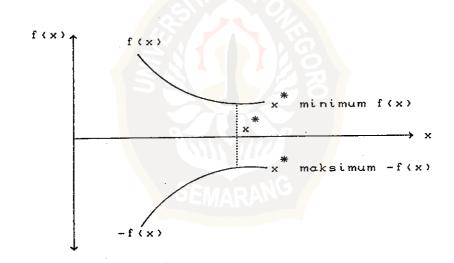
BAB IV OPTIMASI NON LINIER

4.1. Tinjauan Umum Optimasi

4.1.1. Definisi

Optimasi adalah proses efisiensi dalam mencari kondisi yang akan memaksimumkan atau meminimumkan suatu fungsi. Secara umum optimasi dapat didefinisikan sebagai proses meminimumkan suatu fungsi dengan dasar bahwa mencari nilai maksimum suatu fungsi sama dengan mencari nilai minimum dari negatif fungsi tersebut (RaO S.S., 1979), seperti yang diperlihatkan pada gambar 4.1.



Gambar 4.1. Nilai minimum suatu fungsi x* berada pada absis yang sama dengan nilai maksimum negatif fungsi tersebut (Rao S.S., 1979).

Proses optimasi dikenal juga sebagai program matematik yang merupakan bagian Operasi Riset (OR). Metode-metode yang digunakan dalam OR diperlihatkan pada bagan 4.1.

METODA OPERASI RISET

Program Matematik Proses Stochastic	Metode Statistik
1. Metoda Kalkulus 1. Proses Markov	1.Analisa
2. Kalkulus variasi 2. Teori Antrian	Regresi
3. Pemrograman 3. Teori Renewal	2.Disain
Non Linier 4. Metoda Simu <mark>la</mark> si	Percobaan
4. Pemrograman Linier 5. Teori Reliabilit	y 3.Analisa
5. Pemrograman Integer 6. Teoeri Keputusan	Cluster
o. Pemrograman Dinamik Statistik	4.Analisa
7. Pemrograman Quadratik	Diskriminasi
8. Pemrograman Sep <mark>a</mark> rabel	
9. Pemrograman Stochastik	
10. Pemrograman Multi obyektiv	
11. Pemrograman Geometrik	
12.Teori Permainan	
13. Metoda Jaringan ; CPM & PERT	

Bagan 4.1. Metode-metode dalam Operasi Riset (Rao S.S., 1979).

4.1.2. Perumusan Masalah Optimasi

Program matematik adalah menentukan

 $X = [x_1, x_2, \dots, x_n]^T \in \mathbb{R}^n$ (Rao S.S., 1978) yang meminimumkan atau memaksimimumkan f(X), terhadap konstrain (syarat pembatas)

$$g_{i}(X) \leq 0,$$
 $i = 1, 2, ..., m$ (4.1)

dan

$$e_{j}(X) = 0, \quad j = m+1, m+2, \dots, p \quad (4.2)$$

dengan

- * f(X) disebut fungsi obyektif
- * $g_{i}(X)$ disebut konstrain ketaksamaan
- * e (X) disebut konstrain kesamaan

Bila konstrain tidak ada (p = 0), maka masalah di atas sebagai masalah optimasi tanpa konstrain. Vektor $X \in \mathbb{R}^n$ yang mewakili persamaan (4.1) dan (4.2) disebut fisibel sedang yang tidak memenuhi persamaan tersebut di atas disebut infisibel. Solusi fisibel yang mengoptimasikan fungsi obyektif disebut solusi optimal.

4.2. Teori Optimasi Klasik

Metode ini menggunakan differensial kalkulus untuk menetukan titik optimum. Walaupun metode ini mempunyai keterbatasan karena hanya berlaku untuk fungsi obyektif yang dapat diturunkan, tetapi merupakan dasar untuk pengembangan teknik numerik dalam pengoptimasian (Rao S.S., 1978).

4.2.1. Optimasi Fungsi Satu Variabel

Teorema 4.1

Misal f(x) didefinisikan dalam interval $a \le x \le b$ dan mempunyai minimum relatif pada $x = x^*$. dimana $a \le x^* \le b$, dan apabila df(x)/dx = f'(x) ada pada $x = x^*$ maka $f'(x^*) = 0$ (Rao, S.S., 1978).

Teorema 4.2

Misal $f'(x^*) = f''(x^*) = ... = f^{(n-1)}(x^*) = 0$, dan $f^{(n)}(x^*) \neq 0$. Maka $f(x^*)$ adalah :

- * Nilai minimum f(x) apabila $f^{(n)}(x^*) > 0$ dan nigenap.
- * Nilai maksimum f(x) apabila $f^{(n)}(x^*) < 0$ dan n genap.
- * Bukan minimum atau maksimum bila n ganjil (Rao, S.S., 1978).

4.2.2. Optimasi F<mark>u</mark>ngsi Multi Variabel

Definisi differensial ke-r dari fungsi f(X), X & Rⁿ

Misal semua turunan parsial fungsi f(x) ada dan kontinu pada X* maka polinomial

$$d^{r}f(X^{*}) = \sum_{i=1}^{n} \sum_{j=1}^{n} \dots \sum_{k=1}^{n} h_{i}h_{j} \dots h_{k} \frac{\partial^{r}f(X^{*})}{\partial x_{i}\partial x_{j}\dots \partial_{k}}$$

disebut turunan ke-r dari fungsi f(X) pada X* (Rao, S.S., 1978).

Teorema 4.3

Apabila f(X) mempunyai titik ekstrim (maksimum

atau minimum) pada $X = X^*$ dan apabila turunan parsial pertama f(X) ada pada X^* , maka $\frac{\partial f(X^*)}{\partial x_1} = \frac{\partial f(X^*)}{\partial x_2} = \dots = \frac{\partial f(X^*)}{\partial x_n} = 0 \quad (Rao, S. S., 1978).$

Teorema 4.4

Syarat cukup untuk titik stasioner X^* menjadi titik ekstrim apabila matriks turunan parsial kedua f(X) (matrik Hessian-nya) pada X^*

- (i) Definit positip apabila X* titik minimum.
- (ii) Definit negatif apabila X* titik maksimum (Rao, S.S., 1978).

Teorema 4.5

Misal df(X) = 0, $d^2f(X) = 0$, ..., $d^{k-1}f(X) = 0$ dan $d^kf(X) \neq 0$ pada titik stasioner X^* , bila k genap maka f(X) menacapai

- * minimum di X* apabila d*f(X) definif positip
- * maksimum X* apabila dkf(X) definit negatif
- * tidak ada ketentuan apabila d^kf(X) semi definit positip atau semi definit negatif.

Bila k ganjil maka fungsi f(X) tidak mencapai maksimum ataupun minimum di titik X (Rao, S.S., 1978).

Teorema 4.6

Syarat cukup fungsi f(X) dengan konstrain $g_j(X)=0$, $j=1,2,\ldots,m$ mempunyai minimum relatif pada titik X^* , apabila turunan parsial fungsi pertama fungsi Lagrange didefinisikan oleh L=

 $L(x_1, x_2, \dots, x_n; \lambda_1, \lambda_2, \dots, \lambda_n)$ terhadap tiap argumennya adalah nol (*Rao, S.S., 1978*).

4.3. Program Non Linier

4.3.1. Persamaan Non Linier

Persamaan-persamaan linier aljabar mempunyai bentuk

$$a_{1}a_{x} + a_{12}x_{12} + \dots + a_{1n}x_{1n} = C_{1}$$

$$\vdots$$

$$a_{n1}a_{n1} + a_{n2}a_{12} + \dots + a_{nn}a_{nn} = C_{n}$$

dengan C konstsanta dan n adalah banyaknya persamaan. Persamaan aljabar yang tidak cocok dengan bentuk di atas didefinisikan sebagai persamaan non linier. Misalnya persamaan dalam bentuk transenden, kuadratik, polinom berderajat dua (Steven C.C dan Raymond P.C., 1989).

4.3.2. Minimisasi Jumlah Kuadrat Terkecil Parameter Non Linier

Jika model yang akan dicocokkan dengan data adalah

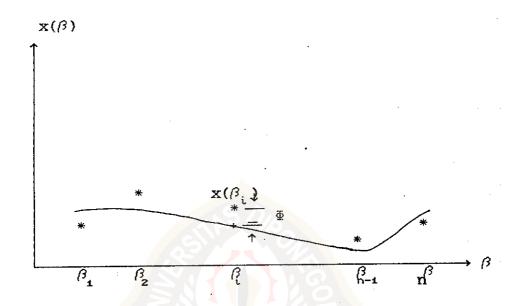
$$E(y) = f(x_1, x_2, ..., x_m; \beta_1, \beta_2, ..., \beta_k)$$

= $f(x,\beta)$ (4.3)

dengan

* x_1, x_2, \ldots, x_m adalah variabel bebas

- * β_1 , β_2 , ..., β_k adalah nilai populasi dari sejumlah k parameter.
- * E(y) adalah nilai yang diharapkan dari variabel tak bebas y



Gambar 4.2. Pencocokan kurva dimana x(β) yang diketahui oleh n pasangan data diperkirakan oleh fungsi non linier f(x,β) yaitu jumlah yang diberikam dari fungsi basis non linier. Permasalahan matematis adalah meminimumkan ralat Φ, pada tiap pasang sampel (Thomas R. Curhbert Jr., 1978)

Apabila titik-titik data dinyatakan dalam bentuk

$$(Y_i, X_{1i}, X_{2i}, \dots, X_{mi})$$
 $i = 1, 2, \dots, n$ (4.4)

Permasalahan yang dihadapi adalah menghitung estimasi dari parameter-parameter yang akan meminimumkan

$$\Phi = \sum_{i=1}^{n} [Y_i - \hat{Y}_i]^2 = \|Y - \hat{Y}\| \qquad (4.5)$$

dengan Ŷ adalah nilai y yang diperkirakan oleh persamaan (4.3) pada titik data ke-i.

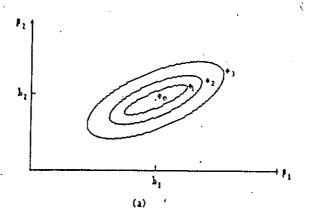
Persamaan (4.3) yang merupakan fungsi jumlah kuadrat residu dari model dan data selanjutnya dipakai sebagai fungsi obyektif. Secara grafik permasalahan matematik tersebut dapat dilihat pada gambar 4.2.

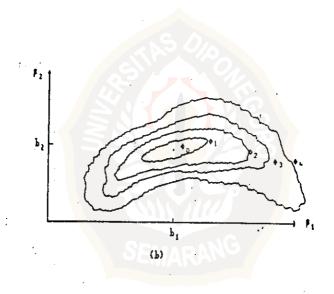
Tafsiran geometris dari kontur \$\Phi\$ diperlihatkan pada gambar 4.3. Apabila fungsi f mumpunyai hubungan linier dengan \$\beta\$, maka kontur \$\Phi\$ berbentuk elipsoid. Sedangkan jika f non linier konturnya terdistorsi menurut derajat ke-nonlinierannya. Tetapi di sekitar titik minimum kontur non linier mendekati elip.

4.3.2.1. Metoda Penurunan Tercuram (steepest descent)

Turunan parsial dari suatu fungsi f terhadap n variabel, secara umum dinamakan gradien fungsi ∇f

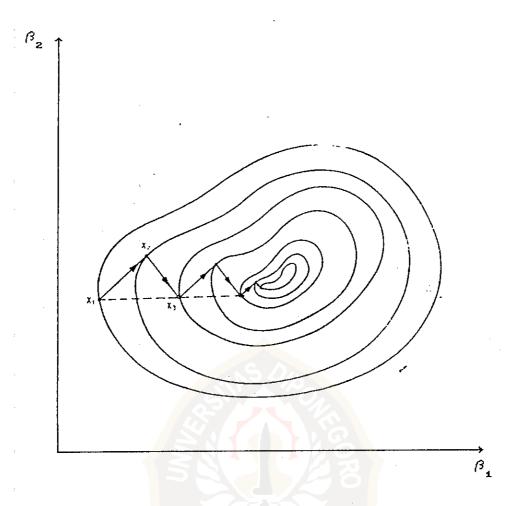
$$\nabla \mathbf{f}_{n \times 1} = \begin{bmatrix} \frac{\partial \mathbf{f}}{\partial \mathbf{x}_{1}}, & \frac{\partial \mathbf{f}}{\partial \mathbf{x}_{2}}, & \dots, & \frac{\partial \mathbf{f}}{\partial \mathbf{x}_{n}} \end{bmatrix}$$
(4.6)





Gambar 4.3. Kontur jumlah kuadrat terkecil dalam ruang parameter (Alkis Constantinides, 1979)

- (a) Model linier
- (b) Model non linier



Gambar 4.4. Proses konvergensi metode Steepest Descent (Rao S.S., 1978).

Sifat penting gradien yang dijadikan dasar optimasi adalah pergerakan sepanjang arah gradien dari titik ke titik dalam ruang dimensi-n, menyebabkan nilai fungsi menaik pada kecepatan maksimum (Rao S.S., 1978).

Penggunaan arah negatif dari vektor gradien untuk mencari minimisasi dilakukan pertama kali oleh A.Cauchy, 1847. Dalam metode ini proses optimasi beranjak dari titik awal x_i dan beriterasi untuk mendapatkan titik minimum berdasarkan persamaan (4.12).

$$\mathbf{x}_{(i+1)} = \mathbf{x}_i + \lambda_i^* \mathbf{s}_i = \mathbf{x}_i - \lambda_i^* \nabla \mathbf{f}_i$$
 (4.7)

dengan λ adalah panjang langkah optimal sepanjang arah pencarian $s_i = -\nabla f_i$. Proses konvergensi metode steepest descent dapat dilihat pada gambar 4.4.

Bila posisi awal jauh dari nilai minimum metode bekerja baik tetapi ketika mendekati nilai minimum metode menjadi lambat karena perilaku zigzagging menjadi semakin sering (N. Drafer dan H. Smith, 1966). Kelemahan lain pula metode steepest descent tidak invarian terhadap skala. Arah yang ditunjukkan berubah bila skala diubah.

4.3.2.2. Metoda Gauss Newton

Konsep kunci yang mendasari metode Gauus Newton adalah uraian deret *Taylor* yang digunakan untuk menyatakan persamaan non linier menjadi bentuk hampiran linier, lihat gambar 4.5 (*Steven C.C dan Raymond P.C., 1989*).

Dari persamaan (4.3) fungsi non linier diuraikan menurut deret Taylor di sekitar nilai parameter dan dihentikan setelah turunan pertamanya,

$$\langle Y(x_i, b + \varepsilon_T) \rangle = f(x_i, b) + \sum_{j=1}^{k} \left[\frac{\partial f_i}{\partial b_j} \right] (\varepsilon_T)_j$$

$$\langle Y \rangle = f_0 + J\varepsilon_T$$
(4.8)

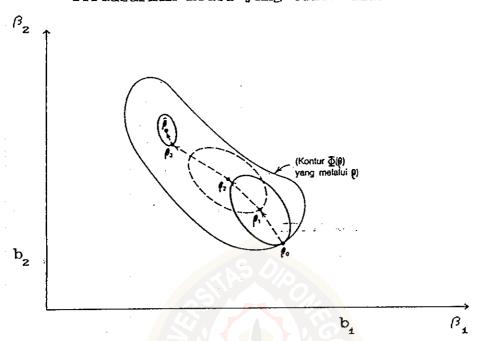
dengan

- * b merupakan estimasi kuadrat terkecil darieta.
- * $arepsilon_{_{
 m T}}$ adalah vektor koreksi terhadap eta hasil

15

uraian Taylor.

< > menunjukkan perkiraan yang dibuat berdasarkan model yang telah dilinierkan.



Gambar 4.5. Rangkaian iterasi suatu proses linierrisasi ditunjukkan konvergen menuju titik $\hat{\beta}$ sehingga meminimumkan $\Phi(\beta)$ (N.Drafer.dan H. Smith, 1981).

Maka bentuk persamaan (4.12) dapat ditulis dalam bentuk

$$\langle \Phi \rangle = \sum_{i=1}^{n} [Y_{i} - \langle Y_{i} \rangle]^{2}$$
 (4.9)

Vektor koreksi ε dapat diperoleh dengan menggunakan syarat perlu $\frac{\partial <\Phi >}{\partial \varepsilon_{_{
m T}}}=0$ yakni dengan menyelesaikan

$$Ae_{_{\mathbf{T}}} = \mathbf{g} \tag{4.10}$$

dengan

$$A_{(k \times k)} = J^{T}J$$

$$J_{(n \times k)} = \frac{\partial f_{i}}{\partial b_{j}}, \quad i = 1, 2, ..., k$$

$$g = \sum_{i=1}^{n} (Y_{i} - f_{i}) = J^{T}(Y - f_{o})$$

Dari pemecahan persamaan (4.10) untuk harga $\varepsilon_{_{
m T}}$ maka parameter baru yang diperbaiki didapatkan, yaitu

$$\mathbf{b}_{(\mathbf{r}+\mathbf{i})} = \mathbf{b}_{(\mathbf{r})} + \varepsilon_{\mathbf{r}} \tag{4.11}$$

Linearisasi akan berhasil bila nilai awal tidak terlalu jauh dari nilai optimum sebab kontur model akan dapat dihampiri dengan baik oleh kontur model yang telah dilinearkan (N. Drafer dan H. Smith, 1989).

4.3.2.3. Metode Marguardt

Dua masalah yang timbul dari kedua metode di atas adalah singularitas matrik A dan skala ruang parameter(*J.C. Nash*, 1979). Metode Marquardt secara simultan memecahkan kedua masalah tersebut.

Seabagaimana telah dibahas dalam metode steepes descent sifat-sifat yang berhubungan dengan solusi vektor koreksi $\varepsilon_{_{\rm T}}$ adalah invarian di bawah transformasi linier di ruang b. Sifat-sifat dari metode gradien adalah tidak scale invariant, sehingga perlu untuk men-scale ruang b. Untuk memenuhi hal tersebut maka ruang b di-scale dalam

satuan standar deviasi dari turunan $\frac{\partial I_i}{\partial b_j}$ untuk titik-titik i = 1, 2, ..., n. Karena pada umumnya turunan tergantung pada b maka nilai terbaru b dipergunakan seperlunya. Pemilihan scale digunakan untuk meningkatkan aspek numerik dari prosedur-prosedur perhitungan.

Dalam hal ini didefinisikan matriks ter-scale A^* dan vektor ter-scale g^*

$$A^* \varepsilon_T^* = g^*$$

$$\varepsilon_j = \frac{\varepsilon_j^*}{(a_{ij})^{1/2}}$$
(4.12)

Sehingga dalam iterasi ke-r akan dibentuk

$$(A_{(r)}^* + \lambda^* I) \varepsilon_{(r)} = g_{(r)}^*$$

$$(4.13)$$

Perubahan bentuk persamaan (4.12) menjadi bentuk persamaan (4.13) yang dilakukan oleh *Donald W. Marguardt* didukung oleh teorema di bawah ini.

Teorema 4.7

Jika λ sembarang dengan $\lambda \geq 0$ dan ε_0 memnuhi persamaan

 $(A + \lambda I)\varepsilon_0 = g$. Maka ε_0 akan meminimumkan $< \Phi >$ pada bola berjejari $\| \varepsilon \|$ yang memenuhi $\| \varepsilon \|^2 = \| \varepsilon_0 \|^2$ (D.W. Marquardt, 1963).

Teorema 4.8

Jika γ sudut antara $\varepsilon_{_{\mathbf{0}}}$ dan $\varepsilon_{_{\mathbf{g}}}$, maka γ adalah fungsi

monoton dari λ sedemikian sehingga $\lambda \longrightarrow \infty$, γ $\longrightarrow 0$. Karena ε_0 , tidak bergantung pada λ maka ε_g berotasi terhadap ε_g sementara $\lambda \longrightarrow \infty$ (D.W. Marquardt, 1963).

Untuk nilai λ yang besar sekali, matriks $(A + \lambda I)$ didominasi oleh diagonal λI . Dapat terlihat dari persamaan (4.13) bahwa λ menuju takhingga maka $\varepsilon_0 = \frac{g}{\lambda}$ dimana ε_0 sebanding dalam limit, sehingga sudut antara keduanya menuju nol. Sebaliknya bila $\lambda = 0$, maka vektor ε_0 dan g membentuk sudut γ dengan $0 \le \gamma \le \Pi/2$ (kecuali kasus trivial dimana A adalah diagonal). Sehingga γ adalah fungsi turun monoton dari λ sedemikian sehingga $\lambda \longrightarrow \infty$ dan $\gamma \longrightarrow 0$.

Persamaan (4.13) dapat diselesaikan dengan didukung oleh teorema berikut

Teorema 4.9

Jika $\varepsilon(\lambda)$ adalah solusi dari persamaan (4.15) untuk suatu nilai λ . Maka $\| \varepsilon(\lambda) \|^2$ adalah fungsi turun monoton dari λ sedemikian sehingga $\lambda \longrightarrow \infty$, $\| \varepsilon(\lambda) \|^2$ menuju nol (*D.W. Marquardt*, 1963).

Persamaan (4.12) digunakan untuk mendapatkan vektor koreksi ϵ^* sehingga parameter baru menjadi

$$b_{(r+1)} = b_{(r)} + \varepsilon_{(r)} \tag{4.18}$$

yang akan mengahasilkan jumlah kuadrat $\Phi_{(r}$, baru-

document is Undip Institutional Repository Collection. The author(s) or copyright owner(s) agree that UNDIP-IR may, without changing the content, translate mission to any medium or format for the purpose of preservation. The author(s) or copyright owner(s) also agree that UNDIP-IR may keep more than one copy of

Nilai λ dipilih sedemikian sehingga memenuhi kriteria $\Phi_{(r+1)} < \Phi_{(r)}.$

Dengan demikian dipilih nilai λ yang kecil jika kondisinya sedemikian sehingga metode Gauss Newton dapat mencapai konvergensi yang baik, yaitu apabila sudah berada di daerah konvergensi. Sedangkan pemilihan harga λ yang besar dilakukan jika masih berada di luar daerah konvergensi (Donald W. Marguardt, 1963).

