BAB II

LANDASAN TEORI

2.1 VEKTOR DAN MATRIKS

2.1.1 VEKTOR

Definisi 2.1.1.1 :

Vektor adalah suatu besaran yang merupakan garis yang mempunyai arah.

Penyajian vektor biasanya digunakan kurung [ ] ini yang dimaksudkan untuk membedakan dengan kurung ( ) yang digunakan untuk koordinat titik ataupun kurung { } yang sering digunakan dalam teori himpunan. Namun beberapa buku mempergunakan kurung ( ) atau { } tersebut untuk suatu vektor. Jadi, $a[a_1, a_2, ..., a_n]$ menjadi $a(a_1, a_2, ..., a_n)$ atau $\{a_1, a_2, ..., a_n\}$ atau dapat juga ditulis :

$$a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

Definisi 2.1.1.2 :

Himpunan m buah vektor $\{u_1, u_2, ... u_m\}$ disebut bergantung linier (linearly dependent, tidak bebas linier) bila terdapat skalar-skalar $\beta_1, \beta_2, ..., \beta_m$ yang tidak semua nol sedemikian sehingga $\beta_1 u_1 + \beta_2 u_2 + \beta_3 u_3 + \ ... \ + \beta_m u_m = 0$.

Definisi 2.1.1.3 :

Sedangkan $\{u_1, u_2, \ldots u_m\}$ disebut bebas linier (linearly independent) apabila $\beta_1 u_1 + \beta_2 u_2 + \beta_3 u_3 + \ldots + \beta_m u_m = 0$, hanya terpenuhi oleh $\beta_1 = \beta_2 = \ldots \beta_m = 0$.

contoh :

Pandang ruang vektor R^3 dengan $a=[3,1,2]$, $b=[1,2,1]$, dan $c=[2,-1,1] \in R^3$. ketiga vektor tersebut adalah bergantung linier karena $\beta_1 a + \beta_2 b + \beta_3 c = 0$, dengan demikian diperoleh $\beta_1 [3,1,2] + \beta_2 [1,2,1] + \beta_3 [2,-1,1] = [0,0,0]$, ada β yang tidak sama dengan nol , yaitu misal $\beta_1 = 1$, $\beta_2 = -1$, $\beta_3 = -1$.

Pandang ruang vektor R^2 dengan $a=[2,3]$, $b=[1,3] \in R^2$ kedua vektor tersebut bebas linier karena $\beta_1 [2,3] + \beta_2 [1,3] = [0,0]$

atau $2\beta_1 + \beta_2 = 0$

$3\beta_1 + 3\beta_2 = 0$

diperoleh hanya $\beta_1 = \beta_2 = 0$

2.1.2 MATRIKS

Definisi 2.1.2.1 :

Matriks adalah sekelompok besaran skalar yang disusun dalam larik segiempat (menurut baris-baris dan kolom-kolom). Skalar-skalar yang dikandungnya dinamakan elemen matriks. Beberapa notasi dan istilah yang erat kaitannya dengan matriks, diberikan dibawah ini :

1. Suatu elemen pada baris i dan kolom j dari matriks

dinyatakan dengan a_{ij}. Bila matriks A terdiri dari m baris dan n kolom maka dapat ditulis sebagai berikut :

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & & a_{mn} \end{bmatrix}$$

2. Suatu matriks dengan baris m dan kolom n dinamakan matriks berorde mxn. Jika m=n, maka matriks tersebut berbentuk bujursangkar dan dikatakan berorde n.
3. Matriks dengan satu baris atau satu kolom saja dinamakan vektor, dan biasanya ditulis sebagai berikut :

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

Beberapa operasi aljabar matriks ialah :

a. Penjumlahan

Penjumlahan dua matriks A dan B yang berorde sama menghasilkan matriks yang berorde sama pula dan dapat ditulis sebagai berikut :

$$c_{ij} = a_{ij} + b_{ij}$$

Proses penjumlahan ini mempunyai sifat-sifat seperti pada penjumlahan skalar, yaitu :

1) $A + (B + C) = (A + B) + C$ (hukum assosiatif)

2) $A + B = B + A$ (hukum komutatif)

3) Jika $A + C = B + C$ maka $A = B$

b. Perkalian Skalar

Diberikan matriks $A = [a_{ij}]$ dan skalar k , maka perkalian antara k dengan A menghasilkan matriks baru dengan mengalikan masing- masing elemen matrik A dengan k, dan dapat ditulis sebagai kA atau Ak.

Jika $B = [b_{ij}] = kA = Ak$, maka $b_{ij} = k\, a_{ij}$.

c. Perkalian Matriks

Bila matriks $A = [a_{ij}]_{mxn}$ dan $B = [b_{ij}]_{qxp}$, maka hasil perkalian $A B = C$ hanya dapat didefinisikan bila dan hanya bila $n = q$, dan matriks C akan berorde $m \times p$ dengan :

$$c_{ij} = \sum_{k=1}^{n} a_{ik}\, b_{jk} \qquad i = 1,2,\ldots,m$$

$$j = 1,2,\ldots,p$$

Dalam hal ini hukum komutatif tidak berlaku karena secara umum tidak benar bahwa $A B = B A$, sedangkan hukum assosiatif dan distributif tetap berlaku disini

$$A(B\ C) = (A\ B)\ C = A\ B\ C$$

$$(A + B)\ C = A\ C + B\ C$$

contoh:

$$A = \begin{bmatrix} 3 & 2 & 1 \\ 1 & 2 & 1 \end{bmatrix} \text{ berukuran } (2x3)$$

$$B = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} \text{ berukuran } (3x1)$$

$$\text{maka } A\ B = \begin{bmatrix} 3 & 2 & 1 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$$

$$A\ B = \begin{bmatrix} 11 \\ 5 \end{bmatrix} \text{ berukuran } (2x1)$$

Definisi 2.1.2.2 :

Pandang suatu matriks $A = (a_{ij})$ berukuran (mxn), maka transpose dari A adalah matriks A^T berukuran (nxm) yang didapatkan dari A dengan menuliskan baris ke-i dari A = 1,2,...,m sebagai kolom ke-i dari A^T. Dengan perkataan lain $A^T = (a_{ji})$

contoh :

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

bila : baris I ditulis sebagai kolom I

baris II ditulis sebagai kolom II

baris III ditulis sebagai kolom III

maka matriks hasilnya disebut A^T

$$A^T = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$$

Definisi 2.1.2.3 :

Matriks simetris adalah matriks yang transposenya sama dengan dirinya sendiri, dengan perkataan lain bila $A = A^T$ atau $a_{ij} = a_{ji}$ untuk semua i dan j. Jelas bahwa matriks simetris adalah bujursangkar.

2.1.3 DETERMINAN

Setiap matriks bujusangkar A selalu dikaitkan dengan suatu skalar yang disebut determinan matriks tersebut, dan ditulis sebagai det (A) atau | A |. Salah satu cara untuk mencari determinan adalah dengan minor dan kofaktor.

Pandang matriks dengan ukuran (nxn) yaitu $A = (A_{ij})$, dan M_{ij} suatu sub matriks dari A dengan ukuran (n-1)x(n-1) dimana baris ke-i dan kolom ke-j dari A dihilangkan.

contoh :

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

$$\text{maka } M_{32} = \begin{bmatrix} 1 & 3 \\ 4 & 6 \end{bmatrix}$$

(baris ke-3 dan kolom ke-2 dihilangkan)

<u>Definisi 2.1.3.1 :</u>

Minor dari elemen a_{ij} suatu matriks A (a_{ij}) adalah $|M_{ij}|$ dan kofaktor dari a_{ij} adalah $(-1)^{i+j}$ $|M_{ij}|$ adalah suatu skalar.

contoh :

$$A = \begin{bmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \\ 8 & 9 & 1 \end{bmatrix}$$

$$\text{Minor dari elemen } a_{32} = \begin{vmatrix} 2 & 4 \\ 5 & 7 \end{vmatrix} = 14 - 20 = -6$$

$$\text{Kofaktor dari elemen } a_{32} = (-1)^{3+2} \cdot (-6) = 6$$

<u>Definisi 2.1.3.2 :</u>

Determinan dari suatu matriks = jumlah perkalian elemen-elemen dari sebarang baris/kolom dengan kofaktor-kofaktornya.

Dengan perkataan lain :

$$| A | = \sum_{j=1}^{n} a_{ij} A_{ij} = a_{i1} A_{i1} + a_{i2} A_{i2} + \ldots + a_{in} A_{in}$$

dengan i sebarang, disebut uraian baris ke-i

$$| A | = \sum_{j=1}^{n} a_{ij} A_{ij} = a_{1j} A_{1j} + a_{2j} A_{2j} + \ldots + a_{nj} A_{nj}$$

dengan j sebarang, disebut uraian kolom ke-j

Kalau elemen-elemen dari suatu baris/kolom dikalikan dengan kofaktor-kofaktor dari elemen-elemen baris/kolom lain, jumlahnya akan nol.

Misalnya $a_{11} A_{21} + a_{12} A_{22} + \ldots + a_{1n} A_{2n} = 0$

atau $a_{11} A_{21} + a_{21} A_{22} + \ldots + a_{n1} A_{n2} = 0$

Jadi : $\sum_{j=1}^{n} a_{ij} A_{kj} = | A |$ bila $i = k$

$= 0$ bila $i \neq k$

$\sum_{j=1}^{n} a_{ij} A_{ik} = | A |$ bila $j = k$

$= 0$ bila $j \neq k$

2.1.4 BEBERAPA MATRIKS KHUSUS

Matriks diagonal

Matriks diagonal adalah matriks bujursangkar yang semua elemen diluar diagonal utama adalah nol. Dengan perkataan lain (a_{ij}) adalah matriks diagonal bila $a_{ij} = 0$ untuk $i \neq j$.

contoh :

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Matriks identitas

Matriks identitas adalah matriks diagonal yang elemen-elemen diagonal utamanya semua = 1. dengan perkataan lain (a_{ij}) adalah matriks identitas bila $a_{ij}=1$, untuk $i = j$, dan $a_{ij}=0$ untuk $i \neq j$. Matriks identitas biasa ditulis I atau I_n, dimana n menunjukkan ukuran matriks bujursangkar tersebut.

contoh :

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

<u>Definisi 2.1.4.1 :</u>

Suatu matriks berbentuk uv^T, dimana u dan v merupakan vektor, dikatakan sebagai matriks rank satu.

Setiap kolom dari uv^T merupakan satu pergandaan dari matriks u, dan setiap barisnya merupakan pergandaan dari vektor v^T. Pemilihan vektor u dan v tertentu memberikan matriks rank satu yang khusus.

contoh :

Jika $v = e_i$ (kolom ke-i dari matriks identitas), maka matriks uv^T adalah nol kecuali untuk kolom ke-i, yang merupakan vektor u.

<u>Definisi 2.1.4.2 :</u>

Suatu matriks berbentuk $uv^T + qr^T$, dimana u,v dan q,r merupakan vektor, dikatakan sebagai matriks rank dua.

<u>Definisi 2.1.4.3 :</u>

Kalau A dan B matriks-matriks bujursangkar berordo n dan berlaku $A\,B = B\,A = I$ maka dikatakan B invers dari A dan ditulis $B = A^{-1}$, sebaliknya A adalah invers dari B, dan ditulis $A = B^{-1}$.

2.1.5 AKAR DAN VEKTOR KARAKTERISTIK (EIGENVALUE DAN EIGENVEKTOR)

Definisi 2.1.5.1 :

A suatu matriks bujursangkar dan γ adalah skalar yang memenuhi persamaan :

$$Au = \gamma u$$

untuk suatu vektor kolom $u \neq 0$, maka dikatakan γ adalah suatu akar karakteristik dari A dan u yang memenuhi persamaan diatas disebut vektor karakteristik yang bersangkutan dengan γ.

contoh :

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

γ skalar dan $u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ vektor yang memenuhi maka

persamaan diatas dapat ditulis :

$(A - \gamma I)\, u = 0$

jadi

$$\begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} - \begin{bmatrix} \gamma & 0 \\ 0 & \gamma \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 1-\gamma & 2 \\ 3 & 2-\gamma \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\det : \begin{vmatrix} 1-\gamma & 2 \\ 3 & 2-\gamma \end{vmatrix} = 0$$

atau $(1-\gamma)(2-\gamma) - 6 = 0$

jadi $\gamma^2 - 3\gamma - 4 = 0$, akan mempunyai akar $\gamma_1 = 4$ dan $\gamma_2 = -1$

Definisi 2.1.5.2 :

Jika semua eigenvalue dari suatu matriks simetris A adalah definit positif, maka matriks simetris itu dikatakan definit positif. Jika A definit positif, maka untuk vektor x yang tidak nol berlaku :

$$x^T A x \geq 0 \qquad \text{dimana } x \neq 0$$

Definisi 2.1.5.3 :

Jika semua eigenvalue dari suatu matriks simetris A adalah non negatif, maka matriks A tersebut dikatakan semi definite positif.

Definisi 2.1.5.4 :

Jika suatu matriks simetris A mempunyai eigenvalue yang positif dan negatif, maka A dikatakan indefinite.

2.2 NORM

Suatu vektor norm yang dinotasikan dengan $|| \quad ||$ merupakan pemetaan dari R^n ke R^1 yang memberlakukan :

a. $||x|| \geq 0$, $\forall\ x \in R^n$ dan $||x|| = 0$ hanya jika x sama dengan nol.

b. $||\gamma x|| = |\gamma|\ ||x||$, $\forall\ x \in R^n$, $\gamma \in R^1$

c. $||x + y|| \leq ||x|| + ||y||$, $\forall\ x , y \in R^n$

Norm-p dari suatu vektor x dinotasikan dengan $||x||_p$, dan didefinisikan sebagai :

$$||x||_p = \left[\sum_{i=1}^{n} |x_i|^p \right]^{1/p} \qquad (2.2.1)$$

untuk p =2,dari persamaan (2.2.1) diperoleh:

$$||x||_2 = \sqrt{x^T x}$$
$$= \sqrt{|x_1|^2 + |x_2|^2 + \ldots\ldots + |x_n|^2}$$

2.3 Pengertian Optimasi Tanpa Kendala

Dalam bidang keteknikan maupun dalam dunia bisnis, banyak ditemukan model-model persamaan yang bila dinyatakan dalam fungsi atau hubungan matematik adalah tidak linier. Hubungan yang tidak linier ini lebih mewakili atau mencerminkan abstraksi masalah-masalah yang sebenarnya dari pada hubungan matematik linier.

Permasalahan optimasi erat kaitannya dengan usaha untuk mendapatkan nilai maksimum atau minimum dari suatu fungsi obyektif dengan n variabel. Suatu fungsi atau hubungan matematik dikatakan tanpa kendala jika atas fungsi tersebut tidak dibatasi dengan kendala.
Karena fungsi ini tidak dibatasi dengan kendala, maka fungsi berdiri sendiri atau dapat berarti fungsi minimisasi atau maksimasi. Secara umum fungsi diatas dapat ditulis dalam bentuk

$$Z = f(X)$$

$$\text{dengan } X \in R^n$$

dimana f(X) merupakan fungsi non linier dengan n variabel bebas dan tanpa kendala. Dengan perkataan lain untuk menyelesaikan fungsi diatas merupakan suatu permasalahan optimasi tanpa kendala.

2.3.1 Optimasi Tanpa Kendala 1 Variabel Dengan Interpolasi Kubik

2.3.1.1 Lokal dan global minimum

Misalkan $D \subseteq R^1$ domain fungsi $f(x)$.

Definisi 2.3.1.1 :

Fungsi $f(x)$ dikatakan mempunyai lokal minimum pada $x^* \in D$ jika ada selang buka $J = (x^*-\delta , x^*+\delta)$, dengan $\delta > 0$ sedemikian sehingga $f(x) \geq f(x^*)$, untuk setiap $x \in J \cap D$.

Definisi 2.3.1.2 :

Fungsi $f(x)$ dikatakan mempunyai global minimum pada $x^* \in D$ jika ada selang buka $J = (x^*-\delta , x^*+\delta)$, dengan $\delta > 0$ sedemikian sehingga $f(x) \geq f(x^*)$. untuk setiap $x \in D$

Bentuk umum dari minimisasi tanpa kendala satu variabel adalah

$$\text{meminimalkan } f(x), \text{ dimana } x \in R^1$$

jika $f(x)$ dapat dideferensialkan secara kontinue dua kali dan mempunyai lokal minimum di titik x^*, maka ada syarat perlu dan syarat cukup yang harus dipenuhi titik x^*.

Syarat perlu untuk minimum tanpa kendala 1 variabel:

1. $f'(x^*) = 0$ dan
2. $f''(x^*) \geq 0$

Syarat pertama dibuktikan dengan kontradiksi. Dimana akan ditunjukkan bahwa jika $f'(x^*)$ tidak nol, maka setiap persekitaran (neighborhood) dari x^* berisi titik-titik dengan nilai fungsi yang lebih rendah dari pada $f(x^*)$.

Karena tidak ada kendala, semua titik-titik feasible dan dititik beratkan pada nilai f di sekitar titik-titik persekitaran.

Kita ekspansikan f dalam deret Taylor disekitar titik x^*:

$$f(x^*+\varepsilon) = f(x^*) + \varepsilon f'(x^*) + \frac{1}{2}\varepsilon^2 f''(x^*+\theta\varepsilon) \qquad (2.3.1)$$

untuk $\theta (0 \leq \theta \leq 1)$.

Andaikan bahwa x^* adalah suatu lokal minimum, dan ambil $f'(x^*) < 0$, maka harus ada $\bar{\varepsilon}$ ($\bar{\varepsilon} > 0$) sedemikian sehingga $\varepsilon f'(x^*)+\frac{1}{2}\varepsilon^2 f''(x^*+\theta\varepsilon) < 0$ untuk semua $0 < \varepsilon \leq \bar{\varepsilon}$. Maka persamaan diatas menjadi $f(x^*+\varepsilon) < f(x^*)$ untuk setiap ε. Jadi setiap persekitaran dari x^* berisi titik-titik dengan nilai fungsi yang lebih rendah. Hal ini mengakibatkan kontradiksi dengan asumsi optimalitas.

Sekarang andaikan $f'(x^*) > 0$ maka harus ada $\bar{\varepsilon}$ ($\bar{\varepsilon} > 0$) sedemikian sehingga $\varepsilon f'(x^*)+\frac{1}{2}\varepsilon^2 f''(x^*+\theta\varepsilon) < 0$ untuk semua $0 < \varepsilon \leq \bar{\varepsilon}$. Dengan demikian menyebabkan $f(x^*+\varepsilon) < f(x^*)$. Hal ini juga kontradiksi dengan asumsi optimalitas.

Oleh karena itu $f'(x^*)$ harus nol agar x^* menjadi minimum .

Syarat 2 juga dibuktikan dengan kontradiksi.

Asumsikan bahwa x^* merupakan suatu lokal minimum. Karena $f'(x^*) = 0$ maka persamaan (2.3.1) menjadi :

$f(x^*+\varepsilon) = f(x^*)+\frac{1}{2}\varepsilon^2 f''(x^*+\theta\varepsilon)$ untuk θ dimana $0 \leq \theta \leq 1$.

Jika $f''(x^*)$ negatif, maka f'' akan tetap negatif dalam persekitaran x^*. Jika $|\varepsilon|$ cukup kecil maka $x^*+\varepsilon$ berada dalam persekitaran itu, sehingga persamaan diatas menjadi $f(x^*+\varepsilon) < f(x^*)$.

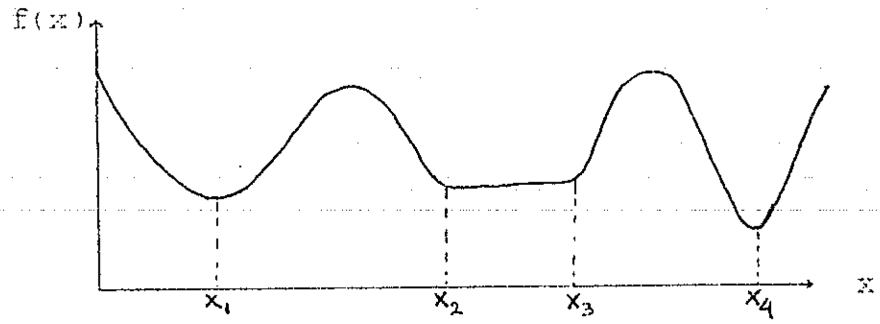
Ini mengakibatkan x^* tidak dapat menjadi lokal minimum apabila $f''(x^*) < 0$.

Berikut ini ada dua syarat cukup untuk menjamin bahwa titik x^* adalah suatu lokal minimum yang kuat.

1. $f'(x^*) = 0$
2. $f''(x^*) > 0$

Untuk syarat yang pertama, telah dibuktikan dalam syarat perlu yang pertama. Sekarang tinggal menunjukkan syarat kedua. Jika syarat kedua berlaku yaitu $f''(x^*) > 0$, maka secara kontinue $f''(x^*+\varepsilon)$ akan menjadi positif untuk setiap $|\varepsilon|$ yang cukup kecil. Karena $|\varepsilon|$ cukup kecil maka persamaan $f(x^* + \varepsilon) = f(x^*) + \frac{1}{2}\varepsilon^2 f''(x^* + \theta\varepsilon)$ untuk θ dimana $0 \le \theta \le 1$, menjadi $f(x^*+\varepsilon) > f(x^*)$. Hal ini menandakan bahwa $f(x^*)$ nilainya lebih kecil dari pada f pada titik lain dalam persekitaran dan dengan demikian x pasti suatu lokal minimum yang kuat.

Dibawah ini contoh grafik minima dalam masalah satu variabel.



keterangan:

x_1 = suatu lokal minimum kuat

x_2 , x_3 = suatu lokal minimum lemah

x_4 = suatu global minimum

Untuk mendapatkan nilai minimum fungsi dengan satu variabel dapat digunakan suatu metoda interpolasi kubik. Dalam hal ini mencari suatu nilai λ^*, yaitu nilai terkecil tidak negatif λ ($\lambda \geq 0$) sehingga fungsi

$$f(\lambda) = f(X + \lambda p) \qquad (2.3.2)$$

mencapai minimum.

Dengan λ = panjang langkah optimal dalam arah p

p = vektor arah

Jika fungsi asal f(X) dinyatakan secara eksplisit dari x_i, i = 1,2,3,...,n, maka untuk setiap vektor khusus p, dengan mudah dapat ditulis pernyataan persamaan (2.3.2), kemudian menyelesaikan

$$f'(\lambda) = \frac{df}{d\lambda}(\lambda) = 0 \qquad (2.3.3)$$

sehingga diperoleh nilai λ^* dalam bentuk vektor X dan p.

Pada masalah praktis, persamaan (2.3.2) pada umumnya tidak dapat dinyatakan dalam bentuk λ secara eksplisit. Sehingga akan terasa sulit memperoleh nilai λ^*. Untuk mengatasi hal ini, nilai λ^* dapat dicari menggunakan metoda interpolasi kubik.

2.3.1.2 Metoda Interpolasi Kubik

Dalam metoda interpolasi kubik, panjang langkah optimal λ^* dicari dalam empat tahap. Teknik yang dipakai adalah menggunakan turunan fungsi $f(\lambda)$, yaitu

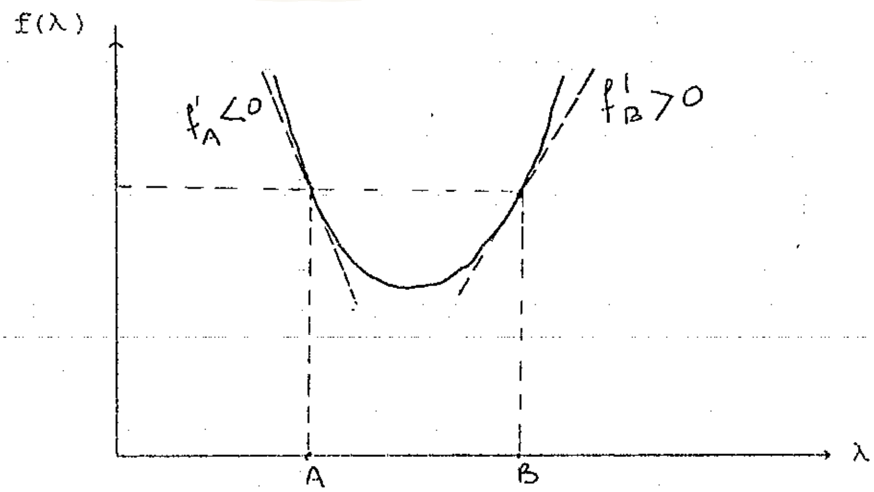
$$f'(\lambda) = f'(x+\lambda p).$$

Tahap pertama

Dalam tahap ini vektor p dinormalkan (jika masalah peminimuman fungsi banyak variabel) yaitu $p_{nor} = \dfrac{p}{||\ p\ ||}$ dengan $||\ p\ || = \sqrt{p_1^2 + p_2^2 + \ldots + p_n^n}$

Tahap kedua

Dibuat batas bawah dan atas dari panjang langkah optimal λ^*, kemudian dicari dua titik A dan B pada kemiringan $f'(\lambda)$ sehingga tanda dari $f'(\lambda)$ pada titik A berbeda dengan tanda dari $f'(\lambda)$ pada titik B. Jika diandaikan pada $\lambda = A$, $f'(\lambda=A) < 0$ maka tinggal dicari titik lain $\lambda = B$ yang mempunyai kemiringan $f'(\lambda=B) > 0$. Untuk titik B dapat diambil salah satu dari $B = A + i\gamma$, $i=1,2,3,\ldots.$ dengan $\gamma>0$ sehingga $f'(\lambda=B) > 0$. Kemudian batasi panjang langkah optimal λ^* dalam selang $A < \lambda^* \leq B$. proses ini ditunjukkan dalam gambar 2



gambar 2 : minimum $f(\lambda)$ bergerak antara A dan B

Tahap ketiga

Misalkan polinom kubik yang digunakan sebagai pendekatan fungsi f(λ) antara titik A dan titik B adalah

$$h(\lambda) = a + b\lambda + c\lambda^2 + d\lambda^3$$

Atau secara matematis ditulis

$$f(\lambda) \approx h(\lambda) = a + b\lambda + c\lambda^2 + d\lambda^3 \qquad (2.3.1.2.1)$$

dengan a,b,c,dan d adalah konstanta yang akan dicari.

Untuk mencari konstanta-konstanta tersebut, harus dicari nilai $f_A = f(\lambda=A)$, $f'_A = f'(\lambda=A)$, $f_B = f(\lambda=B)$, $f'_B = f'(\lambda=B)$. Asumsikan $A \neq 0$, maka rumus umum pendekatan $\tilde{\lambda}^*$ dapat diturunkan sebagai berikut.

Dari persamaan (2.3.1.2.1) diperoleh:

$$\left.\begin{aligned} f_A &= a + bA + cA^2 + dA^3 \\ f_B &= a + bB + cB^2 + dB^3 \\ f'_A &= b + 2cA + 3dA^2 \\ f'_B &= b + 2cB + 3dB^2 \end{aligned}\right] \qquad (2.3.1.2.2)$$

Jika persamaan (2.3.1.2.2.) diselesaikan didapat:

$$a = f_A - bA - cA^2 - dA^3 \qquad (2.3.1.2.3)$$

$$b = \frac{1}{(A-B)^2}(B^2 f'_A + A^2 f'_B + 2ABZ) \qquad (2.3.1.2.4)$$

$$c = -\frac{1}{(A-B)^2}((A+B)Z + Bf'_A + Af'_B) \qquad (2.3.1.2.5)$$

$$d = \frac{1}{3(A-B)^2}(2Z + f'_A + f'_B) \qquad (2.3.1.2.6)$$

$$\text{dengan } Z = \frac{3(f_A + f_B)}{B - A} + f'_A + f'_B \qquad (2.3.1.2.7)$$

Misalkan $\tilde{\lambda}^*$ titik pendekatan minimum f(λ), maka

$$f'(\tilde{\lambda}^*) = \frac{dh}{d\lambda}(\lambda^*) = b + 2c\lambda + 3d\lambda^2 = 0$$

$$\tilde{\lambda}^* = \frac{-2c \pm \sqrt{(4c^2 - 12bd)}}{6d}$$

$$\tilde{\lambda}^* = \frac{-c \pm \sqrt{(c^2 - 3bd)}}{3d} \qquad (2.3.1.2.8)$$

Dengan mensubstitusikan konstanta-konstanta a,b,c dan d yang diberikan persamaan (2.3.1.2.3) sampai (2.3.1.2.7) dalam persamaan (2.3.1.2.8) diperoleh :

$$\tilde{\lambda}^* = A + \frac{(f'_A + Z \pm Q)}{(f'_A + f'_B + 2Z)}(B - A) \qquad (2.3.1.2.9)$$

dengan $Q = \sqrt{Z^2 - f'_A\, f'_B}$ (2.3.1.2.10)

Asumsikan A = 0, maka persamaan (2.3.1.2.3) sampai dengan (2.3.1.2.10) menjadi :

$$a = f_A$$

$$b = \frac{1}{B^2} B^2 f'_A$$

$$= f'_A$$

$$c = -\frac{1}{B^2}(BZ + Bf'_A)$$

$$= -\frac{1}{B}(Z + f'_A)$$

$$d = \frac{1}{B^2}(2Z + f'_A + f'_B)$$

$$\text{dan } \tilde{\lambda}^* = \frac{\frac{1}{B}(Z + f'_A) \pm \frac{1}{B}\sqrt{(Z + f'_A)^2 - (2Z + f'_A + f'_B)f'_A}}{\frac{1}{B^2}(2Z + f'_A + f'_B)}$$

$$\tilde{\lambda}^* = B\,\frac{f'_A + Z \pm Q}{f'_A + f'_B + 2Z} \qquad (2.3.1.2.11)$$

dengan $Q = \sqrt{Z^2 - f'_A\, f'_B}$ (2.3.1.2.12)

$$Z = \frac{3(f_A - f_B)}{B} + f'_A + f'_B \qquad (2.3.1.2.13)$$

Dua nilai $\tilde{\lambda}^*$ dalam persamaan (2.3.1.2.9) dan persamaan (2.3.1.2.11) memberikan dua kemungkinan untuk $h'(\lambda)$ yaitu nilai maksimum dan nilai minimum dari $h(\lambda)$. Untuk menghindari harga imaginair Q, maka persamaan (2.3.1.2.12) harus diberikan syarat supaya $Z^2 - f'_A f'_B \geq 0$. Pertidaksamaan ini secara otomatis dipenuhi,sebab dijamin oleh asumsi $f'_A < 0$ dan $f'_B > 0$. Selanjutnya menghitung $\tilde{\lambda}^*$ dengan menggunakan persamaan (2.3.1.2.9), kemudian dilanjutkan ke tahap keempat.

Tahap keempat

Nilai pendekatan $\tilde{\lambda}^*$ yang diperoleh dalam tahap ini adalah pendekatan minimum fungsi $h(\lambda)$, dan nilai disini belum tentu nilai minimum fungsi $f(\lambda)$. Supaya nilai pendekatan λ^* minimum $f(\lambda)$, kriteria berikut dapat dipakai untuk menguji nilai pendekatan $\tilde{\lambda}^*$.

$$|f'(\tilde{\lambda}^*)| \leq a \qquad (2.3.1.2.14)$$

dengan a adalah bilangan positif cukup kecil yang diambil secara khusus tergantung nilai ketelitian yang diinginkan. Jika kriteria (2.3.1.2.14) tidak dipenuhi, maka persamaan kubik baru $\tilde{h}(\lambda) = \tilde{a} + \tilde{b}\lambda + \tilde{c}\lambda^2 + \tilde{d}\lambda^3$ dapat digunakan sebagai pendekatan baru fungsi $f(\lambda)$.

Untuk mencari konstanta-konstanta $\tilde{a}, \tilde{b}, \tilde{c}$, dan $\tilde{d}$, dapat digunakan nilai fungsi $f(\lambda)$ dan turunannya pada 2 titik terbaik, yang dipilih dari tiga titik A,B dan $\tilde{\lambda}^*$.

Kemudian rumus umum yang diberikan oleh persamaan (2.3.1.2.9) digunakan untuk mencari $\tilde{\lambda}^*$ yang baru. Selanjutnya nilai baru $\tilde{\lambda}^*$ diuji keoptimalannya. Jika $|f'(\tilde{\lambda}^*)| \leq a$, maka $\lambda^* = \tilde{\lambda}^*$ dan proses iterasi berhenti. Selanjutnya jika $|f'(\tilde{\lambda}^*)| > a$, uji tanda fungsi $f'(\lambda)$ pada pendekatan titik baru $\tilde{\lambda}^*$. Jika $f'(\tilde{\lambda}^*) < 0$, maka ambil $A = \tilde{\lambda}^*$ dan B tetap, sebaliknya jika $f'(\tilde{\lambda}^*) > 0$, ambil A tetap dan $B = \tilde{\lambda}^*$. Ulangi proses iterasi sampai diperoleh nilai $\lambda^* = \tilde{\lambda}^*$.

2.3.2 Optimasi Tanpa Kendala Beberapa Variabel

Misalkan $D \subseteq R^n$ domain dari $f(X)$

Definisi 2.3.2.1 :

Fungsi $f(X)$ dikatakan mempunyai lokal minimum atau relatif pada X^*, jika ada bola buka $\mathbb{B}[X^*,\delta]$, $\delta > 0$ sedemikian sehingga $f(X) \geq f(X^*)$, $\forall \in \mathbb{B}[X^*,\delta] \cap D$.

Definisi 2.3.2.2 :

Fungsi $f(X)$ dikatakan mempunyai global minimum atau absolut pada X^*, jika $f(X\ ) \geq f(X^*)$, $\forall X \in D$.

catatan: Pengertian bola buka $\mathbb{B}[X^*,\delta]$ didefinisikan oleh $\mathbb{B}[X^*,\delta] = \{\ X \in R^n |\ \ ||X - X^*|| < \delta\ \}$, yaitu himpunan titik didalam bola berdimensi n dengan pusat X^* dan jari-jari $\delta > 0$.

Definisi 2.3.2.3 :

Misalkan fungsi $f(X)$ dengan domain $D \subseteq R^n$ terdeferinsial pada $X^* \in D$, maka vektor

$$\nabla f(X^*) = \begin{bmatrix} f_{x_1}(X^*) \\ \vdots \\ \vdots \\ f_{x_n}(X^*) \end{bmatrix}, \text{ dengan } f_{x_i} = \frac{\partial f}{\partial x_i}(X^*)$$

$$i = 1,2,3,\ldots n$$

disebut **gradien** dari $f(X)$ pada X^*

Definisi 2.3.2.4 :

Jika fungsi $f(X)$ mempunyai turunan parsial orde kedua yang kontinue pada X^*, maka

$$H_f(X^*) = \left[\frac{\partial^2 f(X^*)}{\partial X_i \, \partial X_j} \right], \; i,j = 1,2,3,\ldots,n$$

disebut **matriks Hessian** $f(X)$ pada X^*.

Pengertian mengenai optimasi tanpa kendala telah disinggung diatas dalam awal bab 2. Dan sekarang akan dibahas mengenai persoalan minimisasi tanpa kendala dalam n dimensi atau bisa ditulis dalam bentuk:

$$\text{meminimalkan } F(X)$$

$$\text{dimana } X \in R^n$$

Seperti dalam masalah 1 variabel yang membicarakan syarat-syarat perlu dan cukup untuk X^* menjadi suatu lokal minimum dari permasalahan minimisasi tanpa kendala.

Berikut syarat-syarat perlu untuk minimum dengan n variabel:

1. $\| g(X^*) \| = 0$, yaitu X^* suatu titik stasioner.
2. $G(X^*)$ adalah positif semi definite.

Seperti dalam satu variabel, syarat optimal berasal dari ekspansi deret Taylor pada F disekitar X^*.

$$F(X^*+\varepsilon p)=F(X^*)+\varepsilon p^T g(X^*)+\frac{1}{2}\varepsilon^2 p^T G(X^*+\varepsilon\theta p)p \qquad (2.3.2.1)$$

dimana θ memenuhi $0\le\theta\le 1$, ε adalah suatu skalar dan p adalah suatu vektor n.

Syarat pertama dibuktikan dengan kontradiksi.

Diasumsikan bahwa X^* adalah suatu lokal minimum, tetapi bukan titik stasioner. Jika $g(X^*)$ bukan nol maka ada suatu vektor p sedemikian, maka

$$p^T g(X^*)\le 0 \qquad (2.3.2.2)$$

untuk contoh, p dapat diambil sebagai $-g(X^*)$. Maka dapat diartikan sebarang vektor p yang memenuhi (2.3.2.2) adalah arah menurun pada X^*.

Diberikan sebarang arah menurun p, maka ada suatu skalar positif $\bar{\varepsilon}$ sedemikian hingga untuk semua ε positif yang memenuhi $\varepsilon \le \bar{\varepsilon}$ berlaku:

$$\varepsilon p^T g(X^*)+\frac{1}{2}\varepsilon^2 p^T G(X^*+\varepsilon\theta p)p<0$$

sehingga dari persamaan (2.3.2.1) berlaku $F(X^*+\varepsilon p)<F(X^*)$ untuk semua ε. Dengan demikian $g(X^*)$ adalah nol, disetiap persekitaran dari X^* yang berisi titik-titik dengan nilai fungsi yang lebih rendah daripada $F(X^*)$. Ini membuktikan bahwa setiap lokal minimum harus menjadi titik stasioner.

Syarat 2 juga dibuktikan dengan kontradiksi.

Dengan memperhatikan persamaan (2.3.2.1) diatas dan syarat 1 untuk persoalan minimisasi tanpa kendala beberapa variabel maka: $F(X^*+\varepsilon p)=F(X^*)+\frac{1}{2}\varepsilon^2 p^T G(X^*+\varepsilon\theta p)p$ (2.3.2.3)

Apabila $G(X^*)$ indefinite, secara kontinue G akan menjadi indefinite untuk semua titik-titik dalam persekitaran X^*, dan kita dapat memilih $|\varepsilon|$ yang cukup kecil, maka $X^*+\varepsilon p$ berada dalam persekitaran itu.

Dengan definisi dari suatu matriks indefinite, p dapat dipilih sedemikian maka $p^T G(X^*+\varepsilon\theta p)p<0$. Oleh karena itu dari persamaan (2.3.2.3) diatas, dapat diambil kesimpulan bahwa setiap persekitaran dari X^* berisi titik-titik dimana nilai dari F lebih rendah daripada di X^*, yang mana terjadi kontradiksi dengan syarat optimal pada X^*.

Syarat-syarat cukup untuk minimum pada optimasi tanpa kendala beberapa variabel:

1. $\|g(X^*)\| = 0$
2. $G(X^*)$ adalah positif definite.

Untuk bukti yang pertama, telah ditunjukkan seperti syarat perlu diatas. Sekarang tinggal menunjukkan syarat cukup yang kedua.

Kita ekspansikan F disekitar X^* menurut deret Taylor.

$$F(X^*+\varepsilon p)=F(X^*)+\varepsilon p^T g(X^*)+\frac{1}{2}\varepsilon^2 p^T G(X^*+\varepsilon\theta p)p$$

Jika $G(X^*)$ positif definite, maka secara kontinue G adalah positif definite untuk semua titik-titik dalam persekitaran dari X^*. Jika diberikan $|\varepsilon|$ cukup kecil, maka $X^*+\varepsilon p$ akan berada dalam persekitaran itu. Dengan demikian untuk semua ε dan setiap arah p berlaku bahwa $p^T G(X^*+\varepsilon\theta p)p>0$. Dari persamaan diatas, secara tak langsung $F(X^*)$ nilainya lebih kecil dari F untuk setiap titik di persekitaran X^* dan dengan demikian x^* adalah suatu lokal minimum.