BAB II

MATERI PENUNJANG

Dalam regresi, analisis data digunakan untuk menerangkan hubungan antara variabel bebas yang berpengaruh terhadap variabel terikat yang ditaksir. Nilai variabel terikat ini ditentukan oleh parameter-parameter pada variabel bebas. Analisis regresi memberikan suatu cara untuk menaksir parameter tersebut, yang akan membentuk suatu fungsi pendekatan terhadap fungsi dari data hasil pengamatan.

Pada penulisan tugas akhir ini akan dibahas suatu metode untuk mendapatkan parameter pada fungsi regresi, yang merupakan kombinasi dari metode kuadrat terkecil dan regresi MINMAD. Sebelum melangkah pada materi inti, perlu adanya uraian tentang unsur pendukung dari metode kombinasi tersebut.

2.1. Matriks

Matriks A_{mxn} dapat ditinjau dari vektor-vektor baris atau vektor-vektor kolom. Sub ruang dari R^n yang direntang oleh vektor-vektor baris dinamakan ruang baris dari A dan sub ruang dari R^m yang direntang oleh vektor-vektor kolom dinamakan ruang kolom dari A.

Definisi 2.1.1

Matriks berukuran mxn adalah kumpulan bilangan yang disusun secara khusus berbentuk empat persegi panjang dengan m baris dan n kolom.

$$M = \begin{pmatrix} m_{11} & m_{12} & \dots m_{1n} \\ m_{21} & m_{22} & \dots m_{2n} \\ \vdots & \vdots & \vdots \\ m_{m1} & m_{m2} & \dots m_{mn} \end{pmatrix} \text{ atau } (M_{ij})_{mxn}$$

Definisi 2.1.2

Jika $A = (a_{ij})_{mxn}$ dan $B = (b_{ij})_{mxn}$, maka penjumlahan matriks A+B adalah suatu matriks $C = (c_{ij})_{mxn}$ demgan $c_{ij}=a_{ij}+b_{ij}$ untuk i=1,2,...,n dan j=1,2,...,n

Definisi 2.1.3

Jika $A = (a_{ij})_{mxp}$ dan $B = (b_{ij})_{pxn}$, maka perkalian A.B adalah $C = (c_{ij})_{mxn}$ dengan $c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+ \ldots +a_{in}b_{nj}=\sum_{k=1}^{p} a_{ik} b_{kj}$ untuk setiap i=1,2,...,n dan j=1,2,...,n.

Definisi 2.1.4

Transpose dari matriks $A = (a_{ij})_{mxn}$ adalah pengubahan dari baris menjadi kolom dan kolom menjadi baris, dinotasikan dengan $A^T = (a_{ji})_{nxm}$

Definisi 2.1.5

Suatu matriks disebut matriks diagonal, jika $a_{ij}=0$, $i\neq j$ dan a_{ii}= konstan atau matriks yang semua elemennya berharga nol, kecuali elemen diagonalnya. Misal :

$$D=\begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}$$

Definisi 2.1.6

Matriks diagonal yang elemen diagonalnya 1 disebut Matriks Identitas dan dinotasikan dengan I_n. Misal : $I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

Definisi 2.1.7

Suatu bentuk kuadrat dalam variabel $x_1, x_2, \ldots, x_n$

$$f(x_1, x_2, \ldots, x_n) = a_{11}x_1^2 + \ldots + a_{nn}x_n^2 + 2a_{12}x_1x_2 + \ldots + 2a_{1n}x_1x_n + 2a_{(n-1)n}x_{n-1}x_n$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j = \mathbf{X^TAX}$$

dengan $x = (x_1, x_2, \ldots, x_n)^T$ dan A_{nxn} adalah matriks bentuk simetris kuadrat dalam **X** dan **A** disebut matriks bentuk kuadrat.

Definisi 2.1.8

Jika **A** suatu matriks bujursangkar, maka :

(i) Matriks **A** disebut matriks definit positif, jika $\mathbf{X^TAX > 0}, \forall x \neq 0$

Contoh :

Misal $A = \begin{pmatrix} 3 & 0 \\ 5 & 0 \end{pmatrix}$, akan ditunjukkan **A** definit positif

$$\mathbf{F = X^TAX}$$

$$= (x_1 \quad x_2)\begin{pmatrix} 3 & 0 \\ 0 & 5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= 3x_1^2 + 5x_2^2$$

adalah definit positif, karena $\mathbf{X^TAX}$ selalu berharga positif untuk $\forall x \neq 0$

(ii) Matriks **A** disebut matriks semidefinit positif, jika $\mathbf{X^TAX \geq 0}, \forall x \neq 0$

Contoh :

Misal $A = \begin{pmatrix} 4 & -2 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$, akan ditunjukkan **A** semidefinit positif

$$F = X^T AX$$

$$= (x_1 \quad x_2 \quad x_3)\begin{pmatrix} 4 & -2 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$= \quad 4x_1^2 + x_2^2 - 4x_1x_2 + 3x_3^2$$

$$= \quad (2x_1 - x_2)^2 + 3x_3^2$$

adalah semi definit positif, karena selalu positif , kecuali jika $x_2 = 2x_1$ maka

$X^T AX = 0$

(iii) Matriks A disebut matriks definit negatif, jika $X^T AX < 0, \forall x \neq 0$

Contoh:

Misal $A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$, akan ditunjukkan A definit negatif

$$F = X^T AX$$

$$= (x_1 \quad x_2)\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = -x_1^2 - x_2^2$$

adalah definit negatif, karena selalu negatif untuk $\forall x \neq 0$

2.2. METODE KUADRAT TERKECIL

Metode ini dikemukakan oleh Carl Friedrich Gauss, seorang ahli matematika Jerman. Diberikan suatu model regresi linier , sebagai berikut :

$$Y_i = \beta_0 + \beta_1 X_{1i} + \ldots + \beta_p X_{pi} + \varepsilon_i \qquad (2.2.1)$$

dengan ε_i adalah deviasi/sesatan dan $\beta_0, \beta_1, \ldots, \beta_p$ adalah parameter regresi

Pada sejumlah n pengamatan dengan p-1 variabel bebas, n≥p dan x_{ij} merupakan nilai pengamatan ke-j dengan j=1,2,...,n dari variabel bebas ke-i, dengan i=1,2,...,p.

Untuk p=1, maka model (2.2.1) dapat dituliskan dalam bentuk matriks, sebagai berikut :

$$Y = X\beta + \varepsilon \qquad (2.2.2)$$

dengan $Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$, $X = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}$, $\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$ dan $\varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$

Model regresi seperti tersebut di atas, dibangun berdasarkan asumsi-asumsi klasik, sebagai berikut :

1. Nilai rata-rata dari unsur kesalahan pengganggu ε_i adalah nol atau $E(\varepsilon_i)=0$, untuk setiap i, $i = 1,2,\ldots,n$
2. Varian dari ε_i adalah konstan atau homoskedastik

   $$Var(\varepsilon_i) = E(\varepsilon_i - E(\varepsilon_i))^2$$
   $$= E(\varepsilon_i) = \sigma^2$$

3. Tidak ada otokorelasi dalam kesalahan pengganggu ε_i

   $$Cov(E(\varepsilon_i,\varepsilon_j)) = E(\varepsilon_i - E(\varepsilon_i))(\varepsilon_j - E(\varepsilon_j))$$
   $$= E(\varepsilon_i \varepsilon_j)$$
   $$= 0, \text{ untuk } i \neq j$$

4. Variabel bebas X adalah non stokastik (tetap dalam penyampelan berulang), atau jika stokastik, didistribusikan secara independen dari kesalahan pengganggu ε_i
5. Tidak ada multikolineritas di antara variabel bebas X

Pada metode kuadrat terkecil, dilakukan penaksiran parameter regresi dengan ketentuan mutlak didapat sesatan yang sekecil mungkin atau dengan kata lain mempunyai Jumlahan Sesatan Kuadrat yang minimal, sehingga dirumuskan :

$$S = \sum (F_{obs} - F_{approx})$$

$$= (\varepsilon_1{}^2 + \varepsilon_2{}^2 + \ldots + \varepsilon_n{}^2)$$

$$= \sum_{i=1}^{n} \varepsilon_i{}^2$$

Definisi 2.2.1

Jika titik $A(x_0, y_0, z_0)$ merupakan titik ekstrem dari suatu fungsi $Z=f(x,y)$, dengan semua turunan parsial keduanya ada dan kontinyu, maka

(i) Relatif Maksimum, jika

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0, \frac{\partial^2 f}{\partial x^2} < 0 \text{ dan} \left(\frac{\partial^2 f}{\partial x^2}\right)\left(\frac{\partial^2 f}{\partial y^2}\right) > \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2$$

(ii) Relatif Minimum, jika

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0, \frac{\partial^2 f}{\partial x^2} > 0 \text{ dan} \left(\frac{\partial^2 f}{\partial x^2}\right)\left(\frac{\partial^2 f}{\partial y^2}\right) > \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2$$

(iii) Tidak relatif Maksimum atau Minimum

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0 \text{ dan} \left(\frac{\partial^2 f}{\partial x^2}\right)\left(\frac{\partial^2 f}{\partial y^2}\right) < \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2$$

Misalkan dari persamaan (2.2.2) diinginkan suatu persamaan garis penduga :

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i \quad \text{dan} \quad \hat{\varepsilon} = y_i - \hat{y}_i$$

dengan harga x_i adalah harga pendekatan x dan ε_i adalah sesatan.

Secara prosedural, kita bentuk formula jumlahan kuadrat dari sesatan.

$$S = (\varepsilon_1{}^2 + \varepsilon_2{}^2 + \ldots + \varepsilon_n{}^2)$$

$$= \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

$$= \sum_{i=1}^{n} \{y_i - (\hat{\beta}_1 x_i + \hat{\beta}_0)\}^2$$

Selanjutnya untuk meminimalkan jumlahan kuadrat sesatan, maka kita harus mendifferensialkan secara parsial untuk a dan b ,serta berikan harga sama dengan nol untuk setiap turunannya. Demikian pula untuk turunan parsial keduanya juga memenuhi kriteria minimum, seperti pada definisi (2.2.1).

Meminimalkan jumlahan sesatan

$$S = \sum_{i=1}^{n} \varepsilon_i^{\ 2}$$

$$= \sum_{i=1}^{n} \left[y_i - \left(\hat{\beta}_1 x_i + \hat{\beta}_0\right)\right]^2$$

$$\frac{\partial S}{\partial \hat{\beta}_1} = -2\sum_{i=1}^{n}\left(y_i - \hat{\beta}_1 x_i - b\right)\left(-x_i\right) = 0$$

$$2\sum_{i=1}^{n} y_i x_i = 2\hat{\beta}_1 \sum_{i=1}^{n} x_i^{\ 2} + 2\hat{\beta}_0 \sum_{i=1}^{n} x_i$$

$$\sum_{i=1}^{n} y_i x_i = \hat{\beta}_1 \sum_{i=1}^{n} x_i^{\ 2} + \hat{\beta}_0 \sum_{i=1}^{N} x_i \qquad (2.2.3)$$

$$\frac{\partial^2 S}{\partial \hat{\beta}_1^{\ 2}} = \frac{\partial}{\partial \hat{\beta}_1}\left(-2\sum_{i=1}^{n} x_i y_i + 2\hat{\beta}_1 \sum_{i=1}^{n} x_i^{\ 2} + 2\hat{\beta}_0 \sum_{i=1}^{n} x_i\right)$$

$$= 2\sum_{i=1}^{n} x_i^{\ 2}\text{, sehingga } \frac{\partial^2 S}{\partial \hat{\beta}_1} > 0$$

$$\frac{\partial S}{\partial \hat{\beta}_0} = 2\sum_{i=1}^{n}\left(y_i - x_i - \hat{\beta}_0\right)\left(-1\right) = 0$$

$$2\sum_{i=1}^{n} y_i = 2\hat{\beta}_1 \sum_{i=1}^{n} x_i + 2\sum_{i=1}^{n} \hat{\beta}_0$$

$$\sum_{i=1}^{n} y_i = \hat{\beta}_1 \sum_{i=1}^{n} x_i + n\hat{\beta}_0 \qquad (2.2.4)$$

$$\frac{\partial^2 S}{\partial \hat{\beta}_0^{\ 2}} = \frac{\partial}{\partial \hat{\beta}_0}\left(2\sum_{i=1}^{n} y_i + 2\hat{\beta}_1 \sum_{i=1}^{n} x_i + 2n\hat{\beta}_0\right)$$

$$= 2n\text{, sehingga } \frac{\partial^2 S}{\partial \hat{\beta}_0^{\ 2}} > 0$$

$$\frac{\partial^2 S}{\partial\hat{\beta}_0\partial\hat{\beta}_1} = \frac{\partial S}{\partial\hat{\beta}_0}\left(-2\sum_{i=1}^{n} y_i x_i + 2\hat{\beta}_1\sum_{i=1}^{n} x_i^2 + 2\hat{\beta}_0\sum_{i=1}^{n} x_i\right)$$
$$= 2\sum_{i=1}^{n} x_i$$

$$\left(\frac{\partial^2 S}{\partial\hat{\beta}_1^2}\right)\left(\frac{\partial^2 S}{\partial\hat{\beta}_0^2}\right) = 2\sum_{i=1}^{n} x_i \cdot 2n = 4n\sum_{i=1}^{n} x_i^2, \text{ sehingga } \left(\frac{\partial^2 S}{\partial\hat{\beta}_1^2}\right)\left(\frac{\partial^2 S}{\partial\hat{\beta}_0^2}\right) > \left(\frac{\partial^2 S}{\partial\hat{\beta}_1\partial\hat{\beta}_0}\right)$$

Karena definisi 2.2.1 (ii) dipenuhi, maka untuk mendapatkan parameter regresi, kita lihat persamaan (2.2.4)

$$\sum_{i=1}^{n} y_i = \hat{\beta}_1\sum_{i=1}^{n} x_i + n\hat{\beta}_0 \geq 0$$
$$n\hat{\beta}_0 = \sum_{i=1}^{n} y_i - \hat{\beta}_1\sum_{i=1}^{n} x_i$$
$$\hat{\beta}_0 = \frac{\sum_{i=1}^{n} y_i}{n} - \hat{\beta}_1\frac{\sum_{i=1}^{n} x_i}{n}$$
$$= \bar{y} - \hat{\beta}_1\bar{x} \qquad (2.2.5)$$

Kemudian persamaan (2.2.5) kita substitusikan pada persamaan (2.2.3) untuk mendapatkan $\hat{\beta}_1$:

$$\sum_{i=1}^{n} y_i x_i = \hat{\beta}_1\sum_{i=1}^{n} x_i^2 + \hat{\beta}_0\sum_{i=1}^{n} x_i$$
$$\sum_{i=1}^{n} y_i x_i - \hat{\beta}_1\sum_{i=1}^{n} x_i^2 + \left(\frac{\sum_{i=1}^{n} y_i}{n} - \hat{\beta}_1\frac{\sum_{i=1}^{n} x_i}{n}\right)\sum_{i=1}^{n} x_i = 0$$
$$\sum_{i=1}^{n} y_i x_i - \hat{\beta}_1\sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} y_i\frac{\sum_{i=1}^{n} x_i}{n} - \hat{\beta}_1\frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n} = 0$$
$$\sum_{i=1}^{n} y_i x_i + \sum_{i=1}^{n} y_i - \frac{\sum_{i=1}^{n} x_i}{n} - \hat{\beta}\left(\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}\right) = 0$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} Y_i X_i - \frac{\left(\sum_{i=1}^{n} Y_i\right)\left(\sum_{i=1}^{n} X_i\right)}{n}}{\sum_{i=1}^{n} X_i^2 - \frac{\left(\sum_{i=1}^{n} X_i\right)^2}{n}} \tag{2.2.6}$$

Maka dapat diperoleh model regresi :

$$\hat{y}_i = \hat{\beta}_1 x_i + \hat{\beta}_0$$

Contoh 2.2 :

Diberikan data untuk penaksiran suatu persamaan regresi, sebagai berikut :

Tabel 2.1.1

X	1	2	3	4	5
Y	4	3	4	5	5

Berdasarkan data diatas dapat kita hitung nilai penduga untuk koefisien regresi:

Tabel 2.1.2

N	X	Y	X^2	XY
1	1	4	1	4
2	2	3	4	6
3	3	4	9	12
4	4	5	16	20
5	5	5	25	25

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} Y_i X_i - \frac{\left(\sum_{i=1}^{n} Y_i\right)\left(\sum_{i=1}^{n} X_i\right)}{n}}{\sum_{i=1}^{n} X_i^2 - \frac{\left(\sum_{i=1}^{n} X_i\right)^2}{n}}$$

$$= \frac{67 - \frac{(15.21)}{5}}{55 - \frac{(15)^2}{5}} = \frac{67-63}{55-45} = 0,4$$

$$\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X}$$

$$= 4,2 - 0,4(3)$$

$$= 3$$

Sehingga persamaan regresi $\hat{Y} = 0,4X + 3$

2.3. Regresi MINMAD (Minimizing Mean Absolute Deviation)

Di samping metode kuadrat terkecil (LS) yang cukup populer dan luas pemakaiannya , dikenal pula suatu metode penaksiran yang disebut metode norm L_1 (L_1 - norm Minimization). Berbeda dengan metode kuadrat terkecil, dalam metode ini penyelesaian masalah regresi linier dilakukan melalui pendekatan pemrograman matematik, yang tidak membutuhkan asumsi kenormalan.

Berdasarkan masalah penaksiran parameter dalam model regresi linier :

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \qquad (2.3.1)$$

dengan Y_i adalah vektor pengamatan berukuran nx1

X_i adalah matriks pengamatan berukuran nxp

β_0 dan β_1 adalah vektor parameter berukuran px1

ε_i adalah deviasi/ sesatan berukuran nx1 dan n>p

Ketelitian persamaan regresi diukur dengan menggunakan sebuah norm dari vektor sesatan $(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)$ dengan $\hat{\beta}_0$ dan $\hat{\beta}_1$ adalah nilai taksiran untuk β_0 dan β_1

Definisi 2.3.1

Norm-L_q didefinisikan untuk vektor v dengan elemen v_i sebagai berikut :

$$\|v\|_q = \left(\sum_i |v_i|^q\right)^{\frac{1}{q}} \tag{3.2}$$

Di bawah kriteria L_q minimal, estimator $\hat{\beta}$ didapat dengan menyelesaikan

masalah : $\min \|Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i\| = \left(\sum_i |Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i|^q\right)^{\frac{1}{q}}$

untuk q = 1, kriteria ini adalah nilai mutlak terkecil yang dikenal dengan MINMAD, yaitu :

$$\min \|Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i\| = \sum_i |Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i| \tag{2.3.3}$$

Regresi MINMAD didefinisikan sebagai regresi yang didapat dengan meminimalkan rata-rata deviasi mutlak (MAD) antara pengamatan Y_i dan nilai penaksir $\hat{Y}_i$ untuk pengamatan ke-i.

$$\text{MAD} = \frac{1}{n}\sum_{i=1}^{n} |\varepsilon_i|$$

atau untuk memudahkan perhitungan, meminimalkan MAD sama dengan meminimalkan jumlah sesatan mutlak (SAD)

$$\text{SAD} = \sum_{i=1}^{n} |\varepsilon_i|.$$

Koefisien regresinya didapatkan dengan metode Simpleks, yang merupakan satu metode yang digunakan untuk memecahkan persoalan pemrograman linier

Berdasarkan persamaan (2.3.1), masalah regresi MINMAD dapat disusun sebagai berikut :

$$\text{Minimalkan} \sum_{i=1}^{n} |\varepsilon_i| \tag{2.3.4}$$

Dengan kendala $\beta_0 + \beta_1 X_i + \varepsilon_i = Y_i$

$\varepsilon, \beta_0, \beta_1$ tak dibatasi tanda

Untuk menyelesaikan persamaan (2.3.4) diperlukan definisi, sebagai berikut :

Definisi 2.3.2

Sebarang $(\beta_0, \beta_1, d_1, d_2)_{nx1}$ yang memenuhi $\beta_0+\beta_1 X+Id_1-Id_2=Y$ disebut solusi untuk masalah (2.3.4) dengan d_1 merupakan vektor nx1 dari deviasi atas gaaris regresi, d_2 merupakan vektor nx1 dari deviasi bawah garis regresi dan I matriks identitas orde n.

Ambil $A_{nx(p+2n)} = (X, I, -I)$ dan $W_{(p+2n)x1} = (\beta_0, \beta_1, d_1, d_2)$.

Sebarang kolom dari A dinyatakan dengan a_j untuk j =1,2,...,p+2n, sehingga untuk sebarang W yang memenuhi AW = Y merupakan solusi untuk (2.3.4).

Misal $C^T_{1x(p+2n)} = (0, e^T, e^T)$ dengan 0 adalah vektor berukuran 1xp dan $e^T = (1,...,1)$ adalah vektor berukuran 1xn ,maka $C^T W$ disebut fungsi tujuan dari masalah (2.3.4).

Definisi 2.3.3

Himpunan n kolom bebas linier dari matriks A disebut basis dari A

Misal $B=(a_{i1},...,a_{in})$ merupakan basis, maka terdapat p+n kolom yang tak terdapat dalam B disebut kolom non basis

Definisi 2.3.4

(i) Sebarang solusi W, jika memenuhi $w_j > 0$, $j = p+1,..., p+2n$, maka solusi di atas disebut solusi fisibel.

(ii) Sebarang solusi fisibel W disebut solusi basis fisibel, jika kolom A yang sesuai dengan komponen tak nol dari W membentuk himpunan bebas linier dari vektor pada R^n

(iii) Solusi optimal adalah suatu titik pada daerah fisibel dengan nilai fungsi tujuan yang paling menguntungkan. Untuk masalah minimasi, solusi optimalnya adalah suatu titik pada daerah fisibel dengan nilai fungsi tujuan terkecil dan pada masalah maksimasi adalah suatu titik pada daerah fisibel dengan nilai fungsi tujuan terbesar.

2.3.1. Algoritma Metode Simpleks untuk Regresi MINMAD

Metode Simpleks merupakan suatu prosedur aljabar yang bersifat iteratif, yang bergerak selangkah demi selangkah, dimulai dari titik ekstrem pada daerah fisibel menuju titik ekstrem yang optimal. Metode ini digunakan untuk mencari koefisien regresi dan yang menjadi fungsi tujuannya adalah jumlah deviasi/sesatan absolut pada harga pengamatan ke-i., i=1,2,...,n

Misalkan B merupakan basis yang sesuai dengan solusi basis fisibel W

W_B menunjukkan vektor variabel W_j yang sesuai dengan kolom basis B dan koefisien yang sesuai dengan C_j dinyatakan dengan C_B, sedang $Z_j = \sum_{i=1}^{n} C_{B_i} \alpha_{ij}$

dengan B_i adalah kolom basis ke-i. Bentuk tabel yang digunakan untuk menyelesaikan masalah regresi MINMAD adalah sebagai berikut :

Tabel 2.3.1. Tabel awal untuk regresi MINMAD

C_B	Vektor basis	W_B	α_1	$\alpha_2 \dots$	$\alpha_j \dots$	α_{p+2n}
1	a_p atau a_{p+n+1}	$\lvert Y_i \rvert$	α_{11}	$\alpha_{12} \dots$	$\alpha_{1j} \dots$	$\alpha_{1,p+2n}$
:	:	:	:	:	:	:
1	a_{p+n} atau a_{p+2n}	$\lvert Y_n \rvert$	α_{n1}	$\alpha_{n2} \dots$	$\alpha_{nj} \dots$	α_{np+2n}
	$C_k - Z_k$	$Z = \sum \lvert Y_i \rvert$		$C_j - Z_j$		

Langkah-langkah yang dilakukan dalam metode Simpleks, adalah sebagai berikut:

1. Menentukan basis awal dan solusi basis fisibel

   - untuk α_j, $j = 1,\ldots, p$ bukan calon basis untuk tabel awal
   - untuk α_j, $j = p+1,\ldots,p+2n$ merupakan calon basis untuk tabel awal

   Untuk menentukan kolom yang menjadi basis dan solusi basis fisibel dalam tabel awal dilakukan pemilihan sebagai berikut :

   - jika $1 \le j \le p$, maka $W_j = 0$
   - jika $p+1 \le j \le p+n$, maka : $w_j = w_{p+r} = \begin{cases} Y_r, & Y_r > 0 \\ 0, & \text{yang lain} \end{cases}$

   sehingga yang menjadi kolom basis adalah α_j, $p+1 \le j \le p+n$

   - jika $p+n+1 \le j \le p+2n$, maka : $w_j = w_{p+n+r} = \begin{cases} -Y_r, & Y_r < 0 \\ 0, & \text{yang lain} \end{cases}$

   sehingga yang menjadi kolom basis adalah α_j, $p+n+1 \le j \le p+2n$

2. Menentukan C_B awal :

   Mengingat $C^T = (0, e^T, e^T)$, maka :

   - untuk $1 \le j \le p$, $C_j = 0$
   - untuk $p+1 \le j \le p+2n$, $C_j = 1$

   jadi C_B awal, koefisiennya berharga 1

3. Menentukan baris $c_k - Z_k$

   Karena C_B awal berharga satu , maka $Z = \sum |Y_i|$ dan untuk mencari $c_j - Z_j$ digunakan :

   $$c_j - Z_j = c_j - \sum_{i=1}^{n} c_{Bi}\,\alpha_{ij}$$

   untuk kolom basis, $c_j - Z_j = 0$

4. Menentukan kolom kunci atau vektor yang akan masuk basis

a. Misal $C_{j1} - Z_{j1} = \max_{C_k - Z_k > 0} C_k - Z_k$, W_k variabel yang tak dibatasi

b. Misal $|C_{j2} - Z_{j2}| = \max_{C_k - Z_k < 0} |C_k - Z_k|$, pilih j

untuk $|C_j - Z_j| = \max \left[C_{j1} - Z_{j1}, |C_{j2} - Z_{j2}| \right]$

jika j_1 dan j_2 tak ditemukan, maka dilanjutkan ke langkah 8. Jika ditemukan lanjutkan ke langkah 5.

5. Memilih elemen pivot dengan cara :

- Jika C_j-Z_j>0, pilih r sedemikian sehingga $\frac{W_{Br}}{\alpha_{rj}} = \max_{i \in R_B} \left[\frac{W_{Bi}}{\alpha_{ij}}, \alpha_{ij} < 0 \right]$

dengan R_B menunjukkan himpunan indeks variabel terbatas pada basis.

- Jika C_j-Z_j<0, sedemikian sehingga $\frac{W_{Br}}{\alpha_{rj}} = \min_{i \in R_B} \left[\frac{W_{Bi}}{\alpha_{ij}}, \alpha_{ij} > 0 \right]$

6. Lakukan operasi baris elementer dan akan diperoleh tabel baru B, sebagai berikut:

$$\hat{W}_{Br} = \frac{W_{Br}}{\alpha_{rj}}, \hat{\alpha}_{rl} = \frac{\alpha_{rl}}{\alpha_{rj}}$$,dengan α_{rj} adalah elemen pivot

$$\hat{W}_{Bi} = W_{Bi} - \frac{W_{Br}}{\alpha_{rj}} \alpha_{ij}, i \neq r, i = 1,\dots, n$$

$$\hat{\alpha}_{il} = \alpha_{il} - \frac{\alpha_{ij}\alpha_{rl}}{\alpha_{rj}}, i \neq r, i = 1,\dots, n \quad l = 1,\dots, p + 2n$$

$$\hat{Z} = Z + (C_j - Z_j)\frac{W_{Br}}{\alpha_{rj}}$$

$$\hat{C}_l - \hat{Z}_l = (C_l - Z_l) - (C_j - Z_j)\frac{\alpha_{rl}}{\alpha_{rj}}, l = 1,\dots, p + 2n$$

7. Jika masih ada $C_k - Z_k < 0$, maka kembali ke langkah 4. Jika tidak, dilanjutkan langkah 8

8. Proses berhenti, solusi telah optimal dengan $c_j - Z_j \geq 0$ untuk semua variabel non basis dan $c_j - Z_j = 0$, untuk semua variabel tak terbatas w_j.

Hasil akhir dari algoritma Simpleks untuk regresi MINMAD terlihat pada tabel di bawah ini:

Tabel 2.3.2 Tabel Akhir

C_B	Vektor basis	w_B	$\alpha_1 \ldots \alpha_p$	$\alpha_{p+1} \ldots \alpha_{p+2n}$
0	a_1	β_0	1...... 0	$\hat{\alpha}_{1,p+1} \cdots \hat{\alpha}_{1,p+2n}$
:	:	:	: :	: :
0	a_{p-1}	β_{p-1}	0 1	$\hat{\alpha}_{p,p+1} \cdots \hat{\alpha}_{p,p+2n}$
1	a_{p+r} atau a_{p+n+r}	d_{1r} atau d_{2r}	0 0	$\hat{\alpha}_{r,p+1} \cdots \hat{\alpha}_{r,p+2n}$
	$C_k - Z_k$		0 0	

Dari tabel akhir dapat dilihat bahwa untuk vektor basis $a_1, a_2, \ldots, a_p$ dengan C_B berharga nol, diperoleh nilai koefisien regresi $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_p$ dan harga fungsi

tujuan : $\mathbf{Z} = \sum_{r=1}^{n} (d_{1r} + d_{2r})$

2.3.2. Algoritma Barrodale and Roberts

Tabel akhir yang terbentuk dalam metode simpleks biasa, seringkali tidak memberikan hasil yang optimal. Hal ini ditunjukkan oleh diperolehnya harga $\hat{\beta}_0$ dan $\hat{\beta}_1$ (vektor basis $a_1, a_2, \ldots, a_p$ dengan C_B berharga nol), tetapi masih terdapat $C_j - Z_j < 0$. Untuk mengatasi hal ini, digunakan algoritma Barrodale and Roberts.

Langkah-langkah pada algoritma ini adalah sebagai berikut :

1. Memilih r seprti pada langkah 3 algoritma simpleks biasa
2. Hitung $\hat{c}_j - \hat{z}_j = c_j - z_j + 2\alpha_{rj}$
   - jika C_j-Z_j dan $\hat{c}_j - \hat{z}_j$ berlawanan tanda (positif dan negatif), kita masukkan variabel yang berhubungan dengan j dan mengeluarkan B_r dari basis
   - Kalau tidak, ubah semua C_k-Z_k menjadi C_k-Z_k+$2\alpha_{rk}$ dan Z menjadi Z-$2W_B$
3. Mengganti a_{p+r} atau a_{p+n+r} dalam basis pada baris r, dengan a_{p+n+r} atau a_{p+r} dan mengalikan dengan (-1)
4. Membuang B_r dari pertimbangan pemindahan dan temukan vektor yang mungkin untuk dipindahkan seperti pada langkah 3 algoritma simpleks biasa.

Contoh 2.3 :

Akan ditentukan persamaan regresi MINMAD $\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_1$ dari data yang diberikan sebagai berikut :

X	1	2	3	4	5
Y	4	3	4	5	5

Dalam bentuk matriks, dapat dituliskan menjadi : $X = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{bmatrix}$, $Y = \begin{bmatrix} 4 \\ 3 \\ 4 \\ 5 \\ 5 \end{bmatrix}$

Parameter β_0 dan β_1 akan diduga dengan menggunakan algoritma simpleks, sebagai berikut :

Langkah 1. Menentukan basis awal yang dibentuk oleh kolom yang sesuai untuk $d_{11},\ldots,d_{15}$. Solusi basis fisibel ditentukan oleh :

$W_j=0$, untuk $j=1,2$

$$W_{p+r}=\begin{cases} Y_r, \text{ jika } Y_r > 0 \\ 0, \text{ yang lainnya}, 1 \leq r \leq 5 \end{cases}$$

$$W_{p+n+r}=\begin{cases} -Y_r, \text{ jika } Y_r < 0 \\ 0, \text{ yang lainnya}, 1 \leq r \leq 5 \end{cases}$$

$$A=(X,I,-I)=\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 1 & 4 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 \\ 1 & 5 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}$$

Akan disajikan a_{ij} yang sesuai dalam bentuk tabel. Basis **B** yang sesuai untuk solusi ini berisi tepat satu kolom e_r atau $-e_r$ untuk semua $r=1,2,\ldots,5$ dan invers **B** adalah dirinya sendiri, yaitu $B^{-1}=B$. Oleh karena itu, $\alpha_j=B^{-1}a_j$ diperoleh dengan mengalikan baris ke-r dari **A** dengan +1, jika W_{p+r} dalam basis dan −1 jika W_{p+n+r} dalam basis dan $B^{-1}Y=W_B$ memberikan solusi basis fisibel yang sesuai dengan p=1,2 dan n=1,2,...,5. Sedang $C_j-Z_j=C_j-\sum_{i=1}^{n} C_B\alpha_{ij}$ untuk setiap j dan $C_j-Z_j=0$ untuk a_j yang merupakan basis. Untuk tabel awal disajikan pada tabel di bawah ini:

Tabel 2.3.3

C_B	BV	W_B	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}	α_{11}	α_{12}
1	a_3	4	1	1	1	0	0	0	0	-1	0	0	0	0
1	a_4	3	1	2	0	1	0	0	0	0	-1	0	0	0
1	a_5	4	1	3	0	0	1	0	0	0	0	-1	0	0
1	a_6	5	1	4	0	0	0	1	0	0	0	0	-1	0
1	a_7	5	1	5	0	0	0	0	1	0	0	0	0	-1
		21	-5	-15	0	0	0	0	0	2	2	2	2	2

Langkah 2. Pilih a_j yang tak terdapat dalam basis. Untuk mengubah vektor dalam basis dilakukan sebagai berikut :

Karena C_k-Z_k berharga negatif untuk k=1,2 maka :

$$|C_j - Z_j| = \max [|-5|, |-15|] = 15 \text{ ,dengan j=2 dan } C_2 - Z_2 < 0$$

Karena j didapat, dilanjutkan ke langkah 3.

Langkah 3. Memilih r sebagai berikut :

$$\frac{W_{Bi}}{\alpha_{r2}} = \min_{i \in R_B} \left[\frac{W_{Bi}}{\alpha_{i2}} \right]$$

$$= \min_{k \in R_n} \left[\frac{4}{1}, \frac{3}{2}, \frac{4}{3}, \frac{5}{4}, \frac{5}{5} \right] = 1$$

Karena nilai ini terletak pada baris ke-5 maka dipilih r=5.Selanjutnya a_7 dalam basis diganti oleh a_2. Kemudian dilanjutkan langkah 4.

Langkah 4. Membentuk tabel baru yang sesuai dengan basis baru B , sebagai berikut :

$$\hat{W}_{B_5} = \frac{W_{B_5}}{\alpha_{52}} = \frac{5}{5} = 1 \qquad \hat{\alpha}_{5l} = \frac{\alpha_{5l}}{5}, l = 1,2,\ldots,12$$

$$\hat{W}_{B_i} = W_{B_i} - l\alpha_{i2} \quad ,i \neq r, i = 1,2,\ldots,5$$

$$\hat{\alpha}_{il} = \alpha_{il} - \frac{\alpha_{i2}\alpha_{5l}}{5}, \ i \neq r, \ i=1,\ldots,5 \quad l=1,\ldots,12$$

$$\hat{Z} = Z + (C_2 - Z_2)1$$

$$\hat{C}_l - \hat{Z}_l = (C_l - Z_l) - (C_2 - Z_2)\frac{\alpha_{5l}}{5}, \quad l = 1,\ldots 12$$

Akan diperoleh tabel baru yang ditunjukkan pada tabel 2.3.4. Pada tabel dapat dilihat bahwa nilai deviasi mutlak Z=6, yang lebih kecil dari nilai sebelumnya.

Tabel 2.3.4

C_B	BV	W_B	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}	α_{11}	α_{12}
1	a_3	3	4/5	0	1	0	0	0	-1/5	-1	0	0	0	1/5
1	a_4	1	3/5	0	0	1	0	0	-2/5	0	-1	0	0	2/5
1	a_5	1	2/5	0	0	0	1	0	-3/5	0	0	-1	0	3/5
1	a_6	1	1/5	0	0	0	0	1	-4/5	0	0	0	-1	4/5
0	a_2	1	1/5	1	0	0	0	0	1/5	0	0	0	0	-1/5
		6	-2	0	0	0	0	0	3	2	2	2	2	-1

Kembali ke langkah 2

$$|C_j - Z_j| = \max\left[|-2|,|-1|\right] = 2 \text{ ,dengan } j=1 \text{ dan } C_1 - Z_1 < 0$$

Pada langkah 3, pilih r untuk C_1-Z_1<0, sebagai berikut:

$$\frac{W_{Br}}{\alpha_{rl}} = \min_{i\in R_B}\left[\frac{15}{4},\frac{5}{3},\frac{5}{2},5,5\right] = \frac{5}{3}$$

Karena terletak pada baris ke-2, maka dipilih r=2. Selanjutnya a_1 menggantikan a_4 pada basis. Tabel baru seaperti pada langkah 4 ditunjukkan pada tabel 2.3.5. Pada tabel terlihat nilai deviasi mutlak Z=8/3, yang lebih kecil dari nilai sebelumnya.

Tabel 2.3.5

C_B	BV	W_B	α_1	α_2	α_3	α_4	a_5	α_6	α_7	α_8	α_9	α_{10}	α_{11}	α_{12}
1	a_3	5/3	0	0	1	-4/3	0	0	1/3	-1	4/3	0	0	-1/3
0	a_1	5/3	1	0	0	5/3	0	0	-2/3	0	-5/3	0	0	2/3
1	a_5	1/3	0	0	0	-2/3	1	0	-1/3	0	2/3	-1	0	1/3
1	a_6	2/3	0	0	0	-1/3	0	1	1/3	0	1/3	0	-1	2/3
0	a_2	1	0	1	0	-1/3	0	0	5/3	0	1/3	0	0	-1/3
		8/3	0	0	0	10/3	0	0	5/3	2	-4/3	2	2	1/3

Pada tabel terlihat nilai deviasi mutlak Z=8/3, yang nilainya lebih kecil daripada nilai Z sebelumnya dan nilai β_0 dan β_1 ditunjukkan oleh W_1 dan W_2, tetapi solusi tidak optimal, karena C_j-Z_j<0, yaitu C_9-Z_9= -4/3<0, sehingga kita kembali ke tabel 2.3.3. Agar lebih efektif tabel kita potong untuk kolom-kolom yang harganya berpengaruh pada solusi optimal, sedang kolom yang tidak tercantum nilainya tidak mempengaruhi syarat optimal (C_j-Z_j≥0)

Tabel 2.3.6

C_B	BV	W_B	α_1	α_2
1	a_3	4	1	1
1	a_4	3	1	2
1	a_5	4	1	3
1	a_6	5	1	4
1	a_7	5	1	5
	C_k-Z_k	Z=21	-5	-15

Catatan bahwa a_7 digantikan oleh a_2 , pada algoritma simpleks biasa (2.3.1).

Dengan mengubah pada langkah 3, kita hitung $\hat{c}_2 - \hat{z}_2$ = -15+(2x5)=-5, dengan α_{52}=5

Kita pertimbangkan bertambahnya β_1 di luar 1($\frac{W_{B_r}}{\alpha_{52}}$ =5/5), sehingga a_7 digantikan a_{12}.

Kemudian mengalikan baris 5 dengan -1,menambah $2\alpha_{52}$ untuk semua C_k-Z_k dan mengubah Z-$2W_{Br}$. .Proses ini menghasilkan baris terakhir :

C_k-Z_k	Z=11	-3	-5

Selanjutnya, mengeluarkan baris 5, kita pertimbangkan suatu r untuk pemindahan. Seperti pada langkah 3 algoritma simpleks biasa, kita dapat r=4 dan $\frac{W_{B_r}}{\alpha_{42}} = \frac{5}{4}$,

sehingga $\hat{c}_2 - \hat{z}_2$ = -5+(2x4)=3. Karena kita dapatkan C_2-Z_2 bernilai positif, sehingga kita dapat memasukkan a_2 dan menggantikan a_6 dari basis. Diperoleh tabel baru :

Tabel 2.3.7

C_B	BV	W_B	α_1	α_2	α_4
1	a_3	11/4	3/4	0	-1/4
1	a_4	1/2	1/2	0	-1/2
1	a_5	1/4	1/4	0	-3/4
0	a_2	5/4	1/4	1	1/4
1	a_{12}	5/4	1/4	0	1/4
	C_k-Z_k	Z=19/4	-7/4	0	9/4

Selanjutnya, kembali ke langkah 3 algoritma simpleks biasa.Karena C_k-Z_k berharga negatif untuk k=1, kita pilih C_1-Z_1= -7/4<0, didapat $\frac{W_{B_1}}{\alpha_{21}} = 1$ yang terletak pada baris ke-2 sehingga r=2. Untuk iterasi selanjutnya kita masukkan a_1 menggantikan a_4 ke dalam basis dan diperoleh tabel akhir :

Tabel 2.3.8

C_B	BV	W_B	α_1	α_4
1	a_3	2	0	-3/2
0	a_1	1	1	2
1	a_5	0	0	-1/2
0	a_2	1	0	-1/2
1	a_{12}	1	0	-1/2
	C_k-Z_k	Z=3	0	7/2

Dari tabel di atas, basis yang diperoleh sudah optimal. W_1=a_1 menunjukkan koefisien β_0 sebesar 1 dan W_2=a_2 menunjukkan koefisien β_1 sebesar 1 . Kemudian W_3 =a_3 menunjukkan deviasi mutlak d_{11} sebesar 2 dan W_{12} menunjukkan deviasi mutlak d_{25} sebesar 2, sehingga jumlah deviasi mutlak = 2+1=3

Sesuai dengan tabel 2.3.8,koefisien β_0 dan β_1, membentuk persamaan regresi MINMAD, yaitu : $\hat{Y} = 1 + X$