\sim}{X}'\underset{\sim}{Y} - n\bar{Y}^2$$

$$= (b_0\ b_1\ b_2\ b_3) \begin{bmatrix} \Sigma Y_i \\ \Sigma Y_{i1} Y_i \\ \Sigma X_{i2} Y_i \\ \Sigma X_{i3} Y_i \end{bmatrix} - n\bar{Y}^2$$

$$=(39,1574\ \ 1,0161\ \ -1,8616\ \ -0,343) \begin{bmatrix} 377,5 \\ 1877,567 \\ -2246,661 \\ 3337,789 \end{bmatrix} -(13)\frac{377,5}{13}$$

$$= 399,45$$

JKS = JKT - JKR

= 438,13 - 399,45

= 138,68

$$RKS = \frac{JKS}{n-p} = \frac{36,68}{13-4} = 4,298$$

Matriks Variansi-Covariansi untuk $\tilde{b}$ adalah

$$s^2(\tilde{b}) = RKS(\tilde{X}'\tilde{X})^{-1}$$

$$= 4,290 \begin{bmatrix} 8,0648 & -0,0826 & -0,0942 & -0,7905 \\ & 0,0085 & 0,0017 & 0,0027 \\ & & 0,0166 & -0,0021 \\ & & & 0,0886 \end{bmatrix}$$

$s^2(b_1) = 0,036533$

$s^2(b_2) = 0,07135$

$s^2(b_3) = 0,3808.$

Dari penghitungan didapat estimator dari persamaan regresi ganda yaitu b_0=39,1574 b_1=1,0161 b_2=-1,8616 b_3=-0,3433. Sehingga persamaan regresi gandanya yaitu :

$Y = 39,1574 + 1,0161\ X_1 - 1,8616\ X_2 - 0,3433\ X_3$

Juga diperoleh matriks variansi-covariansi dari b serta jumlah kuadrat regresi, jumlah kuadrat sesatan dan jumlah kuadrat total.