BAB II

MATERI PENUNJANG

Dalam mempermudah penurunan rumus - rumus optimalisasi suatu fungsi, maka kita memerlukan beberapa definisi dan beberapa teorema dasar sebagai pendukung pembuktian. Pada bab ini akan dibahas pula, beberapa akibat dari teorema yang diperoleh sebelumnya dan beberapa penjelasan agar lebih mudah dimengerti.

Pokok -pokok yang akan diuraikan di sini adalah tentang vektor dan matriks, norma, himpunan konveks, dan teknik optimasi.

2.1 VEKTOR

Definisi 1

Vektor adalah kumpulan dari bilangan real yang sejenis. Misalkan $\mathbf{X} = (x_1, x_2, x_3, \ldots, x_N)$ adalah vektor dari N elemen atau komponen.

Definisi 2

Misal himpunan V sedemikian sehingga jika $\mathbf{X},\mathbf{Y},\mathbf{Z} \in V$ dan $\alpha, \beta \in R$, maka memenuhi postulat sebagai berikut:

Penambahan

i. $\mathbf{X} + \mathbf{Y} \in V$

ii. $\mathbf{X} + \mathbf{Y} = \mathbf{Y} + \mathbf{X}$

iii. $(X + Y) + Z = X + (Y + Z)$

iv. Terdapat $0 \in V$, $X + 0 = X$

v. Terdapat $-0 \in V$, $X + (-X) = 0$

Perkalian

vi. $\alpha X \in V$

vii. $\alpha(X + Y) = \alpha X + \alpha Y$

ix. $(\alpha \beta) X = \alpha(\beta X)$

x. $IX = XI = X$

Maka V disebut ruang vektor dan elemen-elemennya disebut vektor.

Definisi 3

Jika $X = (x_1, x_2, x_3, \ldots, x_N)$ dan $Y = (y_1, y_2, y_3, \ldots, y_N)$ adalah sembarang vektor pada R^N, maka dot product atau hasil kali dalam Euclidis didefinisikan sebagai: 12

$$X.Y = x_1 y_1 + x_2 y_2 + \ldots + x_N y_N$$

$$= \sum_{i=1}^{N} x_i \; y_i$$

Jika hasil kali dalam Euclidis dari dua vektor sama dengan nol, maka dua vektor tersebut dikatakan ortogonal.

Definisi 4

Sebuah vektor X dinamakan kombinasi linier dari vektor-vektor $X_1, X_2, \ldots, X_N$ jika terdapat bilangan real α_i, tidak semuanya nol, sedemikian sehingga:

$X = \sum_{i=1}^{N} \alpha_i X_i$, diman $\alpha_1, \alpha_2, \ldots, \alpha_N$ adalah skalar.

Definisi 5

Subhimpunan W dari sebuah ruang vektor V dinamakan subruang V jika W itu sendiri adalah ruang vektor di bawah penambahan dan perkalian skalar yang didefinisikan pada V.

Misalkan $W \subseteq R^N$ sedemikian sehingga

$$W = \{X \mid X = (x_1, x_2, \ldots, x_N)\},$$

maka W adalah subhimpunan dari R^N. Jika $\mathbf{X}, \mathbf{Y} \in W$, maka:

$$\mathbf{X} + \mathbf{Y} \in W,$$

$$\mathbf{X} + \mathbf{Y} = (x_1 + y_1, x_2 + y_2, \ldots, x_N + y_N) \in R^N$$

$$\alpha \mathbf{X} = (\alpha x_1, \alpha x_2, \ldots, \alpha x_N) \in W, \ \alpha \in R$$

Teorema 1

Jika W adalah himpunan semua kombinasi linear dari $X_i \in V$, $i = 1, 2, \ldots, M$ maka W adalah subruang V.

Bukti :

Ambil $\mathbf{X} = \sum_{i=1}^{M} \alpha_i X_i$, $Y = \sum_{i=1}^{M} \beta_i X_i$, sedemikian sehingga $\mathbf{X}, \mathbf{Y} \in W$, maka

$$\mathbf{X} + \mathbf{Y} = \sum_{i=1}^{M} (\alpha_i + \beta_i) X_i \in W$$

$$\lambda \mathbf{X} = \sum_{i=1}^{M} (\lambda \alpha_i) X_i \in W, \ \lambda \in R$$

jadi W adalah ruang vektor dan sehingga merupakan subruang dari V. ∎

Definisi 6

Jika $X_1, X_2, \ldots, X_N$ adalah vektor - vektor pada ruang vektor V dan Jika masing-masing vektor pada V dapat dinyatakan sebagai kombinasi linear $X_1, X_2, \ldots, X_N$ maka dikatakan bahwa vektor-vektor ini membangun V.

Definisi 7

Jika V adalah sebarang ruang vektor dan S = { $X_1, X_2, \ldots, X_N$} merupakan himpunan berhingga dari vektor-vektor pada V, maka S dinamakan basis untuk V jika

i. S bebas linear

ii. S membangun V

Definisi 8

Jika S = { $V_1, V_2, V_3, \ldots, V_N$} adalah himpunan vektor maka persamaan vektor :

$$k_1V_1 + k_2V_2 + k_3V_3 + \ldots + k_NV_N = 0 \qquad 13$$

mempunyai paling sedikit satu pemecahan, yakni

$k_1 = 0$, $k_2 = 0$, $k_3 = 0$, ... , $k_N = 0$

jika ini adalah satu-satunya solusi, maka S dikatakan himpunan bebas linear. Jika ada pemecahan lain maka S dikatakan himpunan tak bebas linear.

Kita akan menggunakan notasi matriks

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \text{ atau}$$

$X = [x_1 \; x_2 \; \ldots \; x_N]^T$ untuk menyatakan vektor $X=(x_1, x_2, \ldots, x_N)$ pada R^N. Hal ini dibenarkan karena operasi matriks

$$X + Y = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_N + y_N \end{bmatrix}$$

$$\lambda X = \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_N \end{bmatrix}$$

Satu-satunya perbedaan bahwa hasil-hasilnya dinyatakan secara vertikal dalam satu kasus dan secara herisontal untuk kasus yang lain.

ambil $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}$ dan $Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}$ sebagai matrik Nx1 untuk vektor dan menghilangkan tanda kurung besar pada matriks 1x1, maka selanjutnya

$$X^T Y = [x_1, x_2, \ldots, x_N] \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = [y_1 x_1 + y_2 x_2 + \ldots + y_N x_N]$$

$$= [\, Y \,.\, X \,]$$

$$= Y \,.\, X$$

Jadi, bagi vektor pada notasi vertikal, kita punya rumus

$$X^T Y = Y \,.\, X$$

sehingga hasil kali dalam Euclidis dapat dinyatakan sebagai

$$X^T Y = \sum_{i=1}^{N} x_i y_i \qquad 14$$

2.2 NORMA DARI VEKTOR

Ambil sebarang R^N notasi untuk himpunan dari semua vektor kolom dengan bilangan real sebagai koefisien. Untuk mendefinisikan jarak dalam R^N kita gunakan notasi norma.

Definisi 9

Norma vektor pada R^N adalah suatu fungsi yang dinotasikan dengan $\| \cdot \|$ dari R^N menuju R dengan memenuhi ketentuan :

i. $\| X \| \geq 0$, untuk semua $X \in R^N$

ii. $\| X \| = 0$ jhj $X = (0,0,\ldots,0)^T \equiv 0$

iii. $\| \alpha X \| = | \alpha | \, \| X \|$, untuk setiap $\alpha \in R$ dan $X \in R^N$

iv. $\| X + Y \| \leq \| X \| + \| Y \|$, untuk setiap $X, Y \in R^N$

Definisi 10

Norma l_2 dan l_∞ untuk vektor $X = (x_1, x_2, \ldots, x_N)^T$ didefinisikan dengan

i. $\| X \|_2 = \left\{ \sum_{i=1}^{N} x_i^2 \right\}^{1/2}$

ii. $\| X \|_\infty = \max_{1 \leq i \leq N} | x_i |$

Teorema 2

Untuk setiap $X, Y \in R^N$ dan $\alpha \in R$

i. $\| X \|_\infty \geq 0$,

ii. $\| X \|_\infty = 0$ jhj $X = 0$

iii. $\| \alpha X \|_\infty = | \alpha | \, \| X \|_\infty$

iv. $\| X + Y \|_\infty \leq \| X \|_\infty + \| Y \|_\infty$

sehingga $\| \cdot \|_\infty$ adalah sebuah norma pada R^N

Bukti :

Dibuktikan iv. Jika $X = (x_1, x_2, \ldots, x_N)^T$ dan $Y = (y_1, y_2, \ldots, y_N)$ maka

$$\| X + Y \|_\infty = \max_{1\le i\le N} \{ | x_i | + | y_i | \}$$

$$\le \max_{1\le i\le N} | x_i | + \max_{1\le i\le N} | y_i |$$

$$\le \| X \|_\infty + \| Y \|_\infty$$

untuk 1,2,3 tak perlu dibuktikan. ∎

Teorema 3

Untuk setiap $X, Y \in R^N$, dan $\alpha \in R$

i. $\| X \|_2 \ge 0$

ii. $\| X \|_2 = 0$ jhj $X = 0$

iii. $\| \alpha X \|_2 = | \alpha | \| X \|_2$

iv. $\| X + Y \|_2 \le \| X \|_2 + \| Y \|_2$

sehingga $\| \cdot \|_2$ adalah norma pada R^N

Bukti :

Untuk setiap $X = (x_1, x_2, \ldots, x_N)^T$, dan $\alpha \in R$,

$$\| X \|_2 = \left\{ \sum_{i=1}^{N} x_i^2 \right\}^{1/2} \ge 0$$

dan $\| X \|_2 = 0$ jhj $x_i^2 = 0$, untuk setiap $i=1,2,..,N$

Jadi $\| X \|_2 = 0$ jhj $x = 0$, dan ketentuan i dan ii terbukti. ∎

Untuk membuktikan iii ,

$$\| \alpha X \|_2 = \left\{ \sum_{i=1}^{N} (\alpha x_i)^2 \right\}^{1/2} = \left\{ \sum_{i=1}^{N} \alpha^2 x_i^2 \right\}^{1/2}$$

$$\| \alpha X \|_2 = (\alpha^2)^{1/2} \left\{ \sum_{i=1}^{N} x_i^2 \right\}^{1/2}$$

$$= | \alpha | \| X \|_2$$

untuk membuktikan ketentuan iv, kita butuhkan lemma sebagai berikut:

lemma 1

Untuk setiap $X = (x_1, x_2, \ldots, x_N)^T$ dan $Y = (y_1, y_2, \ldots, y_N)^T$ dalam R^N,

$$\sum_{i=1}^{N} | x_i y_i | \leq \left\{ \sum_{i=1}^{N} x_i^2 \right\}^{1/2} \left\{ \sum_{i=1}^{N} y_i^2 \right\}^{1/2}$$

bukti :

Jika $X = 0$ dan $Y = 0$, didapat kedua ruas sama dengan nol. Misalkan $X \neq 0$ atau $Y \neq 0$, untuk setiap $\lambda \in R$

$$0 \leq \| X - \lambda Y \|_2^2 = \sum_{i=1}^{N} (x_i - \lambda y_i)^2$$

$$= \sum_{i=1}^{N} x_i^2 - 2\lambda \sum_{i=1}^{N} x_i y_i + \lambda^2 \sum_{i=1}^{N} y_i^2$$

dan $2\lambda \sum_{i=1}^{N} x_i y_i \leq \sum_{i=1}^{N} x_i^2 + \lambda^2 \sum_{i=1}^{N} y_i^2 = \| X \|_2^2 + \| Y \|_2^2$

selama bagian i dan ii dari teorema (3) mengakibatkan bahwa $\| X \|_2 > 0$ dan $\| Y \|_2 > 0$, pilih

$$\lambda = \frac{\| X \|_2}{\| Y \|_2}$$

memberikan

$$2 \frac{\| X \|_2}{\| Y \|_2} \sum_{i=1}^{N} x_i y_i \leq \| X \|_2^2 + \frac{\| X \|_2}{\| Y \|_2} \| Y \|_2^2 = 2 \| X \|_2^2$$

sehingga

$$2 \sum_{i=1}^{N} x_i y_i \leq 2 \| X \|_2^2 \frac{\| Y \|_2}{\| X \|_2} = 2 \| X \|_2 \| Y \|_2$$

jadi $\sum_{i=1}^{N} x_i y_i \leq \| X \|_2 \| Y \|_2$

misalkan x_i diganti dengan $-x_i$, meskipun $x_i y_i < 0$ dan sebut saja sebagai vektor $\overline{X}$ memberikan

$$\sum_{i=1}^{N} | x_i y_i | \leq \| \overline{X} \|_2 \| Y \|_2 = \| X \|_2 \| Y \|_2$$

$$\text{jadi} \quad \sum_{i=1}^{N} | x_i y_i | \leq \left\{ \sum_{i=1}^{N} x_i^2 \right\}^{1/2} \left\{ \sum_{i=1}^{N} y_i^2 \right\}^{1/2} \quad \blacksquare$$

pembuktian iv selanjutnya adalah

untuk setiap $X, Y \in R^N$,

$$\| X + Y \|_2^2 = \sum_{i=1}^{N} (x_i + y_i)^2 = \sum_{i=1}^{N} x_i^2 + 2 \sum_{i=1}^{N} x_i y_i + \sum_{i=1}^{N} y_i^2$$

dengan lemma diatas

$$\| X + Y \|_2^2 = \| X \|_2^2 + 2 \| X \|_2 \| Y \|_2 + \| Y \|_2^2$$

$$= (\| X \|_2 + \| Y \|_2)^2$$

sehingga $\| X + Y \|_2 \leq \| X \|_2 + \| Y \|_2$ $\blacksquare$

Teorema 4 (Ketaksamaan Cauchy - Schwartz)

Jika **X** dan **Y** adalah vektor -vektor pada sebuah ruang hasil kali dalam, maka $X^T Y \leq \| X \|_2 \| Y \|_2$.

bukti :

Ambil sebarang $\alpha \in R$,

$$(\alpha X + Y)^T (\alpha X + Y) = \sum_{j=1}^{N} (\alpha x_j + y_j)^2 \geq 0$$

$$\Rightarrow \alpha^2 \| X \|_2^2 + \alpha X^T Y + \| Y \|_2^2 \geq 0$$

Pandang sebagai kuadratis dalam α, ini non negatif untuk semua harga α, yang berarti bahwa

$$D \leq 0$$

$$(2 X^T Y)^2 - 4 \| X \|_2^2 \| Y \|_2^2 \leq 0$$

$$(X^T Y)^2 \leq \| X \|_2^2 \| Y \|_2^2$$

$$X^T Y \leq \| X \|_2 \| Y \|_2 \quad \blacksquare$$

Dari teorema (4), dapat dipergunakan untuk mendefinisikan sudut antara vektor **X** dan **Y**. Anggap bahwa X dan Y adalah vektor-vektor tak nol dalam V, kita mengambil

$$(X^T Y)^2 \le \| X \|_2^2 \| Y\|_2^2$$

$$\frac{(X^T Y)^2}{\| X \|_2^2 \| Y \|_2^2} \le 1$$

$$\left(\frac{X^T Y}{\| X \|_2 \| Y \|_2} \right)^2 \le 1 \quad \text{atau secara equivalen}$$

$$-1 \le \frac{X^T Y}{\| X \|_2 \| Y \|_2} \le 1$$

kini jika θ adalah sudut yang mengukur radian dari 0 hingga π , maka cosinus θ mengasumsikan setiap nilai antara -1 dan 1. Jadi kita peroleh

$$\cos \theta = \frac{X^T Y}{\| X \|_2 \| Y \|_2} \qquad 15$$

Dalam hal ini, bukan hanya mencari sudut – sudut diantara vektor dalam R^N dengan hasil kali dalam Euclidis, akan tetapi permasalahan utama yang tidak kalah pentingnya pada semua ruang hasil kali dalam adalah nenentukan apakah kedua vektor itu adalah ortogonal. Yaitu, apakah vektor tersebut mempunyai sudut $\theta = \pi/2$ diantara vektor-vektor tersebut.

Definisi 11

Jika X dan Y adalah vektor -vektor di ruang R^N dan θ adalah sudut diantaranya, maka hasil kali titik atau hasil kali dalam Euclidis X.Y didefinisikan oleh

$$X \cdot Y = \begin{cases} \| X \|_2 \| Y \|_2 \cos\theta & \text{jika } X \neq 0 \text{ dan } Y \neq 0 \\ 0 & \text{jika } X = 0 \text{ atau } Y = 0 \end{cases}$$

Definisi 12

Jika $X = (x_1, x_2, \ldots, x_N)^T$ dan $Y = (y_1, y_2, \ldots, y_3)^T$ adalah vektor-vektor dalam R^N, l_2 dan l_∞ jarak antara X dan Y didefinisikan oleh

$$\| X - Y \|_2 = \left\{ \sum_{i=1}^{N} (x_i - y_i)^2 \right\}^{1/2} \quad \text{dan}$$

$$\| X - Y \|_\infty = \max_{1 \le i \le N} | x_i - y_i | .$$

Konsep dari jarak dalam R^N juga digunakan untuk mendefinisikan suatu limit dari barisan vektor dalam ruang ini.

2.2 MATRIKS

Definisi 13

Sebuah matriks adalah susunan segi empat siku-siku dari bilangan-bilangan. Bilangan-bilangan dalam susunan tersebut dinamakan entri dalam matriks.

Definisi 14

Matriks A berukuran MxN adalah susunan bilangan berbentuk array segiempat dengan M baris dan N kolom.

Definisi 15

Elemen ke-i,j dari matriks A, didefinisikan a_{ij}, adalah elemen baris ke-i dan kolom ke-j dari matriks A. Dinotasikan sebagai $A_{MXN} = [\ a_{ij}]$ $i=1,2,...,M$

$j=1,2,...,N$

Definisi 16

Setiap kolom dari matriks didefinisikan sebagai vektor kolom dan setiap baris dari matriks didefinisikan sebagai vektor baris.

Notasi $X = (x_1, x_2, x_3, ..., x_N)^T$ adalah vektor kolom

$Y^T = (y_1, y_2, y_3, ..., y_N)$ adalah vektor baris

Definisi 17

Dua matriks A dan B adalah matriks sama jika kedua ukurannya sama, katakanlah NXM, dan jika $a_{ij} = b_{ij}$ untuk setiap $i = 1,2,...,N$ dan $j = 1,2,...,M$

Definisi 18

Jika A dan B adalah dua matriks yang ukurannya sama, maka jumlah A + B adalah matriks yang diperoleh dengan menambahkan bersama-sama entri yang bersesuaian dalam kedua matriks tersebut.

contoh $C = A + B$ dimana $c_{ij} = a_{ij} + b_{ij}$

Definisi 19

Jika A_{NXM} adalah suatu matriks dan λ adalah suatu skalar, maka hasil kali skalar dari λ dan A dinotasikan λA, adalah matriks MXN yang mana entri-entrinya adalah λa_{ij}, $i = 1,2,...,N$ $j=1,2,...,M$

Definisi 20

Jika A adalah matriks MXp dan B adalah matriks pXN, maka hasil kali A dan B dinotasikan AB, adalah C matriks $M \times N$ dimana entri-entri c_{ij} diberikan oleh

$$c_{ij} = \sum_{k=1}^{P} a_{ik}\ b_{kj}\ , \text{ untuk setiap } i = 1,2,\ldots,M$$
$$j = 1,2,\ldots,N$$

matriks hasil kali A dan B dinotasikan AB

Definisi 21

Matriks diagonal orde N adalah matriks $D = [d_{ij}]$ dengan ketentuan bahwa $d_{ij} = 0$, bilamana $i \neq j$

Definisi 22

A Matriks bujursangkar orde N adalah sebuah matriks dengan N baris dan N kolom dan entri-entri $a_{11}, a_{22}, \ldots, a_{NN}$ berada pada diagonal utama dari A.

Definisi 23

Matriks identitas dinotasikan dengan I, adalah matriks bujursangkar yang memiliki elemen diagonal semuanya satu dan semua elemen bebas diagonal adalah nol.

$$I_{NXN} = \begin{cases} i_{ij} = 0 \ , \text{ untuk } i \neq j \\ i_{ij} = 1 \ , \text{ untuk } i = j \end{cases}$$

Definisi 24

Matriks nol adalah matriks $M \times N$ yang mana semua elemennya sama dengan nol.

Definisi 25

U matriks $N \times N$ segitiga atas dimana $U = [\ u_{ij}]$ untuk setiap $j = 1,2,\ldots,N$, dan entri-entri u_{ij} untuk setiap $i = j+1, j+2, \ldots, N$ dan

L matriks segitiga bawah yaitu $L = [l_{ij}]$ untuk setiap $j = 1.2,\ldots,N$ dan entri-entri $u_{ij} = 0$. untuk setiap $i = 1,2,\ldots,j-1$

Definisi 26

Jika A adalah matriks bujursangkar, dan jika kita dapat mencari matriks B sehingga AB = BA = I , maka A dikatakan mempunayi invers dan B dinamakan invers matriks dari A.

Definisi 27

Determinan dari matriks bujursangkar A_{NXN} dinotasikan dengan $|\ A\ |$ adalah didefinisikan iterasi yang memenuhi $|\ A\ | = \sum_{i=1}^{N} m_{ij}M_{ij}$

disini M_{ij} adalah submatriks (kofaktor) dan m_{ij} didefinisikan sebagai $(-1)^{i+j}$ suatu determinan dari submatriks A dibentuk dengan menghapus baris ke-i dengan kolom j , dan dimana determinan dari skalar adalah skalar itu sendiri.

Definisi 28

Matriks A berdimensi MxN dapat dipartisi menjadi submatriks.

Contoh matriks A_{MXN} dapat dipartisi menjadi

$$A = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right]$$

dimana $A_{11} = N_1 \times N_1$, $A_{12} = M_1 \times N_2$, $A_{22} = M_2 \times N_2$,

$A_{21} = M_2 \times N_1$, $M = M_1 + M_2$, dan $N = N_1 \times N_2$

Definisi 29

A adalah matriks NXN dikatakan nonsingular jika determinannya tidak sama dengan nol. Jika determinannya sama dengan nol dikatakan singular.

Definisi 30

Jika A matriks bujursangkar nonsingular, invers dari A, ada , didefinisikan sebagai A^{-1} adalah matriks bujursangkar nonsingular sedemikian sehingga

$$A\,A^{-1} = A^{-1}\,A = I$$

Definisi 31

Transpose matriks $A = [\,a_{ij}]_{NXM}$, dinotasikan dengan A^T adalah dicapai dengan menukar baris dan kolom dari matriks A, yaitu $A^T = [\,a_{ij}]_{MXN}$, dimana $a^T_{ij} = a_{ji}$.

Definisi 32

A dan B matriks NXN dikatakan equivalen jika ada S matriks nonsingular sedemikian sehingga $A = S^{-1}B\,S$.

Definisi 33

A matriks NXN dikatakan simetris jika $A = A^T$, dimana $a_{ij} = a_{ji}$, untuk semua i dan j.

Definisi 34

A matriks MXN maka rank dari A adalah bilangan maksimum dari baris yang bebas linear, equivalen dengan bilangan terbesar dari kolom yang bebas linear dari matriks A. Jika rank dari A adalah sama dengan minimum {M,N} , maka A dikatakan mempunyai rank penuh.

Teorema 5

Jika A adalah matriks NxN yang memiliki invers, maka untuk setiap vektor B yang berdimensi N, sistem persamaan AX = B mempunyai persis satu pemecahan, yakni $X = A^{-1}B$.

Bukti :

Pandang $B = B$ dan $AA^{-1} = I$

$$IB = B$$

$$(AA^{-1})B = B$$

$$A(A^{-1}B) = B$$

maka $X = A^{-1}B$ adalah pemecahan $AX = B$.

Untuk membuktikan solusi tunggal, maka anggap X_o adalah sebarang solusi dan kemudian diperlihatkan bahwa X_o harus merupakan pemecahan $A^{-1}B$. Jika X_o suatu pemecahan, maka $AX_o = B$

$$AX_o = B$$

$$A^{-1}AX_o = A^{-1}B$$

$$X_o = A^{-1}B$$ 16

Jadi $X = X_o = A^{-1}B$ merupakan solusi tunggal. ■

Teorema 6

Jika P matriks transisi dari basis B' ke basis B, maka

a. P yang memiliki invers

b. P^{-1} adalah matriks transisi dari B ke B'

Bukti :

Misalkan Q adalah matriks transisi dari B ke B', akan diperlihatkan $PQ = I$ sehingga $Q = P^{-1}$. Anggap bahwa $B = (u_1, u_2, \ldots, u_N)$ dan misalkan

$$PQ = \begin{bmatrix} c_{11} & c_{12} \cdots & c_{1N} \\ c_{21} & c_{22} & c_{2N} \\ \vdots & \vdots & \vdots \\ c_{N1} & c_{N2} & c_{NN} \end{bmatrix}$$

maka $[X]_B = P\,[X]_{B'}$ dan $[X]_{B'} = Q\,[X]_B$, untuk setiap $X \in V$. Dimana $[X]_B$ menyatakan matriks X relatif terhadap basis B dan $[X]_{B'}$ menyatakan matriks X relatif terhadap basis B'. Sehingga

$P\,[X]_{B'} = P\,Q\,[X]_B$

$[X]_B = PQ\,[X]_B$, untuk setiap X di V

$I\,[X]_B = PQ\,[X]_B$

jadi $I = PQ$ atau $I = P\,P^{-1}$ ∎

Definisi 35

Jika A adalah matriks $N \times N$, maka vektor tak nol X didalam R^N dinamakan vektor eigen dari A jika AX adalah kelipatan skalar dari X ; yakni

$$AX = \lambda X \qquad 17$$

untuk suatu skalar λ. Skalar λ dinamakan nilai eigen dari A dan X dikatakan vektor eigen yang bersesuaian.

Untuk mencari nilai eigen matriks A, maka $AX = \lambda X$ dapat ditulis sebagai

$$AX = \lambda I X$$

atau secara equivalen

$$(A - \lambda I)\,X = 0 \qquad 17$$

yang mana mengakibatkan matriks $(A - \lambda I)$ adalah singular jika λ adalah nilai eigen dari A, berarti

$$\det(A - \lambda I) = 0$$

$$\det(A - \lambda I) = \lambda^N + c_1\lambda^{N-1} + c_2\lambda^{N-2} + \ldots + c_N = 0$$

dinamakan persamaan karakteristik dari A.

Definisi 36

Suatu matriks bujursangkar A dapat didiagonalisasi jika terdapat matriks P yang mempunyai invers sehingga $P^{-1}AP$ diagonal; matriks P dikatakan mendiagonalisasi A.

Teorema 6

Jika A adalah matriks $N \times N$, maka pernyataan - pernyataan berikut equivalen satu sama lain.

a. A dapat didiagonalisasi

b. A mempunyai N vektor eigen bebas linear

Bukti : a ⇒ b

Asumsi A dapat didiagonalisasi, maka terdapat matriks yang mempunyai invers, yaitu

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1N} \\ p_{21} & p_{22} & \cdots & p_{2N} \\ \vdots & \vdots & & \vdots \\ p_{N1} & p_{N2} & \cdots & p_{NN} \end{bmatrix}$$

sehingga $P^{-1}AP$ diagonal, katakanlah $P^{-1}AP = D$

dimana

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_N \end{bmatrix}$$

maka AP = PD ; yakni

$$AP = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1N} \\ p_{21} & p_{22} & \cdots & p_{2N} \\ \vdots & \vdots & & \vdots \\ p_{N1} & p_{N2} & \cdots & p_{NN} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \lambda_N \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 p_{11} & \lambda_2 p_{12} & \ldots\lambda_N p_{1N} \\ \lambda_1 p_{21} & \lambda_2 p_{22} & \ldots\lambda_N p_{2N} \\ \vdots & \vdots & \vdots \\ \lambda_1 p_{N1} & \lambda_2 p_{N2} & \ldots\lambda_N p_{NN} \end{bmatrix} \qquad 19$$

Misalkan $P_1, P_2, \ldots, P_N$ menyatakan vektor-vektor kolom P, maka bentuk persamaan (19) kolom-kolom AP yang berturutan adalah $\lambda_1 P_1$, $\lambda_2 P_2, \ldots, \lambda_N P_N$. Disisi lain kolom -kolom AP yang berturutan adalah AP_1, $AP_2, \ldots, AP_N$. Sehingga didapat

$$AP_1 = \lambda_1 P_1,\ AP_2 = \lambda_2 P_2,\ \ldots,\ AP_N = \lambda_N P_N \qquad 20$$

karena P mempunyai invers, maka vektor-vektor kolomnya semuanya tak nol. Jadi $\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_N$ adalah nilai-nilai eigen dari A, dan $P_1, P_2, P_3, \ldots, P_N$ adalah vektor - vektor eigen yang bersesuaian. Karena vektor-vektor kolom A membangun ruang kolom, maka vektor-vektor kolom A bebas linear. Jadi A mempunyai N vektor eigen bebas linear.

Bukti b ⇒ a

Asumsi A mempunyai N vektor eigen bebas linear, maka $P_1, P_2, P_3, \ldots, P_N$ dengan nilai eigen yang bersesuaian $\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_N$ dan misalkan

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1N} \\ p_{21} & p_{22} & \cdots & p_{2N} \\ \vdots & \vdots & & \vdots \\ p_{N1} & p_{N2} & \cdots & p_{NN} \end{bmatrix}$$

adalah matriks yang vektor -vektor kolomnya adalah $P_1, P_2, P_3, \ldots, P_N$. Sehingga kolom-kolom dari hasil kali AP adalah AP_1, $AP_2, \ldots, AP_N$. Tetapi $AP_1 = \lambda_1 P_1$, $AP_2 = \lambda_2 P_2$, $AP_3 = \lambda_3 P_3$, ... , $AP_N = \lambda_N P_N$, sehingga

$$AP = \begin{bmatrix} \lambda_1 p_{11} & \lambda_2 p_{12} & \cdots \lambda_N p_{1N} \\ \lambda_1 p_{21} & \lambda_2 p_{22} & \cdots \lambda_N p_{2N} \\ \vdots & \vdots & \vdots \\ \lambda_1 p_{N1} & \lambda_2 p_{N2} & \cdots \lambda_N p_{NN} \end{bmatrix}$$

$$= \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1N} \\ p_{21} & p_{22} & \cdots & p_{2N} \\ \vdots & \vdots & & \vdots \\ p_{N1} & p_{N2} & \cdots & p_{NN} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots 0 \\ 0 & \lambda_2 & \cdots 0 \\ & & \\ 0 & 0 & \cdots \lambda_N \end{bmatrix}$$

$$= PD \qquad (21)$$

dimana D adalah matriks diagonal yang mempunyai nilai-nilai eigen λ_1, λ_2, $\ldots, \lambda_N$ pada diagonal utama, karena vektor-vektor kolom dari P bebas linear, maka P mempunyai invers; jadi persamaan (21) dapat ditulis sebagai $P^{-1}AP = D$; yakni A terdiagonalisasi. ∎

Definisi 37

A matriks bujursangkar dinamakan dapat didiagonalisasi secara ortogonal jika terdapat matriks P ortogonal sehingga $P^{-1}AP$ ($= P^{T}AP$) diagonal; matriks P dikatakan mendiagonalisasi A.

Definisi 38

Himpunan vektor $\{ X_1, X_2, \ldots, X_N \}$ disebut ortogonal jika $(x_i)^T x_j = 0$ untuk setiap $i \neq j$. Jika ditambah dengan $(x_i)^T x_j = 1$, untuk semua $i,j = 1,2,3,\ldots,N$, maka himpunan tersebut disebut ortonormal.

Definisi 39

P matriks NXN dikatakan matriks ortogonal jika $P^{-1} = P^T$.

Definisi 40

A matriks bujursangkar dikatakan matriks ortogonal (ortonormal) jika kolom-kolomnya ortogonal (ortonormal).

Teorema 7

Jika matriks $Q = [\, q_1 \; q_2 \; q_3 \; \ldots \; q_N]$ dan Q adalah ortonormal maka $Q^TQ = I_{NXN}$.

Bukti :Ambil matriks $Q = [q_1 \; q_2 \; q_3 \; \ldots \; q_N]$ dan Q adalah ortonormal. sehingga

$$Q^TQ = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_N \end{bmatrix} [\; q_1 \;\; q_2 \; \ldots \; q_N \;]$$

$$= \begin{bmatrix} q_1q_1 & q_1q_2 & q_1q_3 & \cdots & q_1q_N \\ q_2q_1 & q_2q_2 & q_2q_3 & \cdots & q_2q_N \\ \vdots\;\vdots & \vdots\;\vdots & \vdots\;\vdots & & \vdots\;\vdots \\ q_Nq_1 & q_Nq_2 & q_Nq_3 & \cdots & q_Nq_N \end{bmatrix}$$

sesuai definisi (38) dimana $(q_i)^T \; q_j = 0$ untuk setiap $i \neq j$ dan $(q_i)^T q_j = 1$, untuk setiap $i = j$ dengan $i,j = 1,2,3,\ldots,N$. Selanjutnya didapat

$$Q^T Q = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix} = I_{NXN}$$ ∎

Teorema 8

Jika A adalah matriks NXN, maka pernyataan berikut equivalen satu sama lain

a. A dapat didiagonalisasi secara ortoganal

b. A mempunyai himpunan dari N vektor eigen

c. A adalah simetrik

Bukti a ⇒ b

Karena A dapat didiagonalisasi secara ortogonal, maka terdapat matriks P yang ortogonal sehingga $P^{-1}AP$ diagonal, maka N vektor kolom dari P adalah vektor eigen dari A. Karena P ortogonal maka vektor - vektor kolom ini ortonormal sehingga A mempunyai N vektor eigen ortonormal.

bukti b ⇒ c

Pandang A mempunyai himpunan ortogonal dari N vektor eigen $\{P_1, P_2, \ldots, P_N\}$, maka P ortogonal sehingga akan mendiagonalisasi A secara ortogonal. Matriks A ortonormal berarti kolom - kolomnya membentuk himpunan ortogonal dari vektor - vektor eigen yang berjumlah N. Kita misalkan bahwa D adalah matriks diagonal $D = P^{-1}AP$. Jadi $A = P\ D\ P^{-1}$ atau karena P ortogonal, maka

$$A = P\ D\ P^T$$

sehingga $A^T = (P\ D\ P^T)^T = P\ D^T P^T$

$$= P\ D\ P^T = A \qquad 22$$

yang menunjukkan bahwa A simetris.

bukti c $\longrightarrow$ a

Matriks simetris jelas suatu matriks bujursangkar, maka berdasarkan definisi (37) jelas bahwa A didiagonalisasi oleh P secara Ortogonal. ∎

Definisi 41

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

$$X^TAX = [\ x_1\ x_2\ x_3 \ldots\ x_N\ ]\ A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}$$

$$= a_{11}x_1^2 + a_{22}x_2^2 + \ldots + a_{NN}x_N^2 + \sum_{i \neq j} a_{ij}x_i x_j \qquad 23$$

dengan A adalah matriks simetrik NXN.

Definisi 42

Bentuk kuadratik X^TAX kita sebut definit positif jika $X^TAX > 0$ untuk setiap $X \neq 0$, sedangkan matriks simetris A kita sebut matriks definit positif jika X^TAX adalah bentuk kuadratik definit positif.

contoh $A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$

Definisi 43

A matriks semi definit positif jika hanya jika bentuk kuadratik $X^TAX \geq 0$ untuk semua X dan terdapat $X \neq 0$ sedemikian sehingga $X^TAX = 0$.

contoh $A = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$

Definisi 44

A adalah matriks definit negatif jika hanya jika $-A$ adalah definit positif. Dengan kata lain , A adalah definit negatif ketika $X^TAX < 0$, untuk semua $X \neq 0$. contoh $A = \begin{bmatrix} -2 & 1 \\ 1 & -3 \end{bmatrix}$

Definisi 45

A adalah matriks semi definit negatif jika $-A$ adalah semi definit positif.

contoh $A = \begin{bmatrix} -1 & 1 \\ 1 & -3 \end{bmatrix}$

Definisi 46

A adalah matriks indefinit jika $X^TA\ X$ adalah positif untuk beberapa X dan negatif untuk X yang lain. contoh $A = \begin{bmatrix} 1 & -1 \\ 1 & -2 \end{bmatrix}$

Faktorisasi Cholesky L^TL dan LDL^T untuk matriks A simetris definit positif.

Definisi 47

Faktorisasi Cholesky dari matriks simetris definit positif dapat ditulis sebagai $A = L\ L^T$

dimana L adalah matriks segitiga bawah berbentuk

$$L = \begin{bmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ \vdots & \vdots & & \\ l_{N1} & l_{N2} & \cdots & l_{NN} \end{bmatrix}$$

sedemikian sehingga

$$\begin{bmatrix} a_{11} & a_{12} \dots a_{1N} \\ a_{21} & a_{22} \dots a_{2N} \\ \vdots & \vdots \quad\quad \vdots \\ a_{N1} & a_{N2} \dots a_{NN} \end{bmatrix} = \begin{bmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ \vdots & \vdots & \ddots & \\ l_{N1} & l_{N2} & \dots & l_{NN} \end{bmatrix} \begin{bmatrix} l_{11} & l_{12} & \dots & l_{1N} \\ & l_{22} & \dots & l_{2N} \\ & & \ddots & \vdots \\ & & & l_{NN} \end{bmatrix}$$

$$\text{atau} \quad a_{ij} = \sum_{k=1}^{\min(i,j)} l_{ik} l^T_{kj} = \sum_{k=1}^{\min(i,j)} l_{ik} l_{jk} \qquad 25$$

Faktorisasi dapat disajikan juga dalam bentuk $A = L\,D\,L^T$ dimana L adalah matriks segitiga bawah dan D adalah matriks diagonal, $D = \text{diag}(\, d_1 \; d_2 \; \dots \; d_N)$ yang keduanya memiliki entri positif, dan A matriks definit positif. Dapat ditulis sebagai :

$$A = L\,D\,L^T = (L\,D^{1/2})\,(L\,D^{1/2})^T = L'\,L'^T \qquad 26$$

dimana elemen-elemen dari $D^{1/2}$ merupakan akar kuadrat dari setiap elemen matriks diagonal positif D, yaitu $D = \text{diag}\;(\, d_1 \; d_2 \dots d_N)$, $d_j > 0$, dan L' adalah matriks segitiga atas. Kedua representasi dari faktor Cholesky tersebut adalah equivalen. Jika $D = \text{diag}(\; d_1 \; d_2 \dots \; d_N)$ matrik diagonal, dijumpai $d_j < 0$ dapat diambil konstanta positif kecil, yang menghasilkan matrik definit positif $\bar{A} = A + E$, dimana E adalah matriks diagonal nonnegatif.

Definisi 48

Matriks A simetris dengan harga eigen semuanya real dan vektor eigennya berbentuk basis ortonormal dalam R^N. Ambil harga eigen dari A berkorespondensi dengan elemen - elemen matriks diagonal

$W = \mathrm{diag}(w_1 \quad w_2 \ldots \quad w_N)$ dengan himpunan vektor eigennya adalah $V = \{V_1, V_2, \ldots, V_N\}$ dan didefinisikan :

$$AV_i = w_i V_i \;, \; i = 1,2,\ldots,N \qquad 27$$

dan dapat ditulis dalam bentuk matriks

$$AV = WV \qquad 28$$

selama V ortonormal, $VV^T = I$ dan persamaan (28) menghasilkan $A = V W V^T$ dan sebaliknya $V^T A V = W$ adalah transformasi khusus, atau secara eksplisit dapat disajikan dalam bentuk

$$A = \sum_{i=1}^{N} w_i V_i V_i^T \;, \qquad 29$$

adalah dekomposisi spektral dari matriks A.

2.4 HIMPUNAN KONVEKS

Definisi 49

Persekitaran ε dari X dalam R^N didefinisikan sebagai himpunan dari semua titik Y dalam R^N sedemikian sehingga $\| Y - X \|_2 < \varepsilon$, dimana $\varepsilon > 0$, $\varepsilon \in R$.

Definisi 50

Barisan $\{ X_k \}_{k=1}^{\infty}$ dari vektor dalam R^N dikatakan konvergen ke X terhadap norma $\| \cdot \|_2$ jika diberikan suatu $\varepsilon > 0$, terdapat bilangan integer $N(\varepsilon)$ sedemikian sehingga $\| X_k - X \|_2 < \varepsilon$, untuk setiap $k \geq N(\varepsilon)$.

Definisi 51

Ambil titik X_1, X_2 dalam $S \subseteq R^N$, himpunan $L = \{ X \mid X = (1-\lambda)X_1 + \lambda X_2 \}$, $\lambda \in R$, adalah suatu garis dalam R^N yang melewati X_1 dan X_2 .

Definisi 52

Titik $X = (1-\lambda)X_1 + \lambda X_2$, $0 \le \lambda \le 1$, disebut kombinasi linear konveks dari X_1 dan X_2.

Definisi 53

Himpunan $K \subseteq R^N$ dikatakan himpunan konveks jika setiap kombinasi linear konveks dari dua titik dalam K adalah anggota K. Dengan kata lain setiap $X_1, X_2 \in K$, dan $\lambda \in R$ dimana $0 \le \lambda \le 1$ berlaku $X = (1 - \lambda)X_1 + \lambda X_2 \in K$.

Definisi 54

Ambil $X_i \in R^N$ dan ambil λ_i bilangan nonnegatif sedemikian sehingga $\sum_{i=1}^{M} \lambda_i = 1$, maka

$$X = \lambda_1 X_1 + \lambda_2 X_2 + \lambda_3 X_3 + \dots + \lambda_M X_M$$

disebut kombinasi linear konveks dari titik - titik X_i ,untuk $i = 1,2,\dots,M$.

Teorema 10

Irisan dari dua himpunan konveks adalah himpunan konveks.

Bukti :

Ambil k_1 dan k_2 dua himpunan konveks, ambil $X_1, X_2 \in k_1 \cap k_2$ dan diambil

$$X = (1 - \lambda)X_1 + \lambda X_2,\ 0 \le \lambda \le 1$$

$$X_1, X_2 \in k_1 \longrightarrow X \in k_1$$

$X_1, X_2 \in k_2 \longrightarrow X \in k_2$

sehingga $X_1, X_2 \in k_1 \cap k_2$ maka $X \in k_1 \cap k_2$. ∎

Definisi 55

Titik X dari himpunan konveks K adalah titik ekstrim dari K jika tidak mungkin untuk mencari dua titik X_1, X_2 dalam K , sedemikian sehingga

$$X = (1 - \lambda) X_1 + \lambda X_2,\ 0 \le \lambda \le 1.$$

Teorema 11

Ambil K himpunan tertutup dalam R^N , dan titik $X_o \notin K$, maka ada hiperplane , yang disebut hiperplane pemisah yang mana memuat X_o sedemikian sehingga K termuat dalam salah satu setengah ruang yang dihasilkan dari hiperplane itu. ($CX = \alpha$), $X \in K$, $CX < \alpha$, $CX_o > \alpha$.

Bukti :

Ambil W titik tunggal dalam K terdekat menuju X_o sedemikian sehingga $| W - X_o| \le | X - X_o|$, X suatu titik dalam K. Juga untuk suatu titik $X \in K$,

$$Y = \lambda X + (1 - \lambda)W,\ 0 \le \lambda \le 1.$$

adalah dalam K.

oleh karena itu, $| W - X_o| \le | Y - X_o|$

atau $| W - X_o| \le | \lambda X + (1 - \lambda)W - X_o|$

$$| W - X_o|^2 \le | W - X_o + \lambda(X - W)|^2$$

$\lambda^2(X - W)^T(X - W) \ge -2(W - X_o)^T(X - W)$ atau

$\lambda(X - W)^T(X - W) \ge -2C\ (X - W\ )$

dimana $C = (W - X_o)^T$ adalah vektor konstan sebab X_o ditentukan dan W adalah tunggal dalam K.

Sekarang ambil $\lambda \longrightarrow 0$, maka dalam limit

$$C(X - W) \geq 0 \text{ atau } CX \geq CW$$

tetapi selama $C(W - X_o) = CC' > 0$, $CW > CX_o$

selanjutnya $CX \geq CW > CX_o = \alpha$

Jadi untuk sebarang titik $X \in K$, $CX > \alpha$, yaitu K bergantung dalam suatu setengah bidang yang dihasilkan oleh hiperplane $CX = \alpha$. ∎

2. 5 TEKNIK OPTIMASI TAK BERKENDALA

Metode optimasi adalah cara mencari nilai optimum dari suatu fungsi. Studi kalkulus dari bentuk optimasi adalah dasar untuk membangun sebagian besar teknik numerik dari optimasi. Dalam pasal ini dijabarkan syarat perlu dan syarat cukup tentang nilai optimum dari fungsi univariabel maupun multivariabel yang tak berkendala.

2.5.1 OPTIMASI MASALAH UNIVARIABEL TAK BERKENDALA

Berikut ini disajikan beberapa definisi dan teorema berikut tentang minimal lokal maupun global dari fungsi univariabel.

Definisi 56

Misalkan f(x) terdefinisi pada interval I, maka

i. Jika $x_1 < x_2$ mengakibatkan bahwa $f(x_1) < f(x_2)$, untuk setiap $x_1, x_2 \in I$ maka f dikatakan fungsi monoton naik pada interval I.

ii. Jika $x_1 < x_2$ mengakibatkan bahwa $f(x_1) > f(x_2)$, untuk setiap $x_1, x_2 \in I$ maka f dikatakan fungsi monoton turun pada I.

Definisi 57

Fungsi $f \in R$ dikatakan mempunyai harga minimum lokal pada $x = x^*$ jika terdapat persekitaran $N(x^*, \varepsilon)$ yang memuat x^* sedemikian sehingga $f(x^*) \leq f(x)$ untuk setiap x dalam persekitaran $N(x^*, \varepsilon)$.

Definisi 58

Fungsi f dikatakan mempunyai harga minimum global pada $x = x^*$ jika $f(x^*) \leq f(x)$ untuk setiap $x \in R^N$.

Definisi 59

Fungsi $f \in R^N$, jika $f(x^*) < f(x)$ untuk setiap $x \in N(x^*, \varepsilon)$, $x \neq x^*$, untuk suatu $\varepsilon > 0$, maka x^* disebut minimum lokal terbatasi.

Teorema 12

Misalkan $f : R \longrightarrow R$ mempunyai turunan dengan orde $(N+1) f^{(N+1)}(x)$ diselang $(a-r, a+r)$ dan misalkan ada konstanta $M > 0$ demikian sehingga berlaku $| f^{(N+1)}(x) | \leq M$, untuk semua x diselang tersebut. Maka untuk setiap x diselang itu, f(x) dapat diuraikan menjadi bentuk

$$f(x) = f(a) + f'(a)(x-a) + f''(a)(x-a)^2/2! + \dots + f^{(N)}(a)(x-a)^N/N! + R_N \qquad 30$$

$$\text{dengan } | R_N | \leq M \frac{| x - a |^{(N+1)}}{(N+1)!} \qquad 31$$

Bukti :

$| f^{(N+1)}(x) | \leq M$ kita tulis sebagai

$- M \leq f^{(N+1)}(x) \leq M$

Jika semua ruas diintegrasikan dari a hingga x , kita peroleh

$- M (x - a) \leq f^{(N)}(x) - f^{(N)}(a) \leq M(x - a)$

Hasil ini kita integrasikan lagi dari a sampai x, maka didapat :

$$\frac{-M (x - a)^2}{2!} \leq f^{(N-1)}(x) - f^{(N-1)}(a) - f(a)^{(N)}(x-a) \leq \frac{M (x - a)^2}{2!}$$

Proses integrasi ini kita ulang (N-1) kali lagi dan hasilnya sebagai berikut :

$$\frac{-M (x - a)^{(N+1)}}{(N + 1)!} \leq f(x) - f(a) - f'(a)(x-a) - \ldots$$

$$f''(a)(x-a)^2/2! - f^{(N)}(a) \frac{(x-a)^N}{N!} \leq \frac{M (x - a)^{(N+1)}}{(N + 1)!}$$

Jika ruas tengah $f(x) - f(a) - f'(a)(x-a) -$

$$f''(a)(x-a)^2/2! - \ldots - f^{(N)}(a) \frac{(x - a)^N}{N !} = R_N$$

maka diperoleh :

$$f(x) = f(a) + f'(a)(x-a) + f''(a)(x-a)^2/2! + \ldots + f^{(N)}(a) \frac{(x-a)^N}{N!} + R_N$$

dengan $| R_N | \leq M \dfrac{| x - a |^{N+1}}{(N+1)!}$

Hasil ini kita sebut deret Taylor untuk f(x) disekitar x = a , dengan R_N sebagai suku sisa. ∎

Teorema 13

Jika fungsi f(x) terdefinisi dalam interval $a \le x \le b$ dan mempunyai minimum lokal pada $x = x^*$, dimana $a < x^* < b$, dan jika derivatif $df(x)/dx = f'(x)$ ada (sebagai bilangan berhingga) pada $x = x^*$, maka $f'(x) = 0$.

Bukti : Pandang

$$f'(x^*) = \lim_{\delta \to 0} \frac{f(x^* + \delta) - f(x^*)}{\delta} \qquad 32$$

ada, sebagai bilangan tertentu yang mana akan kita buktikan menuju nol. Selama x^* adalah minimum lokal, kita dapatkan

$$f(x^*) \le f(x^* + \delta) \qquad 33$$

untuk semua harga dari δ menuju nol.

sehingga $$\frac{f(x^* + \delta) - f(x^*)}{\delta} \ge 0 \quad \text{jika } \delta > 0$$

$$\frac{f(x^* + \delta) - f(x^*)}{\delta} < 0 \quad \text{jika } \delta < 0$$

Jadi dari persamaan (33) dengan memberikan limit δ menuju nol memenuhi harga positif seperti

$$f'(x^*) \ge 0 \qquad 34$$

dilain pihak persamaan (33) untuk limit δ menuju nol memenuhi harga negatif seperti

$$f'(x^*) \le 0 \qquad 35$$

Satu-satunya yang memenuhi keduanya (persamaan 34 dan 35) adalah

$$f'(x^*) = 0 \qquad 36$$

■

Teorema 14

Ambil $f'(x^*) = f''(x^*) = \ldots = f^{(N-1)}(x^*) = 0$, tetapi $f^{(N)}(x^*) \neq 0$, maka $f(x^*)$ adalah

i. harga minimum dari $f(x)$ jika $f^{(N)}(x^*) > 0$ dan N adalah genap.

ii. harga maksimum dari $f(x)$ jika $f^{(N)}(x^*) < 0$ dan N adalah genap.

iii. kedua-duanya maksimum atau minimum jika N adalah ganjil.

Bukti :

Sesuai Teorema Taylor dengan suku sisa setelah N suku, kita dapatkan

$$f(x^*+\delta) = f(x^*) + \delta f'(x^*) + \frac{\delta^2}{2!} f''(x^*) + \ldots$$

$$+ \frac{\delta^{N-1}}{(N-1)!} f^{(N-1)}(x^*) + \frac{\delta^{N-1}}{(N-1)!} f^{(N-1)}(x^*+t\delta)$$

untuk $0 < t < 1$. 37

sejak $f'(x^*) = f''(x^*) = \ldots = f^{(N-1)}(x^*) = 0$, persamaan (37) menjadi

$$f(x^*+\delta) - f(x^*) = \frac{\delta^N}{N!} f^{(N)}(x^*+t\delta)$$

Misalkan $f^{(N)}(x^*) \neq 0$, terdapat interval yang melingkupi x^* untuk setiap titik x yang mana derivatif ke-N dari $f^{(N)}(x)$ mempunyai tanda sama, katakanlah bahwa $f^{(N)}(x^*)$. Jadi untuk setiap titik $x^*+\delta$ dalam interval ini, $f^{(N)}(x^*+t\delta)$ mempunyai tanda sama seperti $f^{(N)}(x^*)$. Ketika N adalah genap , $\frac{\delta^{(N)}}{N!}$ adalah positif dengan tak menghiraukan keadaan δ positif atau negatif dan selanjutnya $f(x^*+\delta) - f(x^*)$ akan memiliki tanda sama dengan $f^{(N)}(x^*)$. Jadi x^* akan meminimum lokal jika $f^{(N)}(x^*)$ adalah positif dan maksimum lokal jika $f^{(N)}(x^*)$ adalah negatif.

Saat N ganjil , $\frac{\delta^{(N)}}{N!}$ berarti tanda dari h berubah sehingga titik x^* kedua-duanya tidak maksimum atau minimum. Dalam hal ini titik x^* disebut titik belok. ∎

2.5.2 OPTIMASI MASALAH MULTIVARIABEL TAK BERKENDALA

Masalah tak berkendala dalam hal ini adalah masalah minimasi F(**X**) tanpa suatu kendala pada **X**. Berikut beberapa definisi dan teorema tentang minimasi lokal maupun global dari masalah tak berkendala.

Definisi 60

Jika untuk setiap vektor **X** dalam $D \subseteq R^N$, terdapat bilangan real F(**X**) tunggal maka F(**X**) dikatakan fungsi berharga real dari **X**.

Dinotasikan $F: D \subseteq R^N \longrightarrow R$.

Definisi 61

F(**X**) dikatakan kontinu pada $\mathbf{X}_o$ jika untuk setiap bilangan real $\varepsilon>0$, terdapat bilangan $\eta>0$, sedemikian sehingga

$$\| \mathbf{X} - \mathbf{X}_o \|_2 < \eta \longrightarrow \| F(\mathbf{X}) - F(\mathbf{X}_o) \|_2 < \varepsilon$$

Definisi 62

Misalkan $K \subset R^N$ himpunan konveks, F(**X**) terdefinisi di K untuk setiap $X_1, X_2 \in K$, $\lambda \in R$, $0 \leq \lambda \leq 1$.

Fungsi F disebut konveks jika

$$F(\lambda X_1 + (1-\lambda)X_2) \leq \lambda F(X_1) + (1-\lambda)F(X_2),$$

untuk $\lambda \in R$, $0 \leq \lambda \leq 1$.

Teorema 15

Bila F(X) fungsi konveks pada himpunan K maka F mempunyai paling banyak satu minimum lokal. Bila titik minimum adamaka titik minimum itu merupakan minimal global pada himpunan konveks itu..

Bukti :

Misalkan X_o titik minimum lokal dari F(X), dan F fungsi konveks pada K, maka untuk setiap $X \in K$ dan $\lambda \in R$ berbentuk

$$F(X_o) \leq F((1-\lambda)X_o + \lambda X), \qquad 38$$

Karena F(X) fungsi konveks maka

$$F((1-\lambda)X_o + \lambda X) \leq (1-\lambda)F(X_o) + \lambda F(X) \qquad 39$$

Dari persamaan (38) dan (39) didapat

$$F(X_o) \leq (1-\lambda)F(X_o) + \lambda F(X)$$

atau $\lambda F(X_o) \leq \lambda F(X)$, $0 \leq \lambda \leq 1$.

Khusus untuk $\lambda > 0$ diperoleh $F(X_o) \leq F(X)$ untuk setiap $X \in K$. Jadi F(X) mempunyai minimum global di X_o. Ini mengakibatkan bahwa paling banyak ada satu penyelesaian minimum lokal.

Misalkan $K_1 = \{X \in K \mid F$ mencapai minimum lokal di $X\}$

Akan ditunjukkan bahwa K_1 konveks.

Misalkan $X_o \in K_1$, $X_1 \in K_1$, Z_o nilai minimum F(X), untuk $0 \leq \lambda \leq 1$ berlaku

$$Z_o \leq F((1-\lambda)X_o + \lambda X_1) \leq (1-\lambda)F(X_o) + \lambda F(X_1)$$

$= (1 - \lambda)Z_o + \lambda Z_o$

Jadi $F((1 - \lambda)X_o + \lambda X_1) = Z_o$ untuk setiap $0 \le \lambda \le 1$

Ini berarti bahwa $(1 - \lambda)X_o + \lambda X_1 \in K_1$ yaitu K_1 konveks. ∎

Definisi 63

Derivatif parsial dari F(**X**) terhadap komponen x_j dari **X** didefinisikan sebagai

$$\lim_{\delta x_i \to o} \frac{F(\mathbf{X} + \Delta\mathbf{X}_j) - F(\mathbf{X})}{\delta x_j}$$

dan dinotasikan sebagai $\partial F / \partial x_j$,

dimana : $\Delta\mathbf{X}_j = [\ 0\ ,\ 0\ , \ldots, \delta x_j, \ \ldots,\ 0\ ,\ 0\ ]^T$

$j = 1,2,3,\ldots,N$

Vektor yang komponennya terdiri dari N derivatif parsial dari F(**X**) disebut gradien dari F(**X**) :

$$\text{grad } F = \nabla F = \left[\frac{\partial F}{\partial x_1} \quad \frac{\partial F}{\partial x_2} \quad \frac{\partial F}{\partial x_3} \quad \cdots \quad \frac{\partial F}{\partial x_N} \right]^T$$

Definisi 64

Jika F(**X**) mempunyai derivatif parsial kontinu terhadap masing masing variabelnya, maka dikatakan differensiabel.

Ambil F(**X**) fungsi differensial berharga real dari **X** dalam R^N. Ambil $\mathbf{X} + \Delta\mathbf{X}$ titik sekitar **X** dalam R^N sedemikian sehingga

$$\Delta\mathbf{X} = [\ \delta x_1 \ \ \delta x_2 \ \ \delta x_3 \ldots \ \delta x_N]^T$$

$$\mathbf{X} + \Delta\mathbf{X} = [\ x_1 + \delta x_1 \quad x_2 + \delta x_2 \quad x_3 + \delta x_3 \ldots \ x_N + \delta x_N]^T$$

Analogi teorema deret Taylor diatas untuk fungsi N variabel dapat ditulis sebagai

$$F(X + \Delta X) = F(X) + (\Delta X)^T \nabla F(X) + \frac{1}{2}(\Delta X)^T H(X)(\Delta X) + e(X,\Delta X)|\Delta X|^2$$

dimana $e(X,\Delta X) \longrightarrow 0$ untuk $|\Delta X| \longrightarrow 0$ dan

$$\nabla F(X) = \left[\frac{\partial F}{\partial x_1} \quad \frac{\partial F}{\partial x_2} \quad \frac{\partial F}{\partial x_3} \quad \cdots \quad \frac{\partial F}{\partial x_N} \right]^T$$

$$H(X) = \begin{bmatrix} \frac{\partial^2 F}{\partial x_1^2} & \frac{\partial^2 F}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 F}{\partial x_1 \partial x_N} \\ \frac{\partial^2 F}{\partial x_2 \partial x_1} & \frac{\partial^2 F}{\partial x_2 \partial x_2} & \cdots & \frac{\partial^2 F}{\partial x_2 \partial x_N} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 F}{\partial x_N \partial x_1} & \frac{\partial^2 F}{\partial x_N \partial x_2} & \cdots & \frac{\partial^2 F}{\partial x_N^2} \end{bmatrix}$$

Sehingga dapat didefinisikan sebagai berikut:

Definisi 65

Differensial ke-r dari F, jika semua derivatif parsial dari fungsi F memenuhi orde $r \geq 1$ ada dan kontinu pada titik X^*, maka polinomial

$$d^r F(X^*) = \sum_{i=1}^{N} \sum_{j=1}^{N} \cdots \sum_{k=1}^{N} \delta_i \delta_j \cdots \delta_k \frac{\partial^r F(X^*)}{\partial x_i \partial x_j \cdots \partial x_k}$$

disebut differensial ke-r dari F pada X^*, sehingga ekspansi deret Taylor dari F(X) disekitar X^* diberikan :

$$F(X) = F(X^*) + \frac{1}{2!} d^2 F(X^*) + \frac{1}{3!} d^3 F(X^*) + \cdots + \frac{1}{N!} d^N F(X^*) + R_N(X^*,\delta)$$

dimana $R_N(X^*,\delta) = \frac{1}{N!} d^{N+1} F(X^* + t\delta)$, untuk $0 < t < 1$ dan $\delta = X - X^*$.

Syarat perlu dan cukup untuk meminimumkan fungsi tersebut disajikan dalam teorema berikut :

Teorema 16

Misalkan $F : D \subseteq R^N \longrightarrow R$ adalah differensiabel pada X^*. Jika vektor δ membuat $\nabla F(X^*)^T\delta < 0$, maka terdapat $\varepsilon > 0$ sehingga $F(X^* + t\delta) < F(X^*)$ untuk tiap - tiap $t \in (0,\varepsilon)$, sedemikian sehingga δ adalah arah descent dari F pada X^*.

Bukti :

Misalkan F differensiabel pada X^*, yaitu

$$F(X^* + t\delta) = F(X^*) + t\nabla F(X^*)^T \delta + t\| \delta \|_2^2 N(X^*;t\delta)$$

dimana $N(X^*;t\delta) \longrightarrow 0$, $t \longrightarrow 0$. Sehingga

$$\frac{F(X^* + t\delta) - F(X^*)}{t} = \nabla F(X^*)^T \delta + \| \delta \|_2^2 N(X^*;t\delta),$$

untuk $t \neq 0$. Selama $\nabla F(X^*)^T\delta < 0$ dan $N(X^*;t\delta) \longrightarrow 0$, terdapat $\varepsilon > 0$ sedemikian sehingga

$\nabla F(X^*)^T\delta + \|\delta\|_2^2 N(X^*;t\delta) < 0$, untuk setiap $t \in (0,\varepsilon)$.

sehingga $F(X^* + t\delta) - F(X^*) < 0$ maka

$$F(X^* + t\delta) < F(X^*).$$ ∎

Teorema 17

Jika F(X) mempunyai titik ekstrim (maksimum/minimum) pada $X = X^*$ dan jika derivatif parsial pertama dari F(X) ada pada X^*, maka

$$\frac{\partial F(X^*)}{\partial x_1} = \frac{\partial F(X^*)}{\partial x_2} = \ldots = \frac{\partial F(X^*)}{\partial x_N} = 0 \qquad 40$$

atau $\nabla F(X^*) = 0$

Bukti :

Misalkan bahwa ada satu dari derivatif parsial pertama, katakanlah ke-k, yang tidak hilang pada

X^*,maka dengan teorema Taylor :

$$F(X^*+\delta) = F(X^*) + \sum_{i=1}^{N} \delta_i \frac{\partial F(X^*)}{\partial x_i} + R_i(X^*,\delta)$$

$$\Leftrightarrow F(X^*+\delta) - F(X^*) = \delta_k \frac{\partial F(X^*)}{\partial x_k} + \frac{1}{2!} d^2F(X^*+t\delta)$$

untuk $0 < t < 1$. 41

Selama $d^2F(X^*+t\delta)$ adalah orde δ_i^2, suku dari orde k akan mendominasi suku orde yang lebih tinggi untuk δ kecil. Jadi tanda dari $F(X^*+\delta) - F(X^*)$ ditentukan oleh tanda $\delta_k\ \partial F(X^*)/\partial x_k$. Misalkan

$\partial F(X_k^*)/\partial x > 0$, maka $F(X^*+\delta) - f(X^*) > 0$, $\delta_k > 0$

maka $F(X^*+\delta) - f(X^*) < 0$, $\delta_k < 0$

Ini berarti X^* bukan titik ekstrim.

Kesimpulan secara analogi dapat dicapai juga jika kita asumsikan bahwa $\partial F(X^*)/\partial x < 0$. Karena kesimpulan ini maka terjadi kontradiksi dengan statemen bahwa X^* adalah titik ekstrim, katakanlah $\partial F/\partial x_k = 0$ pada $X = X^*$. ■

Akibat 1

Misalkan $F : R^N \longrightarrow R$ adalah differensiabel pada X^*. Jika X^* adalah minimal lokal, maka $\nabla F(X^*) = 0$.

Bukti :

Misalkan $\nabla F(X^*) \neq 0$. Ambil $\delta = -\nabla F(X^*)$, didapatkan

$$\nabla F(X^*)^T\delta = -\| \nabla F(X^*)\|_2^2 < 0.$$

maka terdapat $\varepsilon>0$ sedemikian sehingga $F(X^*+t\delta) < F(X^*)$ untuk $t \in (0,\varepsilon)$. Kontradiksi dengan asumsi bahwa X^* adalah minimal lokal. Jadi $\nabla F(X^*) = 0$. ■

Teorema 18

Misalkan bahwa $F : D \subseteq R^N \longrightarrow R$ differensiabel kedua pada X^*. Jika X^* minimal lokal maka $\nabla F(X^*)=0$ dan $H(X^*)$ adalah semi definit positif.

Bukti :

Pikirkan sebarang arah δ, maka F differensiabel pada X^*, didapat

$$F(X^* + t\delta) = F(X^*) + t\nabla F(X^*)^T\delta + \frac{1}{2}t^2\delta^T H(X^*)\delta + t^2\|\delta\|_2^2 N(X^*;t\delta) \quad 42$$

dimana $N(X^*;t\delta) \longrightarrow 0$, untuk $t \longrightarrow 0$.

Selama X^* minimal lokal, dari teorema (17), kita dapatkan $\nabla F(X^*) = 0$. Persamaan (42) dapat ditulis dalam bentuk :

$$\frac{F(X^*+t\delta)-F(X^*)}{t^2} = \frac{1}{2}\delta^T H(X^*)\delta + \|\delta\|_2^2 N(X^*;t\delta) \quad 43$$

Sejak X^* minimal lokal, maka $F(X^* + t\delta) \geq F(X^*)$, untuk t cukup kecil. Dari persamaan (43) ternyata

$$\lim_{t \to 0} \frac{F(X^* + t\delta) - F(X^*)}{t^2} = \frac{1}{2}\delta^T H(X^*)\delta \geq 0 \quad 44$$

sehingga $H(X^*)$ adalah semi definit positif. ∎

Teorema 19

Misalkan bahwa $F : D \subseteq R^N \longrightarrow R$ differensiabel kedua pada X^*. Jika $\nabla F(X^*) = 0$ dan $H(X^*)$ adalah definit positif, maka X^* adalah minimum lokal terbatasi.

Bukti :

Pandang F adalah differensiabel kedua pada X^*, kita meski dapatkan, untuk setiap $X \in R^N$,

$$F(X) = F(X^*) + \nabla F(X^*)^T(X - X^*)$$
$$+ \frac{1}{2}(X-X^*)^2 H(X^*)(X-X^*) + \|X - X^*\|_2^2 \; N(X^*;t\delta)$$

dimana $N(X^*;X - X^*) \longrightarrow 0$, untuk $X \longrightarrow X^*$. 45

Misalkan dengan pengandaian, bahwa X^* tidak minimum lokal terbatasi; yaitu misalkan terdapat barisan $\{ X_k \}$ konvergen menuju X^* sedemikian sehingga $F(X_k) \leq F(X^*)$, $X_k \neq X^*$, untuk setiap k.

Pikirkan barisan tersebut. dengan catatan bahwa $\nabla F(X^*) = 0$ dan $F(X_k) \leq F(X^*)$ dan dinotasikan $\delta_k = (X_k - X^*)/ \| X_k - X^*\|_2$. Persamaan (45) mengakibatkan bahwa

$$\frac{1}{2}\delta^T H(X^*)\delta + N(X^*; X_k - X^*) \leq 0, \; \forall \; k. \qquad 46$$

Tetapi $\| \delta_k \|_2 = 1$, $\forall$ k; sehingga $\{ X_k \}_K$ konvergen menuju δ, dimana $\| \delta \|_2 = 1$. Pikirkan subbarisan $k \in K$ mendekati ∞ , maka persamaan (46) mengakibatkan bahwa

$$\delta^T H(X^*)\delta \leq 0$$

Hal ini kontradiksi dengan asumsi bahwa $H(X^*)$ adalah definit positif, sejak $\| \delta \|_2 = 1$. Oleh karena itu X^* sesungguhnya minimum lokal terbatasi.

∎

2.5.3 METODE STEEPEST DESCENT

Metode Steepest Descent adalah salah satu dari sebagian besar prosedur untuk meminimalkan fungsi differensiabel multivariabel. Vektor δ disebut arah menurun dari fungsi ϕ pada X. Jika terdapat $\varepsilon > 0$ sedemikian sehingga $\phi(X + t\delta) < \phi(X)$, $\forall X \in (0,\varepsilon)$, yaitu

$$\lim_{t \to 0^+} \frac{\phi(X + t\delta) - \phi(X)}{t} < 0,$$

maka δ adalah arah penurunan. Metode Steepest Descent bergerak sepanjang arah δ dengan $\| \delta \|_2 = 1$, yang meminimalkan atas limit.

Lemma 2

Misalkan bahwa $\phi : R^N \longrightarrow R$ adalah differensiabel pada X, dan misalkan bahwa $\nabla\phi(X) \neq 0$, maka solusi optimal masalah meminimalkan $\phi'(X,\delta)$, dengan $\|\delta\|<1$ diberikan oleh $d^* = - \dfrac{\nabla \phi(X)}{\| \nabla \phi(X) \|}$

merupakan arah penurunan tercuram dari ϕ pada X.

Bukti :

$$\phi'(X;\delta) = \lim_{t \to 0^+} \frac{F(X^* + t\delta) - F(X^*)}{t} = \nabla F(X^*)^T \delta$$

Masalah dinyatakan mereduksi untuk meminimalkan $\nabla F(X^*)^T\delta$ dengan kendala $\| \delta \|_2 < 1$. Dengan pertidaksamaan Schwartz,

$\nabla\phi(X)^T\delta \geq \|\phi(X)^T\| \, \| \delta \| \geq \| \phi(X) \|$

Sehingga persamaan yang meliputi seluruhnya jhj

$$\delta^* = - \frac{\nabla\phi(X)}{\| \nabla\phi(X) \|}$$

Dan ini merupakan arah penurunan tercuram. ∎

Algoritma optimal Steepest Descent menghasilkan barisan $\{ X_k \}_{k=1}^{\infty}$ diberikan oleh

$$X_{k+1} = X_k + \delta_k$$
$$= X_k - t_k \nabla\phi(X_k) \qquad 47$$

dimana t_k adalah $t > 0$ terkecil sedemikian sehingga didapat t_k^* adalah suatu panjang langkah optimal dalam arah $\delta = -\nabla\phi$ yang meminimalkan $\phi(X_k^* + t\delta_k)$, yaitu

$$\phi(X_k^* + t_k\delta_k) = \min_{t>0} \phi(X_k^* + t\delta_k).$$

Metode Steepest Descent mengerjakan aproksimasi dengan sukses dan menghendaki harga fungsi obektif dan gradiennya pada setiap iterasi. Namun, dalam pencapaian konvergensi titik optimum memakan waktu yang cukup lama. Program iterasi ini berakhir jika dan ketika selisih antara nilai-nilai fungsi obyektif pada dua buah vektor X yang berturutan lebih kecil daripada suatu toleransi yang diperkenankan. Vektor X yang dihitung terakhir menjadi arpoksimasi akhir bagi X^*.

2.5.4 METODE NEWTON

Dalam mengaproksimasi fungsi nonlinear sangat tinggi dengan derivatif orde dua kontinu, menggunakan aproksimasi deret Taylor :

$$\phi(X) \cong \varnothing(X) = \phi(X_k) + \nabla\phi(X_k)^T(X - X_k) + \frac{1}{2}(X-X_k)^T\nabla^2\phi(X_k)(X-X_k)$$

dimana $\nabla^2\phi(X_k)$ adalah matriks Hessian NxN. Selanjutnya dicari derivatif dari $\varnothing(X)$ terhadap $(X - X_k)$, diperoleh system linear :

$$\nabla\phi(X_k)^T + \nabla^2\phi(X_k)(X-X_k) = 0$$

Titik X_{k+1} dipilih sebagai titik yang mengaproksimasi minimal. System linear tersebut bisa mempunyai solusi nol, satu atau infinit.

Syarat perlu untuk aproksimasi kuadratis $\emptyset(X)$ minimal adalah :

$$\nabla \emptyset = \lim_{\delta \to 0} \frac{\phi(X_k+\delta) - \phi(X_k)}{\delta} = 0$$

$$= \lim_{\delta \to 0} (\nabla\phi(X_k) + \nabla^2\phi(X_k)\delta) = 0$$

Jika $\nabla^2\phi(X_k)$ adalah matriks nonsingular maka dapat dicari nilai δ dan nilainya adalah

$$\delta^* = -\frac{\nabla\phi(X_k)}{\nabla^2\phi(X_k)}$$

Asumsikan $\nabla\phi(X^*) = 0$ dan $\nabla^2\phi(X^*)$ adalah definit positif pada minimal lokal pada X^*. Titik - titik pengganti untuk mengaproksimasi X^* yang meminimal ϕ adalah

$$X_{k+1} = X_k + \delta$$

$$= X_k - \nabla^2\phi(X_k)^{-1} \nabla\phi(X_k)$$

Dalam menerapkan prosedur ini ditemukan beberapa kesulitan yaitu:

i. Elemen-elemen dari matriks $\nabla^2\phi(X_k)$ harus dihitung pada setiap perkiraan titik X_k.

ii. Jika $\nabla^2\phi(X_k)$ singular, invers natrik $\nabla^2\phi(X_k)$ tersebut tidak mungkin diperoleh.

iii. Tidak membedakan antara titik minimum, maksimum, maupun titik pelana yang berakibat dapat menghasilkan titik belok yang salah.

Usaha untuk memperbaharui teknis diatas dilakukan dengan pembandingan prosedur optimal persamaan (47) dan (48), kita lihat hal khusus yang memenuhi keduanya, yaitu:

$$X_{k+1} = X_k - t_k M_k \nabla^2\phi(X_k)$$

untuk matriks M_k yang cocok. Di daerah yang jauh dari lokal minimum, dapat dipilih t_k yang membuat algoritma Steepest Descent optimal, dan $M_k = I$. Di daerah yang dekat minimum lokal, diambil $t_k = 1$ dan $M_k = [\nabla^2\phi(X)]^{-1}$ untuk mencapai persamaan (51). Oleh Algoritma Marquardt kekonvergenen global untuk titik stasioner dapat dijamin dengan melihat harga eigen dari $\nabla^2\phi(X_k)$. Pada setiap titik X_k, dapat dihitung harga eigen dari $\nabla^2\phi(X_k)$ dan mengambil λ_k skalar nonnegatif terkecil sedemikian sehingga matriks $\lambda_k I + \nabla^2\phi(X_k)$ yang mempunyai harga eigen lebih besar atau sama dengan harga penyeleksi $\varepsilon > 0$. Ini mengakibatkan matriks $\lambda_k I + \nabla^2\phi(X_k)$ definit positif.

Langkah Marquardt diberikan oleh :

$$X_{k+1} = X_k - t_k[\lambda_k I + \nabla^2\phi(X_k)]^{-1}\nabla\phi(X_k)$$

dimana t_k dipilih untuk meminimalkan atas $t \geq 0$, yang memenuhi :

$$\phi\left\{ X_k - t_k[\lambda_k I + \nabla^2\phi(X_k)]^{-1}\nabla\phi(X_k) \right\}$$

dan ini lebih lanjut akan dibahas dalam tugas akhir ini.