

TINJAUAN TERHADAP METODE PENGOPTIMUMAN PENDEKATAN NEWTON

B. P. Silalahi ¹, Siswandi ², A. Aman ³

Abstrak

Saat ini teknik-teknik pengoptimuman semakin berkembang. Hal ini sejalan dengan berkembangnya teknologi komputer dan juga semakin kompleksnya masalah pengoptimuman. Untuk pengembangan metode-metode baru diperlukan pengetahuan yang mumpuni tentang dasar-dasar pengoptimuman. Pada paper ini kami menyajikan teknik pengoptimuman dengan menggunakan metode Davidon-Fletcher-Powell beserta analisis yang detail tentang teorema-teorema yang mendasarinya. Kemudian dengan menggunakan bantuan suatu perangkat lunak metode ini diaplikasikan untuk menyelesaikan suatu kasus pengoptimuman nonlinear.

Kata Kunci : pengoptimuman, metode DFP, metode quasi-Newton, metode pendekatan Newton

2010 Mathematics Subject Classification: 97M40.

1 Pendahuluan

Pengoptimuman adalah suatu ilmu dalam matematika yang mencari nilai maksimum atau minimum dari suatu fungsi. Ilmu pengoptimuman ini sering digunakan dalam kehidupan. Penelitian-penelitian tentang pengoptimuman baik pengembangan aplikasi maupun teori terus dilakukan. Beberapa contoh penelitian yang lebih ke arah aplikasi tentang pengoptimuman antara lain dapat dilihat pada [23], [1], [19], [18], [16], [26], [15]. Selanjutnya penelitian yang berorientasi pada pengembangan teori antara lain [6], [21], [22], [25], [10], [24], [13], [14].

Pada paper ini kami akan mengkaji suatu metode pengoptimuman yang disebut dengan nama metode pendekatan Newton (quasi-Newton) dalam menyelesaikan masalah pengoptimuman nonlinear. Metode ini berkembang untuk mengatasi ketidak praktisan dalam masalah pengoptimuman nonlinear dengan menggunakan metode Newton. Pada metode Newton di setiap iterasinya dilakukan

¹ Departemen Matematika, FMIPA- IPB, Jalan Meranti Kampus IPB Dramaga Bogor, 16680.
E-mail : bibparuhum@gmail.com

² Departemen Matematika, FMIPA- IPB, Jalan Meranti Kampus IPB Dramaga Bogor, 16680.
E-mail : siswandi_aja@yahoo.com

³ Departemen Matematika, FMIPA- IPB, Jalan Meranti Kampus IPB Dramaga Bogor, 16680.
E-mail : amril.aman@gmail.com

penghitungan nilai balikan matriks Hessian. Idenya adalah menggunakan pendekatan terhadap nilai balikan matriks Hessian pada setiap iterasi.

Terdapat berbagai cara pendekatan nilai balikan matriks Hessian, dari yang paling sederhana yaitu dengan menggunakan nilai pendekatan yang tetap di setiap iterasinya sampai yang lebih rumit yaitu dengan menggunakan informasi yang ada pada iterasi-iterasi sebelumnya. Beberapa metode yang menggunakan pendekatan ini antara lain metode Broyden [2], metode SR1 ([3], [5]), metode DFP ([5], [8]), metode BFGS ([4], [7], [9], [20]), dan metode-metode perbaikan pendekatan-Newton lainnya ([17], [11], [12]).

Pada kajian ini kami akan membahas sifat-sifat metode pendekatan Newton dengan pembuktiannya. Pembuktian sifat-sifat ini kami titik beratkan pada metode Davidon, Fletcher dan Powell (DFP), dengan harapan dapat menjadi landasan teori yang kuat untuk mengembangkan metode-metode baru. Pembuktian-pembuktian teorema bersifat rinci agar dapat mudah dimengerti. Kemudian kami menerapkan metode DFP dalam suatu perangkat lunak matematika dan menyelesaikan suatu kasus permasalahan pengoptimuman nonlinear kuadratik.

Struktur panyajian adalah sebagai berikut. Pertama kami akan menyajikan metode pendekatan Newton : ukuran langkah λ -nya dan kekonvergensiannya. Kemudian pada bagian selanjutnya bagaimana langkah untuk memperoleh pendekatan balikan matriks. Kemudian dibahas algoritme DFP beserta sifat positif definit dan jaminan kekonvergenannya. Selanjutnya disajikan hasil percobaan untuk fungsi nonlinear kuadratik sempurna. Ditutup dengan kesimpulan, saran dan daftar pustaka.

2 Metode Pendekatan Newton

Pandang permasalahan pengoptimuman berikut :

$$\min f(\mathbf{x})$$

Suatu proses iterasi untuk menyelesaikan masalah pengoptimuman tersebut berbentuk :

$$x_{k+1} = x_k - \lambda_k D_k \nabla f(x_k)^T, \quad (1)$$

dimana D_k adalah matriks simetri berukuran $n \times n$ dan λ_k ukuran langkah yang meminimumkan $f(x_{k+1})$. Kemudian $\nabla f(x)$ adalah vektor baris berdimensi n . Jika D_k adalah balikan dari matriks Hessian fungsi f , maka kita memperoleh langkah iterasi metode Newton yaitu :

$$x_{k+1} = x_k - \lambda_k H_k^{-1} \nabla f(x_k)^T.$$

Selanjutnya jika D_k adalah matriks identitas kita memperoleh langkah iterasi metode *steepest descent*, yaitu

$$x_{k+1} = x_k - \lambda_k \nabla f(x_k)^T.$$

Untuk menjamin iterasi metode pengoptimuman pada 1 berlangsung secara menurun untuk suatu nilai λ_k , maka diperlukan D_K positif definit. Untuk itu selanjutnya kita mengasumsikan D_k positif definit.

Berikutnya kecepatan konvergen disajikan untuk masalah fungsi kuadrat $f(x)$ dengan :

$$f(x) = \frac{1}{2}x^T Kx - b^T x,$$

dimana matriks K simetri dan positif definit. Algoritme dapat ditulis

$$x_{k+1} = x_k - \lambda_k D_k g_k,$$

dimana g_k adalah ringkasan penulisan untuk $g(x_k)$, sedangkan

$$g(x_k) = \nabla f(x_k)^T = Kx_k - b.$$

Untuk kasus fungsi kuadrat diperoleh Teorema berikut.

Teorema 2.1 *Misalkan $f(x)$ fungsi kuadrat dengan*

$$f(x) = \frac{1}{2}x^T Kx - b^T x,$$

dimana matriks K simetri dan positif definit. Algoritme pendekatan Newton untuk x_{k+1} adalah

$$x_{k+1} = x_k - \lambda_k D_k g_k, \quad (2)$$

dengan

$$g_k = Kx_k - b, \quad (3)$$

dan

$$\lambda_k = \frac{g_k^T D_k g_k}{g_k^T D_k K D_k g_k}. \quad (4)$$

Bukti : Proses iterasi untuk menyelesaikan masalah pengoptimuman berbentuk :

$$x_{k+1} = x_k - \lambda_k D_k \nabla f(x_k)^T.$$

Untuk $f(x) = \frac{1}{2}x^T Kx - b^T x$, dapat diperoleh $\nabla f(x_k) = x_k^T K - b^T$. Dengan demikian $g_k = \nabla f(x_k)^T = Kx_k - b$. Kemudian akan dicari nilai λ_k sehingga $f(x_k - \lambda_k D_k g_k)$ minimum. Untuk kasus ini kita memiliki:

$$f(x_k - \lambda_k D_k g_k) = \frac{1}{2}(x_k - \lambda_k D_k g_k)^T K(x_k - \lambda_k D_k g_k) - (x_k - \lambda_k D_k g_k)^T b.$$

Turunan $f(x_k - \lambda_k D_k g_k)$ terhadap λ_k adalah :

$$\begin{aligned}
\frac{df}{d\lambda_k} &= \frac{1}{2} [2\lambda_k (D_k g_k)^T K D_k g_k - x_k^T K D_k g_k - (D_k g_k)^T K x_k] + (D_k g_k)^T b \\
&= \frac{1}{2} [2\lambda_k g_k^T D_k K D_k g_k - x_k^T K D_k g_k - g_k^T D_k K x_k] + g_k^T D_k b \\
&= \frac{1}{2} [2\lambda_k g_k^T D_k K D_k g_k - g_k^T D_k K x_k - g_k^T D_k K x_k] + g_k^T D_k b \\
&= \lambda_k g_k^T D_k K D_k g_k - g_k^T D_k K x_k + g_k^T D_k b \\
&= \lambda_k g_k^T D_k K D_k g_k - g_k^T D_k (K x_k - b) \\
&= \lambda_k g_k^T D_k K D_k g_k - g_k^T D_k g_k.
\end{aligned}$$

Kemudian untuk meminimumkan f dalam λ_k kita menset $\frac{df}{d\lambda_k} = 0$. Ini berarti kita memperoleh :

$$\lambda_k g_k^T D_k K D_k g_k - g_k^T D_k g_k = 0.$$

Dengan memindahkan ruas kita peroleh

$$\lambda_k g_k^T D_k K D_k g_k = g_k^T D_k g_k.$$

Oleh karenanya

$$\lambda_k = \frac{g_k^T D_k g_k}{g_k^T D_k K D_k g_k}.$$

□

Berikut ini disajikan kecepatan konvergen dari metode pendekatan-Newton.

Teorema 2.2 Misal x^* adalah titik minimum dan $F(x) = \frac{1}{2}(x - x^*)^T K(x - x^*)$. Untuk Teorema (2.1) pada setiap iterasi ke k berlaku

$$F(x_{k+1}) \leq \left(\frac{E_k - e_k}{E_k + e_k} \right)^2 F(x_k),$$

dimana E_k dan e_k adalah nilai eigen terbesar dan terkecil dari matriks $D_k K$.

Bukti : Pertama akan kita perhatikan bahwa

$$\frac{F(x_k) - F(x_{k+1})}{F(x_k)} = \frac{(g_k^T D_k g_k)^2}{(g_k^T D_k K D_k g_k)(g_k^T K^{-1} g_k)}.$$

Perhatikan

$$\begin{aligned} F(x_k) &= \frac{1}{2}(x_k - x^*)^T K(x_k - x^*) \\ &= \frac{1}{2}(x_k^T K x_k - x_k^T K x^* - x^{*T} K x_k - x^{*T} K x^*) \\ &= \frac{1}{2}x_k^T K x_k - x_k^T K x^* - \frac{1}{2}x^{*T} K x^*. \end{aligned}$$

$$\begin{aligned} F(x_{k+1}) &= \frac{1}{2}(x_{k+1} - x^*)^T K(x_{k+1} - x^*) \\ &= \frac{1}{2}(x_{k+1}^T K x_{k+1} - x_{k+1}^T K x^* - x^{*T} K x_{k+1} - x^{*T} K x^*) \\ &= \frac{1}{2}x_{k+1}^T K x_{k+1} - x_{k+1}^T K x^* - \frac{1}{2}x^{*T} K x^*. \end{aligned}$$

Dengan demikian kita peroleh

$$\begin{aligned} F(x_k) - F(x_{k+1}) &= \frac{1}{2}x_k^T K x_k - x_k^T K x^* - \frac{1}{2}x_{k+1}^T K x_{k+1} + x_{k+1}^T K x^* \\ &= \frac{1}{2}x_k^T K x_k - x_k^T K x^* - \frac{1}{2}x_{k+1}^T K x_{k+1} + x_{k+1}^T K x^*. \end{aligned}$$

Selanjutnya dengan menggunakan(2) ruas kanan persamaan di atas sama dengan

$$\begin{aligned} &\frac{1}{2}x_k^T K x_k - x_k^T K x^* - \frac{1}{2}(x_k - \lambda_k D_k g_k)^T K(x_k - \lambda_k D_k g_k) + (x_k - \lambda_k D_k g_k)^T K x^* = \\ &\frac{1}{2}x_k^T K x_k - x_k^T K x^* - \frac{1}{2}x_k^T K x_k + \lambda_k x_k^T K D_k g_k - \frac{1}{2}\lambda_k^2 x_k^T g_k^T D_k K D_k g_k + x_k^T K x^* - \\ &\lambda_k g_k^T D_k K x^* = \lambda_k x_k^T K D_k g_k - \frac{1}{2}\lambda_k^2 x_k^T g_k^T D_k K D_k g_k - \lambda_k g_k^T D_k K x^*. \end{aligned}$$

Oleh karena itu

$$\begin{aligned} F(x_k) - F(x_{k+1}) &= \lambda_k x_k^T K D_k g_k - \frac{1}{2}\lambda_k^2 x_k^T g_k^T D_k K D_k g_k - \lambda_k g_k^T D_k K x^* \\ &= \lambda_k g_k^T D_k K x_k - \frac{1}{2}\lambda_k^2 x_k^T g_k^T D_k K D_k g_k - \lambda_k g_k^T D_k K x^* \\ &= \lambda_k g_k^T D_k (K x_k - K x^*) - \frac{1}{2}\lambda_k^2 x_k^T g_k^T D_k K D_k g_k. \end{aligned}$$

Putaran gradien $F(x)$ adalah $g(x) = Kx - Kx^*$. Dengan demikian

$$F(x_k) - F(x_{k+1}) = \lambda_k g_k^T D_k g_k - \frac{1}{2}\lambda_k^2 x_k^T g_k^T D_k K D_k g_k.$$

Kemudian dengan menggunakan ukuran langkah λ pada (4) diperoleh

$$\begin{aligned} F(x_k) - F(x_{k+1}) &= \frac{(g_k^T D_k g_k)^2}{g_k^T D_k K D_k g_k} - \frac{(g_k^T D_k g_k)^2}{2g_k^T D_k K D_k g_k} \\ &= \frac{(g_k^T D_k g_k)^2}{2g_k^T D_k K D_k g_k}. \end{aligned}$$

$F(x_k)$ dapat pula dinyatakan sebagai berikut

$$\begin{aligned} F(x_k) &= \frac{1}{2}(x_k - x^*)^T K (x_k - x^*) \\ &= \frac{1}{2}(x_k - x^*)^T g_k \\ &= \frac{1}{2}(K^{-1} g_k)^T g_k \\ &= \frac{1}{2}g_k^T K^{-1} g_k. \end{aligned}$$

Dengan demikian diperoleh

$$\frac{F(x_k) - F(x_{k+1})}{F(x_k)} = \frac{(g_k^T D_k g_k)^2}{(g_k^T D_k K D_k g_k)(g_k^T K^{-1} g_k)}.$$

Ambil $T_k = D_k^{1/2} K D_k^{1/2}$ dan $p_k = D_k^{1/2} g_k$, diperoleh

$$\frac{F(x_k) - F(x_{k+1})}{F(x_k)} = \frac{(p_k^T p_k)^2}{(p_k^T T_k p_k)(p_k^T T_k^{-1} p_k)}.$$

Dengan menggunakan pertidaksamaan Kantorovich diperoleh

$$F(x_{k+1}) \leq \left(\frac{E_k - e_k}{E_k + e_k} \right)^2 F(x_k),$$

dimana E_k dan e_k adalah nilai eigen terbesar dan terkecil dari T_k . Karena $D_k^{1/2} T_k D_k^{-1/2} = D_k K$, $D_k K$ dan T_k memiliki nilai eigen yang sama. \square

Teorema (2.2) bermakna bahwa untuk fungsi kuadratik untuk memperoleh kekonvergenan yang cepat nilai eigen terbesar dan terkecil haruslah dibuat sedekat mungkin. Hal ini akan tercapai bilamana D_k dekat dengan K^{-1} , karena berarti $D_k K$ akan mendekati matriks identitas. Bila $D_k K$ mendekati matriks identitas maka nilai eigen $D_k K$ akan dekat ke satu. Untuk permasalahan optimasi dengan fungsi tujuan berupa fungsi nonkuadratik K didekati dengan matriks Hessian $H(x)$. Oleh karena itu untuk memperoleh kekonvergenan yang relatif cepat kita harus memilih D_k yang dekat dengan $H(x_k)^{-1}$.

3 Pendekatan Balikan Matriks Hessian

Ide dasar dari metode pendekatan Newton adalah mengganti penghitungan nilai balikan matriks Hessian pada setiap iterasi algoritme optimasi metode Newton. Ada beberapa cara yang dapat dilakukan, dari yang paling sederhana dengan mengambil nilai tetap untuk balikan matriks Hessian, sampai dengan yang lebih rumit yaitu dengan menggunakan informasi yang diperoleh pada iterasi sebelumnya. Pada bagian ini akan dibahas bagaimana balikan matriks Hessian dapat diperoleh dengan menggunakan informasi gradien.

Teorema 3.1 *Misalkan f suatu fungsi pada R^n yang memiliki turunan parsial kedua kontinu. Misalkan terdapat dua titik x_{k+1} dan x_k . Didefinisikan $g_{k+1} = \nabla f(x_{k+1})^T$, $g_k = \nabla f(x_k)^T$ serta $p_k = x_{k+1} - x_k$. Maka*

$$q_k \equiv g_{k+1} - g_k \approx H(x_k)p_k. \quad (5)$$

Jika Hessian H konstan, maka

$$q_k \equiv g_{k+1} - g_k = Hp_k. \quad (6)$$

Bukti : Berdasarkan Teorema Taylor di sekitar $x = x_k$ berlaku

$$f(x) \approx f(x_k) + (x - x_k)^T \nabla f(x_k) + \frac{1}{2}(x - x_k)^T H(x_k)(x - x_k).$$

Kemudian gradien formula tersebut adalah sebagai berikut

$$\nabla f(x) \approx \nabla f(x_k) + x^T H(x_k) - x_k^T H(x_k).$$

Sehingga

$$\begin{aligned} \nabla f(x_{k+1}) - \nabla f(x_k) &\approx x_{k+1}^T H(x_k) - x_k^T H(x_k) \\ &= (x_{k+1} - x_k)^T H(x_k) \\ &= p_k^T H(x_k). \end{aligned}$$

Kita memperoleh (5) yaitu

$$g_{k+1} - g_k \approx H(x_k)p_k.$$

Dengan cara yang sama (6) juga dapat dibuktikan. \square

Jika p_0, p_1, \dots, p_{n-1} bebas linear dan q_k diketahui, maka H dapat ditentukan.

Misalkan P dan Q adalah matriks berukuran $n \times n$ dengan kolomnya masing-masing berunsur p_k dan q_k , kita memperoleh

$$H = QP^{-1}.$$

Untuk mendapatkan pendekatan H^{-1} berdasarkan data yang diperoleh dari iterasi sebelumnya, apabila H konstan, kita memiliki

$$H_{k+1}^{-1}q_i = p_i, \quad 0 \leq i \leq k.$$

Setelah n langkah bebas linear kita akan memiliki nilai eksak dari H_n^{-1} .

4 Metode Davidon-Fletcher-Powell

Davidon, Fletcher dan Powell adalah ilmuwan yang memelopori cara pendekatan terhadap balikan matriks Hessian. Ide awalnya berasal dari Davidon yang kemudian dikembangkan oleh Fletcher dan Powell.

Algoritme metode ini dimulai dari matriks simetri positif definit \mathcal{H} , titik awal x_0 dan $k = 0$. Berikut ini adalah algoritme Metode Davidon-Fletcher-Powell (DFP).

Langkah 1. Tentukan d_k menurut $d_k = -\mathcal{H}_k g_k$.

Langkah 2. Minimumkan $f(x_k + \lambda d_k)$ terhadap $\lambda \geq 0$, sehingga diperoleh x_{k+1} , $p_k = \lambda_k d_k$, dan g_{k+1} .

Langkah 3. Tentukan $q_k = g_{k+1} - g_k$ dan

$$\mathcal{H}_{k+1} = \mathcal{H}_k + \frac{p_k p_k^T}{p_k^T q_k} - \frac{\mathcal{H}_k q_k q_k^T \mathcal{H}_k}{q_k^T \mathcal{H}_k q_k}. \quad (7)$$

Ubah nilai k menjadi $k = k + 1$. Kembali ke langkah 1 selama kriteria pemberhentian belum tercapai.

4.1 Sifat Positif Definit

Berikut ini akan diperlihatkan bahwa bila \mathcal{H}_k positif definit, maka \mathcal{H}_{k+1} juga positif definit. Hal ini untuk menjamin langkah menurun menuju solusi minimum pada metode DFP.

Bila kita mengalikan persamaan (7) dengan x^T dan x diperoleh

$$x^T \mathcal{H}_{k+1} x = x^T \mathcal{H}_k x + \frac{(x^T p_k)^2}{p_k^T q_k} - \frac{(x^T \mathcal{H}_k q_k)^2}{q_k^T \mathcal{H}_k q_k}. \quad (8)$$

Definisikan $r = \mathcal{H}_k^{1/2} x$ dan $s = \mathcal{H}_k^{1/2} q_k$, kita dapat menuliskan kembali (8) sebagai berikut

$$\begin{aligned} x^T \mathcal{H}_{k+1} x &= r^T r + \frac{(x^T p_k)^2}{p_k^T q_k} - \frac{(r^T s)^2}{s^T s} \\ &= \frac{(r^T r)(s^T s) - (r^T s)^2}{s^T s} + \frac{(x^T p_k)^2}{p_k^T q_k}. \end{aligned}$$

Selanjutnya karena $q_k = g_{k+1} - g_k$ kita memiliki

$$p_k^T q_k = p_k^T (g_{k+1} - g_k) = p_k^T g_{k+1} - p_k^T g_k.$$

Titik minimum dari $f(x_k + \lambda d_k)$ relatif terhadap λ diperoleh diperoleh dengan cara menurunkan fungsi tersebut terhadap λ dan $\frac{df}{d\lambda} = 0$. Kita memperoleh

$$\begin{aligned} \frac{df}{d\lambda} &= \nabla f(x_k + \lambda d_k) d_k^T \\ &= \nabla f(x_{k+1}) d_k \\ &= d_k^T \nabla^T f(x_{k+1}) \\ &= d_k^T g(x_{k+1}) = 0. \end{aligned}$$

Mengalikan dengan λ_k diperoleh

$$\lambda_k d_k^T g(x_{k+1}) = 0 \Leftrightarrow p_k^T g(x_{k+1}) = 0$$

Oleh karena itu

$$p_k^T g_{k+1} = 0. \quad (9)$$

Kemudian dengan menggunakan definisi p_k didapat

$$p_k^T q_k = \lambda g_k^T \mathcal{H}_k g_k. \quad (10)$$

Oleh karenanya

$$x^T \mathcal{H}_{k+1} x = \frac{(r^T r)(s^T s) - (r^T s)^2}{s^T s} + \frac{(x^T p_k)^2}{\lambda g_k^T \mathcal{H}_k g_k}.$$

Formula pada ruas kanan keduanya tidak negatif. Menggunakan pertidaksamaan Cauchy–Schwarz jelas formula pertama tidak negatif. Formula kedua juga tidak negatif, karena pembilangnya berbentuk kuadrat dan penyebutnya menggunakan asumsi bahwa \mathcal{H}_k positif definit. Kemudian formula pada ruas kanan tidak mungkin bernilai 0, seperti ditunjukkan berikut. Formula pertama bernilai 0

hanya jika r dan s proporsional. Sesuai dengan rumus r dan s , jika r dan s proporsional maka x dan q_k proporsional. Misalkan $x = \gamma q_k$. Sehingga kita memiliki, dengan menggunakan (10)

$$p_k^T x = \gamma p_k^T q_k = \gamma \lambda_k g_k^T \mathcal{H}_k g_k \neq 0.$$

Pertidaksamaan terakhir diperoleh karena \mathcal{H}_k positif definit. Oleh karenanya $x^T \mathcal{H}_{k+1} x > 0$ untuk semua x , yang berarti \mathcal{H}_{k+1} positif definit.

4.2 Kekonvergenan

Pada bagian ini ditunjukkan bahwa metode DFP menghasilkan vektor-vektor p_k yang H -ortogonal dan jika metode ini dilakukan dalam n langkah diperoleh $\mathcal{H}_n = H^{-1}$, seperti dinyatakan dalam teorema berikut.

Teorema 4.1 *Jika f adalah persamaan kuadratik dengan matriks Hessian H positif definit, maka metode Davidon-Fletcher-Powell memenuhi sifat-sifat berikut*

$$p_i^T H p_j = 0, \quad 0 \leq i < j \leq k \quad (11)$$

$$\mathcal{H}_{k+1} H p_i = p_j, \quad 0 \leq i \leq k. \quad (12)$$

Bukti : Berikut ini adalah pendekatan matriks balikan Hessian yang dilakukan pada metode DFP seperti pada (7)

$$\mathcal{H}_{k+1} = \mathcal{H}_k + \frac{p_k p_k^T}{p_k^T q_k} - \frac{\mathcal{H}_k q_k q_k^T \mathcal{H}_k}{q_k^T \mathcal{H}_k q_k}.$$

Persamaan di atas dikalikan dengan $H p_k$ sehingga diperoleh

$$\mathcal{H}_{k+1} H p_k = \mathcal{H}_k H p_k + \frac{p_k p_k^T H p_k}{p_k^T q_k} - \frac{\mathcal{H}_k q_k q_k^T \mathcal{H}_k H p_k}{q_k^T \mathcal{H}_k q_k}.$$

Pada (6) kita memiliki

$$q_k = g_{k+1} - g_k = H p_k. \quad (13)$$

Bila kita substitusikan ke formula sebelumnya diperoleh

$$\mathcal{H}_{k+1} H p_k = \mathcal{H}_k q_k + \frac{p_k p_k^T q_k}{p_k^T q_k} - \frac{\mathcal{H}_k q_k q_k^T \mathcal{H}_k q_k}{q_k^T \mathcal{H}_k q_k}.$$

Formula pertama dan ketiga pada ruas kanan saling menghilangkan sedangkan formula kedua sama dengan p_k , akibatnya

$$\mathcal{H}_{k+1} H p_k = p_k. \quad (14)$$

Fakta-fakta di atas akan kita gunakan dalam membuktikan (11) dan (12). Bukti akan menggunakan induksi matematika.

Basis Induksi

Untuk (11), kita menggunakan $k = 1$, sehingga $i = 0$ dan $j = 1$. Dengan menggunakan definisi p_k dan d_k diperoleh

$$p_0^T H p_1 = \lambda_1 p_0^T H d_1 = -\lambda_1 p_0^T H \mathcal{H}_1 g_1.$$

Menggunakan (13) dengan $k = 0$, yaitu $\mathcal{H}_1 H p_0 = p_0$ didapat

$$p_0^T H p_1 = -\lambda_1 p_0^T g_1.$$

Selanjutnya karena $p_0^T g_1 = 0$ (lihat (9)) berarti

$$p_0^T H p_1 = 0.$$

Untuk (12), kita menggunakan $k = 0$, yang berarti $i = 0$. Menggunakan (14) untuk $k = 0$, basis induksi benar.

Langkah Induksi

Hipotesakan bahwa (11) dan (12) benar untuk $k - 1$, akan ditunjukkan bahwa (11) dan (12) juga benar untuk k . Berdasarkan (13) kita punya

$$g_k = g_{i+1} + H(p_{i+1} + \dots + p_{k-1}).$$

Oleh karenanya dengan menggunakan hipotesa induksi dan (9) didapat

$$p_i^T g_k = p_i^T g_{i+1} = 0, \text{ untuk } 0 \leq i < k.$$

Kemudian dengan menggunakan putaran hipotesa induksi (berkenaan dengan (12)) kita memperoleh

$$p_i^T g_k = p_i^T g_{i+1} = p_i^T H \mathcal{H}_k g_k = 0, \text{ untuk } 0 \leq i < k.$$

Selanjutnya karena $p_k = -\lambda_k \mathcal{H}_k g_k$ dan $\lambda_k \neq 0$,

$$p_i^T H p_k = 0, \text{ untuk } 0 \leq i < k, \tag{15}$$

sehingga (11) terbukti.

Dari perkalian (7) dengan $H p_i$ didapat

$$\mathcal{H}_{k+1} H p_i = \mathcal{H}_k H p_i + \frac{p_k p_k^T H p_i}{p_k^T q_k} - \frac{\mathcal{H}_k q_k q_k^T \mathcal{H}_k H p_i}{q_k^T \mathcal{H}_k q_k}. \tag{16}$$

Kemudian dengan mengalikan q_k^T dengan hipotesa induksi (berkenaan dengan (12)) diperoleh

$$q_k^T \mathcal{H}_k H p_i = q_k^T p_i \text{ untuk } 0 \leq i < k.$$

Mengaplikasikan (13) dan (15) berakibat

$$q_k^T p_i = p_k^T H p_i = 0, \text{ untuk } 0 \leq i < k.$$

dengan mensubstitusikan ke (16) dan menggunakan $\mathcal{H}_k H p_i = p_i$ kita memperoleh

$$\mathcal{H}_{k+1} H p_i = p_i, \text{ untuk } 0 \leq i < k.$$

Formula tersebut bersama dengan (14) membuktikan (12). \square

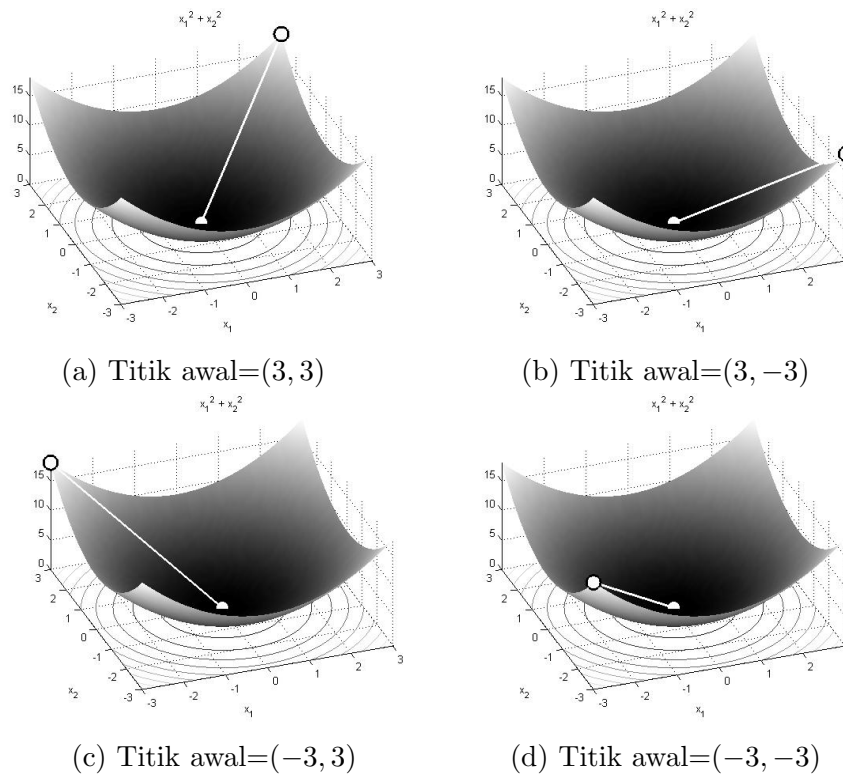
Karena p_k bersifat H -ortogonal dan f diminimumkan menurut arah p_k , metode ini merupakan suatu metode dengan arah konjugat gradien. Jika \mathcal{H}_0 sama dengan matriks identitas metode ini menjadi metode konjugat gradien. Kemudian (12) memperlihatkan bahwa p_0, p_1, \dots, p_k vektor-vektor eigen yang berhubungan dengan nilai eigen satu untuk matriks $\mathcal{H}_{k+1} H$. Vektor-vektor tersebut saling bebas linear dan oleh karenanya $\mathcal{H}_n = H^{-1}$.

5 Hasil Numerik

Untuk percobaan algoritme DFP digunakan fungsi nonlinear kuadratik sempurna yaitu

$$f(x) = (x - x^*)^T K (x - x^*),$$

dimana K matriks simetri berukuran $n \times n$ dan $K_{i,j} = 0$ untuk $i \neq j$. Percobaan menggunakan $\mathcal{H}_0 = H^{-1}$. Hasil percobaan menunjukkan algoritme akan menemukan solusi minimum dalam satu iterasi untuk semua nilai n yang dicoba. Gambar 1 menyajikan contoh untuk $n = 2$ dan $f(x) = x_1^2 + x_2^2$ yang dimulai dari beberapa titik awal yang berbeda. Dalam satu kali iterasi menuju titik minimum $(0, 0)$.



Gambar 1: Iterasi Metode DFP dengan $f(x) = x_1^2 + x_2^2$ dan $\mathcal{H}_0 = H^{-1}$.

6 Kesimpulan dan Saran

Metode Davidon-Fletcher-Powell mendapatkan solusi minimum dari fungsi kuadratik sempurna dalam 1 iterasi.

Penelitian-penelitian selanjutnya yang dapat dilakukan antara lain :

- Membandingkan kecepatan pengoptimuman Metode Davidon-Fletcher-Powell dengan metode pengoptimuman lainnya.
- Mengembangkan metode pendekatan balikan matriks Hessian.
- Mengembangkan metode iterasi pengoptimuman.

Daftar Pustaka

- [1] Bisilisin, F. Y., Herdiyeni, Y., Silalahi, B. P. Optimasi k-means clustering menggunakan particle swarm optimization pada sistem identifikasi tumbuhan obat berbasis citra. *Jurnal Ilmu Komputer dan Agri-Informatika* 3, 1 (2014), 37-46.

- [2] Broyden, C. A class of methods for solving nonlinear simultaneous equations. *Math. Comp.* 19, 92 (1965), 577–593.
- [3] Broyden, C. Quasi-newton methods and their application to function minimization. *Math. Comp.* 21, 368 (1967).
- [4] Broyden, C. A new double-rank minimization algorithm. *Notices American Math. Soc* 16, 670 (1969).
- [5] Davidon, W. Variable metric method for minimization. Tech. rep., Argonne National Laboratories, Ill., 1959.
- [6] Fadhillah, F., Silalahi, B. P., Ilyas, M. Modifikasi stepsize pada metode steepest descent dalam pengoptimuman fungsi: Kasus fungsi kuadrat diagonal. *Jurnal Matematika dan Aplikasinya* 13, 1 (2014), 47–60.
- [7] Fletcher, R. A new approach to variable metric algorithms. *The Computer Journal* 13, 3 (1970), 317.
- [8] Fletcher, R., Powell, M. A rapidly convergent descent method for minimization. *The Computer Journal* 6, 2 (1963), 163–168.
- [9] Goldfarb, D. A family of variable metric updates derived by variational means. *Math. Comp.* 24, 109 (1970), 23–26.
- [10] Haqueqy, N., Silalahi, B. P., Sitanggang, I. S. Uji komputasi algoritme varian metode Newton pada permasalahan optimasi nonlinear tanpa kendala. *Jurnal Matematika dan Aplikasinya* 15, 2 (2016), 63–76.
- [11] Hennig, P. Fast probabilistic optimization from noisy gradients. *In Proceedings of the 30th International Conference on Machine Learning (ICML)* (2013).
- [12] Hennig, P., Kiefel, M. Quasi-newton methods: A new direction. *Journal of Machine Learning Research* 14 (2013), 843–865.
- [13] Idaman, S., Silalahi, B. P., Guritman, S. Penyelarasan arah vektor gradien untuk menentukan step size metode steepest descent pada fungsi nonlinear kuadratik banyak variabel. *Jurnal Matematika dan Aplikasinya* 17, 1 (2018), 47–59.
- [14] Kautsar, M. M. A., Silalahi, B. P., Guritman, S. Paradox in a d-dimensional transportation problem. *International Journal of Pure and Applied Mathematics* 118, 3 (2018), 547–557.
- [15] Making, S. R. M., Silalahi, B. P., Bukhari, F. Multi depot vehicle routing problem dengan pengemudi sesekali. *Jurnal Matematika dan Aplikasinya* 17, 1 (2018), 75–86.

- [16] Mayyani, H., Silalahi, B. P., Aman, A. Frequency determination of bus rapid transit (BRT) applied on service system of trans mataram metro bus to minimize the operational cost. *International Journal of Engineering and Management Research (IJEMR)* 7, 6 (2017), 134–140.
- [17] Nocedal, J. Updating quasi-newton matrices with limited storage. *Math. Comp.* 35, 151 (1980), 773–782.
- [18] Pratamasyari, D. A., Silalahi, B. P., Guritman, S. Kombinasi varian metode Newton dan metode Halley untuk menyelesaikan persamaan tak linier. *Jurnal Matematika dan Aplikasinya* 16, 2 (2017), 1–12.
- [19] Saifudin, M. A., Silalahi, B. P., Sitanggang, I. S. Star catalog generation for satellite attitude navigation using density based clustering. *Journal of Computer Science* 11, 12 (2015), 1082–1089.
- [20] Shanno, D. Conditioning of quasi-newton methods for function minimization. *Math. Comp.* 24, 111 (1970), 647–656.
- [21] Silalahi, B. P. Kasus-kasus buruk penggunaan metode titik interior pada optimisasi linear. *Jurnal Matematika Integratif* 10, 1 (2014), 9–17.
- [22] Silalahi, B. P. Sharper analysis of upper bound for the iteration complexity of an interior-point method using primal-dual full-Newton step algorithm. *Far East Journal of Mathematical Sciences* 95, 1 (2014), 69–80.
- [23] Silalahi, B. P., Dewi, M. S. Comparison of sensitivity analysis on linear optimization using optimal partition and optimal basis (in the simplex method) at some cases. In *IndoMS International Conference on Mathematics and Its Applications* (2014), Indonesian Mathematical Society, pp. 82–90.
- [24] Silalahi, B. P., Laila, R., Sitanggang, I. S. A combination method for solving nonlinear equations. In *IOP Conference Series: Materials Science and Engineering* (2017), vol. 166, IOP Publishing, p. 012011.
- [25] Silalahi, B. P., Wungguli, D., Guritman, S. Steepest descent method with new step sizes. *World Academy of Science, Engineering and Technology, International Journal of Mathematical, Computational, Physical, Electrical and Computer Engineering* 9, 7 (2015), 378–384.
- [26] Wihartiko, F. D., Buono, A., Silalahi, B. P. Integer programming model for optimizing bus timetable using genetic algorithm. In *IOP Conference Series: Materials Science and Engineering* (2017), vol. 166, IOP Publishing, p. 012016.