
METODE LEVENBERG-MARQUARDT UNTUK MASALAH KUADRAT TERKECIL NONLINEAR

Lusia Krismiyati Budiasih
Program Studi Matematika
Universitas Sanata Dharma Yogyakarta
lusia_kris@staff.usd.ac.id

Abstrak

Kuadrat terkecil nonlinear merupakan bentuk dari analisis kuadrat terkecil yang digunakan untuk melakukan *fitting* atas sekumpulan pengamatan dengan suatu model non linear multivariabel. Dalam masalah kuadrat terkecil nonlinear, akan ditentukan vektor parameter model yang akan menghasilkan kecocokan yang paling mungkin antara pengukuran dan prediksi model. Dalam hal ini akan digunakan jumlah kuadrat terbobot sebagai ukuran kecocokan, yang menjadi fungsi obyektif dalam masalah optimisasi. Tujuan dari masalah ini adalah menentukan argumen dari fungsi obyektif yang akan meminimumkan nilai fungsi tersebut.

Pada umumnya metode untuk menyelesaikan masalah optimisasi nonlinear adalah secara iteratif. Metode-metode ini memiliki langkah-langkah yang bertujuan supaya nilai fungsi obyektif yang akan diminimumkan mengalami penurunan pada iterasi selanjutnya. Pada dasarnya dalam setiap iterasi terdiri dari tahap menentukan arah turun (*descent direction*) dan panjang langkah yang akan memberikan penurunan yang baik terhadap nilai fungsi obyektif.

Metode Levenberg-Marquardt merupakan salah satu metode optimasi untuk menyelesaikan masalah kuadrat terkecil yang didasarkan pada metode Gauss-Newton. Pada metode Levenberg-Marquadt, arah turun ditentukan dengan mempertimbangkan parameter damping yang akan mempengaruhi arah dan juga besar langkah. Dalam banyak kasus, metode ini akan mencapai kekonvergenan yang lebih baik daripada konvergen secara linear.

Kata kunci: kuadrat terkecil nonlinear, metode Levenberg-Marquadt, metode Gauss-Newton, *descent methods*.

A. Pendahuluan

Dari sebuah penelitian, sering diperoleh sekumpulan data numerik yang akan digunakan untuk mencari hubungan antar variabel-variabel yang diamati. Hubungan tersebut biasanya dinyatakan dalam suatu fungsi agar dapat juga digunakan untuk memprediksi nilai fungsi di suatu titik tertentu. Banyak metode yang digunakan untuk

menentukan fungsi yang menjadi hampiran dari data-data yang ada. Salah satu metode yang dapat digunakan adalah metode kuadrat terkecil. Kuadrat terkecil dapat diinterpretasikan sebagai metode pencocokan data (*fitting data*). Kesesuaian antara data model dan data hasil pengamatan diukur melalui nilai terkecil dari jumlah kuadrat residu, dengan residu merupakan jarak antara nilai pengamatan dan nilai yang diberikan oleh model. Metode ini pertama kali diperkenalkan oleh Carl Friederich Gauss sekitar tahun 1794.

Kuadrat terkecil nonlinear merupakan bentuk dari analisis kuadrat terkecil yang digunakan untuk melakukan *fitting* atas sekumpulan pengamatan dengan suatu model non linear multivariabel. Dalam masalah kuadrat terkecil nonlinear, akan ditentukan vektor parameter model yang akan menghasilkan kecocokan yang paling mungkin antara pengukuran dan prediksi model. Dalam hal ini akan digunakan jumlah kuadrat terbobot sebagai ukuran kecocokan, yang menjadi fungsi obyektif dalam masalah optimisasi.

Masalah optimisasi tak berkendala bertujuan untuk meminimumkan suatu fungsi bernilai real F dengan n variabel. Hal ini berarti bahwa akan ditentukan suatu peminimal lokal, yakni titik \mathbf{x}^* sedemikian sehingga

$$F(\mathbf{x}^*) \leq F(\mathbf{x}) \text{ untuk semua } \mathbf{x} \text{ dekat } \mathbf{x}^*.$$

Masalah di atas biasanya dinyatakan sebagai

$$\min_{\mathbf{x}} F(\mathbf{x}).$$

Fungsi F disebut sebagai fungsi obyektif dan $F(\mathbf{x}^*)$ merupakan nilai minimum.

Masalah minimisasi lokal berbeda dengan masalah minimisasi global. Suatu peminimal global adalah titik \mathbf{x}^* sedemikian sehingga

$$F(\mathbf{x}^*) \leq F(\mathbf{x}) \text{ untuk semua } \mathbf{x}.$$

Beberapa metode yang dapat digunakan untuk menyelesaikan masalah minimisasi lokal adalah metode Newton maupun metode Gauss-Newton, yang lebih efisien digunakan untuk menyelesaikan masalah kuadrat terkecil nonlinear.

Masalah kuadrat terkecil dapat diselesaikan dengan metode optimisasi pada umumnya. Namun dalam tulisan ini akan dipaparkan metode khusus yang lebih efisien,

yakni metode Levenberg-Marquardt. Dalam banyak kasus, metode ini akan mencapai kekonvergenan yang lebih baik daripada konvergen secara linear; terkadang akan konvergen secara kuadratik.

B. Descent Methods

Semua metode untuk menyelesaikan masalah optimisasi nonlinear adalah iteratif, yakni dari suatu titik awal \mathbf{x}_0 , metode akan menghasilkan suatu deret vektor-vektor $\mathbf{x}_1, \mathbf{x}_2, \dots$, yang diharapkan akan konvergen ke \mathbf{x}^* , suatu peminimal lokal dari fungsi yang diberikan. Jika fungsi yang diberikan memiliki beberapa peminimal, hasil yang diperoleh akan bergantung pada titik awal \mathbf{x}_0 . Pada umumnya metode iteratif tersebut akan memenuhi kondisi

$$F(\mathbf{x}_{k+1}) < F(\mathbf{x}_k).$$

Hal ini mencegah kekonvergenan ke suatu pemaksimal. Metode yang memenuhi kondisi di atas biasanya disebut sebagai *descent method*. Beberapa metode yang termasuk dalam *descent method* adalah metode turun tercuram (*steepest descent method*) atau metode gradient, metode Newton, *line search methods*, dan *trust region methods* (Peressini, 1988).

Dalam metode-metode tersebut, setiap langkah dari proses iterasi akan memenuhi kondisi di atas (Madsen, 2004). Pada dasarnya dalam setiap iterasi terdiri dari:

- 1) Menentukan arah turun (*descent direction*) \mathbf{h}_d , dan
- 2) Menentukan panjang langkah yang memberikan penurunan yang baik terhadap nilai F .

C. Masalah Kuadrat Terkecil Nonlinear

Tujuan yang ingin dicapai dalam masalah kuadrat terkecil adalah menentukan \mathbf{x}^* , yakni peminimal lokal dari

$$F(\mathbf{x}) = \frac{1}{2} \sum_{i=1}^m (f_i(\mathbf{x}))^2,$$

dengan $f_i : R^n \rightarrow R, i = 1, \dots, m$ dan $m \geq n$.

Dalam masalah kuadrat terkecil nonlinear, fungsi obyektif didefinisikan dengan

$$\begin{aligned} F(\mathbf{x}) &= \frac{1}{2} \sum_{i=1}^m (f_i(\mathbf{x}))^2 \\ &= \frac{1}{2} \|\mathbf{f}(\mathbf{x})\|_2^2 = \frac{1}{2} \mathbf{f}(\mathbf{x})^T \mathbf{f}(\mathbf{x}) \end{aligned} \quad (1)$$

Vektor $\mathbf{f} = (f_1, \dots, f_m)$, dengan $\mathbf{f} : R^n \rightarrow R^m$, disebut sebagai *residual*. Masalah ini dapat terjadi seperti dalam pencocokan data.

Misalkan \mathbf{f} memiliki turunan parsial kedua yang kontinu, ekspansi Taylor fungsi tersebut dapat dinyatakan sebagai

$$\mathbf{f}(\mathbf{x} + \mathbf{h}) = \mathbf{f}(\mathbf{x}) + \mathbf{J}(\mathbf{x})\mathbf{h} + O(\|\mathbf{h}\|^2), \quad (2)$$

dengan $\mathbf{J}(\mathbf{x})$ adalah matriks Jacobian, yakni

$$(\mathbf{J}(\mathbf{x}))_{ij} = \frac{\partial f_i}{\partial x_j}(\mathbf{x}), \quad 1 \leq i \leq m, 1 \leq j \leq n.$$

Dari persamaan (1) akan diperoleh

$$\frac{\partial F}{\partial x_j}(\mathbf{x}) = \sum_{i=1}^m f_i(\mathbf{x}) \frac{\partial f_i}{\partial x_j}(\mathbf{x}), \quad (3)$$

dan gradien dari F

$$\begin{aligned} \nabla F(\mathbf{x}) = \mathbf{F}'(\mathbf{x}) &= \begin{pmatrix} \frac{\partial F}{\partial x_1}(\mathbf{x}) \\ \vdots \\ \frac{\partial F}{\partial x_n}(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^m f_i(\mathbf{x}) \frac{\partial f_i}{\partial x_1}(\mathbf{x}) \\ \vdots \\ \sum_{i=1}^m f_i(\mathbf{x}) \frac{\partial f_i}{\partial x_n}(\mathbf{x}) \end{pmatrix} \\ &= \mathbf{J}(\mathbf{x})^T \mathbf{f}(\mathbf{x}). \end{aligned} \quad (4)$$

Dengan mengingat kembali bahwa matriks Hessian untuk F adalah $H = \left(\frac{\partial^2 F}{\partial x_i \partial x_j}(\mathbf{x}) \right)$,

maka dari persamaan (3), elemen matriks Hessian pada posisi (j,k) adalah

$$\frac{\partial^2 F}{\partial x_j \partial x_k}(\mathbf{x}) = \sum_{i=1}^m \left(\frac{\partial f_i}{\partial x_j}(\mathbf{x}) \frac{\partial f_i}{\partial x_k}(\mathbf{x}) + f_i(\mathbf{x}) \frac{\partial^2 f_i}{\partial x_j \partial x_k}(\mathbf{x}) \right).$$

Dengan demikian diperoleh

$$\mathbf{F}''(\mathbf{x}) = \mathbf{J}(\mathbf{x})^T \mathbf{J}(\mathbf{x}) + \sum_{i=1}^m f_i(\mathbf{x}) \mathbf{f}_i''(\mathbf{x}).$$

Dasar dari metode untuk menyelesaikan masalah kuadrat terkecil nonlinear ini adalah metode Gauss-Newton. Metode ini didasarkan pada aproksimasi linear untuk komponen-komponen \mathbf{f} (model linear dari \mathbf{f}) di persekitaran \mathbf{x} , yakni untuk $\|\mathbf{h}\|$ yang kecil, ekspansi Taylor dari persamaan (2) menghasilkan

$$\mathbf{f}(\mathbf{x} + \mathbf{h}) \approx l(\mathbf{h}) \equiv \mathbf{f}(\mathbf{x}) + \mathbf{J}(\mathbf{x})\mathbf{h}. \quad (5)$$

Dengan menerapkan persamaan (5) untuk definisi F pada persamaan (1) maka akan diperoleh

$$\begin{aligned} F(\mathbf{x} + \mathbf{h}) \approx L(\mathbf{h}) &\equiv \frac{1}{2} l(\mathbf{h})^T l(\mathbf{h}) \\ &= \frac{1}{2} (\mathbf{f} + \mathbf{J}\mathbf{h})^T (\mathbf{f} + \mathbf{J}\mathbf{h}) \\ &= \frac{1}{2} (\mathbf{f}^T + \mathbf{h}^T \mathbf{J}^T) (\mathbf{f} + \mathbf{J}\mathbf{h}) \\ &= \frac{1}{2} (\mathbf{f}^T \mathbf{f} + \mathbf{h}^T \mathbf{J}^T \mathbf{f} + \mathbf{f}^T \mathbf{J}\mathbf{h} + \mathbf{h}^T \mathbf{J}^T \mathbf{J}\mathbf{h}) \\ &= \frac{1}{2} \mathbf{f}^T \mathbf{f} + \mathbf{h}^T \mathbf{J}^T \mathbf{f} + \frac{1}{2} \mathbf{h}^T \mathbf{J}^T \mathbf{J}\mathbf{h} \\ &= F(\mathbf{x}) + \mathbf{h}^T \mathbf{J}^T \mathbf{f} + \frac{1}{2} \mathbf{h}^T \mathbf{J}^T \mathbf{J}\mathbf{h} \end{aligned} \quad (6)$$

Gradien dan Hessian dari L masing-masing adalah

$$L'(\mathbf{h}) = \mathbf{J}^T \mathbf{f} + \mathbf{J}^T \mathbf{J}\mathbf{h} \quad \text{dan} \quad L''(\mathbf{h}) = \mathbf{J}^T \mathbf{J}.$$

Dari persamaan (4), dapat dilihat bahwa $L'(\mathbf{0}) = \mathbf{F}'(\mathbf{x})$. Juga tampak bahwa $L''(\mathbf{h})$ tidak bergantung pada \mathbf{h} dan matriks tersebut simetris. Jika \mathbf{J} memiliki rank penuh, yakni setiap kolom dari \mathbf{J} saling bebas linear, maka $L''(\mathbf{h})$ definit positif. Hal ini menunjukkan bahwa $L(\mathbf{h})$ memiliki peminimal tunggal, yang dapat ditentukan dengan menyelesaikan

$$(\mathbf{J}^T \mathbf{J}) \mathbf{h}_{gn} = -\mathbf{J}^T \mathbf{f}. \quad (7)$$

D. Metode Levenberg-Marquardt

Metode *steepest descent* tidak memiliki cara yang baik untuk menentukan panjang langkah. Sedangkan metode Newton didasarkan pada penyelesaian suatu

sistem linear. Jika matriks yang diperoleh adalah singular, maka metode ini gagal untuk menyelesaikan masalah optimisasi. Selain itu, jika iterasi diawali dari suatu titik yang tidak terlalu dekat dengan peminimumnya, maka kadang-kadang metode ini memberikan iterasi yang divergen yang menjauh dari peminimumnya. Oleh karena itu, berikut akan dipaparkan metode yang dapat mengatasi kelemahan-kelemahan di atas.

Levenberg (1944) dan kemudian Marquardt (1963) menggunakan prinsip Gauss-Newton. Langkah \mathbf{h}_{lm} ditentukan dengan memodifikasi persamaan (7), yakni

$$(\mathbf{J}^T \mathbf{J} + \mu \mathbf{I}) \mathbf{h}_{lm} = -\mathbf{g} \quad \text{dengan } \mathbf{g} = \mathbf{J}^T \mathbf{f} \text{ dan } \mu \geq 0.$$

Parameter damping μ memiliki beberapa pengaruh:

- Untuk semua $\mu > 0$, matriks koefisien adalah definit positif, dan hal ini menjamin bahwa \mathbf{h}_{lm} merupakan arah turun (*descent direction*).
- Untuk nilai μ yang besar, akan diperoleh

$$\mathbf{h}_{lm} \approx -\frac{1}{\mu} \mathbf{g} = -\frac{1}{\mu} \mathbf{F}'(\mathbf{x}),$$

yakni langkah yang pendek dalam arah turun tercuram. Langkah ini baik jika iterasi yang sedang berlangsung jauh dari penyelesaiannya.

- Jika μ amat kecil, maka $\mathbf{h}_{lm} \approx \mathbf{h}_{gn}$, yang merupakan langkah yang baik dalam tahap iterasi terakhir, saat \mathbf{x} dekat dengan \mathbf{x}^* . Jika $F(\mathbf{x}^*) = 0$ (atau amat kecil), maka iterasi dapat konvergen kuadratik.

Dengan demikian, parameter μ akan mempengaruhi arah dan juga besar langkah.

Pemilihan nilai μ awal harus disesuaikan dengan banyaknya elemen dari

$\mathbf{A}_0 = \mathbf{J}(\mathbf{x}_0)^T \mathbf{J}(\mathbf{x}_0)$, yakni dengan memilih

$$\mu_0 = \tau \cdot \max_i \{a_{ii}^{(0)}\},$$

dengan τ dipilih sembarang. Algoritma yang dibangun tidak akan terlalu sensitif terhadap pemilihan τ , tetapi biasanya akan digunakan nilai yang kecil. Jika \mathbf{x}_0 diyakini merupakan aproksimasi yang baik untuk \mathbf{x}^* , maka dapat dipilih $\tau = 10^{-6}$. Dapat juga digunakan $\tau = 10^{-3}$ atau bahkan $\tau = 1$.

Selama iterasi, besar nilai μ dapat diperbaharui. Hal tersebut dikendalikan dengan ratio

$$\rho = \frac{F(\mathbf{x}) - F(\mathbf{x} + \mathbf{h}_{lm})}{L(\mathbf{0}) - L(\mathbf{h}_{lm})}.$$

Penyebut dari ratio tersebut diprediksi dengan model linear pada persamaan (6) sehingga diperoleh

$$\begin{aligned} L(\mathbf{0}) - L(\mathbf{h}_{lm}) &= F(\mathbf{x}) - \left(F(\mathbf{x}) + \mathbf{h}_{lm}^T \mathbf{J}^T \mathbf{f} + \frac{1}{2} \mathbf{h}_{lm}^T \mathbf{J}^T \mathbf{J} \mathbf{h}_{lm} \right) \\ &= -\mathbf{h}_{lm}^T \mathbf{J}^T \mathbf{f} - \frac{1}{2} \mathbf{h}_{lm}^T \mathbf{J}^T \mathbf{J} \mathbf{h}_{lm} \\ &= -\frac{1}{2} \mathbf{h}_{lm}^T (2\mathbf{J}^T \mathbf{f} + \mathbf{J}^T \mathbf{J} \mathbf{h}_{lm}) \\ &= -\frac{1}{2} \mathbf{h}_{lm}^T (2\mathbf{g} + (\mathbf{J}^T \mathbf{J} + \mu \mathbf{I} - \mu \mathbf{I}) \mathbf{h}_{lm}) \\ &= -\frac{1}{2} \mathbf{h}_{lm}^T (2\mathbf{g} - \mathbf{g} - \mu \mathbf{h}_{lm}) = \frac{1}{2} \mathbf{h}_{lm}^T (\mu \mathbf{h}_{lm} - \mathbf{g}) \end{aligned}$$

Perhatikan bahwa $\mathbf{h}_{lm}^T \mathbf{h}_{lm}$ dan $-\mathbf{h}_{lm}^T \mathbf{g}$ bernilai positif, maka $L(\mathbf{0}) - L(\mathbf{h}_{lm})$ juga bernilai positif.

Nilai ρ yang besar menunjukkan bahwa $L(\mathbf{h}_{lm})$ merupakan aproksimasi yang baik untuk $F(\mathbf{x} + \mathbf{h}_{lm})$, dan nilai μ dapat ditentukan sedemikian sehingga langkah Levenberg-Marquardt berikutnya lebih dekat dengan langkah Gauss-Newton. Jika ρ kecil (atau bahkan negatif), maka $L(\mathbf{h}_{lm})$ merupakan aproksimasi yang buruk, dan selanjutnya nilai μ ditingkatkan dengan tujuan supaya langkah lebih dekat dengan arah turun tercuram dan menurunkan panjang langkah.

Kriteria penghentian untuk algoritma ini diberikan untuk menunjukkan bahwa pada suatu peminimal global didapatkan $\mathbf{F}'(\mathbf{x}^*) = \mathbf{g}(\mathbf{x}^*) = \mathbf{0}$, sehingga dapat digunakan

$$\|\mathbf{g}\|_{\infty} \leq \varepsilon_1,$$

dengan ε_1 merupakan bilangan positif yang kecil. Kriteria penghentian lain yang dapat digunakan adalah jika perubahan dalam \mathbf{x} kecil, yakni

$$\|\mathbf{x}_{baru} - \mathbf{x}\| \leq \varepsilon_2 (\|\mathbf{x}\| + \varepsilon_2).$$

Algoritma Metode Levenberg-Marquardt

Langkah 1. Inisialisasi $k = 0, v = 2$

Langkah 2. Tentukan $\mathbf{x} = \mathbf{x}_0$

Langkah 3. Hitung $\mathbf{A} = \mathbf{J}(\mathbf{x})^T \mathbf{J}(\mathbf{x})$, $\mathbf{g} = \mathbf{J}^T \mathbf{f}$ dan $\mu = \tau \cdot \max_i \{a_{ii}\}$

Langkah 4. **while** $\|\mathbf{g}\|_\infty \leq \varepsilon_1$ **and** $k < kmax$

$k = k + 1$

selesaikan $(\mathbf{A} + \mu \mathbf{I})\mathbf{h}_{lm} = -\mathbf{g}$

if $\|\mathbf{x}_{baru} - \mathbf{x}\| \leq \varepsilon_2 (\|\mathbf{x}\| + \varepsilon_2)$

break

else

$$\mathbf{x}_{baru} = \mathbf{x} + \mathbf{h}_{lm}, \quad \rho = \frac{F(\mathbf{x}) - F(\mathbf{x} + \mathbf{h}_{lm})}{L(\mathbf{0}) - L(\mathbf{h}_{lm})}$$

if $\rho > 0$

$$\mathbf{x} = \mathbf{x}_{baru}, \quad \mathbf{A} = \mathbf{J}(\mathbf{x})^T \mathbf{J}(\mathbf{x}), \quad \mathbf{g} = \mathbf{J}^T \mathbf{f}$$

$$\mu = \mu \cdot \max\left\{\frac{1}{3}, 1 - (2\rho - 1)^3\right\}, \quad v = 2$$

else

$$\mu = \mu * v, \quad v = 2 * v$$

end

Metode Levenberg-Marquardt untuk Masalah Powell

Sebagai kasus khusus, akan dipertimbangkan masalah berikut, yang dikenal sebagai masalah Powell (1970), yakni

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} x_1 \\ \frac{10x_1}{x_1 + 0,1} + 2x_2^2 \end{bmatrix}$$

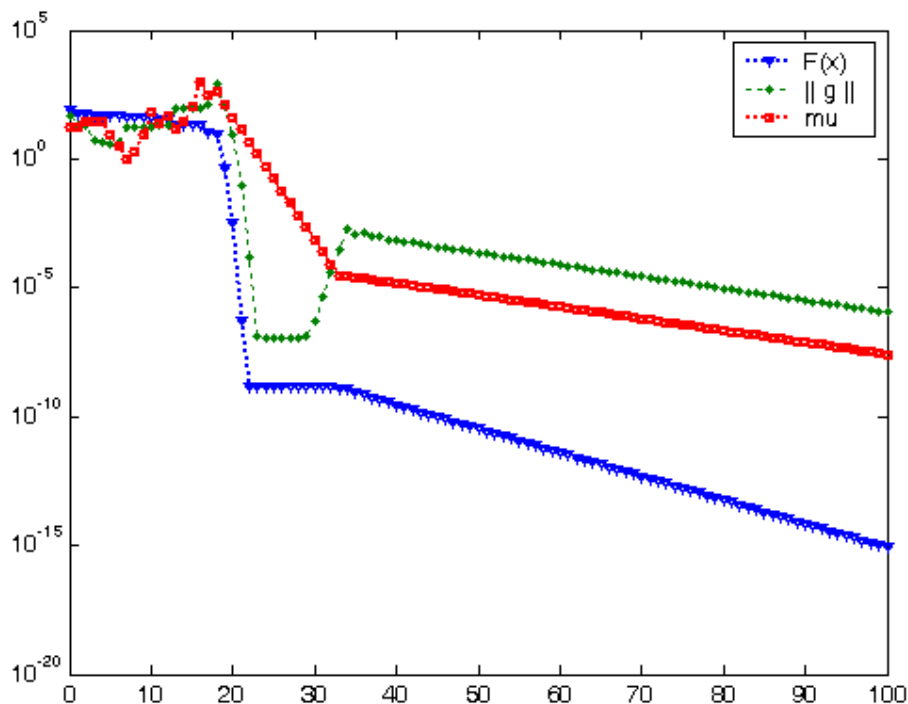
dimana $\mathbf{x}^* = \mathbf{0}$ merupakan satu-satunya penyelesaian dari masalah tersebut.

Perhatikan bahwa matriks Jacobian yang diperoleh adalah

$$\mathbf{J}(\mathbf{x}) = \begin{bmatrix} 1 & 0 \\ (x_1 + 0,1)^{-2} & 4x_2 \end{bmatrix}$$

yang singular pada penyelesaian di atas.

Jika dipilih $\mathbf{x}_0 = [3 \ 1]$, maka akan diperoleh hasil iterasi seperti tergambarkan pada Gambar 1 berikut.



Gambar 1. Metode Levenberg-Marquardt untuk Masalah Powell

Dari Gambar 1 tampak bahwa iterasi mulai stabil terjadi antara iterasi ke-22 dan 30. Hal ini terjadi karena pengaruh dari matriks Jacobian yang hampir mendekati singular. Setelah iterasi tersebut, tampak bahwa iterasi akan konvergen secara linear. Penyelesaian yang didapat dengan metode ini adalah

$$\mathbf{x} = [-3,81 \cdot 10^{-8}; -1,38 \cdot 10^{-3}]$$

Adanya osilasi pada beberapa iterasi di awal dapat menyebabkan kesalahan yang cukup besar dan memperlambat kekonvergenannya.

E. Simpulan dan Saran

Metode Levenberg-Marquardt merupakan salah satu metode optimasi untuk menyelesaikan masalah kuadrat terkecil yang didasarkan pada metode Gauss-Newton. Pada metode Levenberg-Marquadt, arah turun ditentukan dengan mempertimbangkan parameter damping yang akan mempengaruhi arah dan juga besar langkah. Dalam banyak kasus, metode ini akan mencapai kekonvergenan yang lebih baik daripada konvergen secara linear.

Salah satu kelemahan dari metode Levenberg-Marquadt adalah bila terjadi osilasi pada proses iterasinya, maka iterasi akan berjalan lambat. Untuk itu, perlu ada modifikasi terhadap beberapa langkah agar proses iterasi lebih stabil dan berjalan cepat, antara lain dengan menambahkan variabel laju penurunan (Chen, 2003).

F. Daftar Pustaka

Chen, T. and D. Han. (2003). Acceleration of Levenberg-Marquardt Training of Neural Networks with Variable Decay Rate.

<http://hub.hku.hk/handle/123456789/47064> (diakses 15 Oktober 2009)

Madsen, K, H.B. Nielsen, and O. Tingleff. (2004). *Methods For Non-Linear Least Squares Problems, 2nd Edition*, Informatics and Mathematical Modelling Technical University of Denmark.

http://www2.imm.dtu.dk/pubdb/views/edoc_download.../imm3215.pdf

(diakses 20 September 2009)

Peressini, A.L., F.E. Sullivan, and J.J. Uhl, Jr. (1988). *The Mathematics of Nonlinear Programming*. NY: Springer-Verlag.

_____. Levenberg-Marquardt Method.

http://www.physics.utah.edu/~detar/phys6720/handouts/curve_fit/curve_fit/no_de7.html (diakses 5 Oktober 2009)