Jurnal Karya Asli Lorekan Ahli Matematik Vol. 4 No.1 (2011) Page 46-54

Jurnal
Karya Asli Lorekan
Ahli Matematik

Algoritma Berkesan Kaedah BFGS bagi Masalah Pengoptimuman Tak Berkekangan

Mustafa Bin Mamat[1], Ismail Bin Mohd[3]
Jabatan Matematik, Fakulti Malaysia dan Teknologi
Universiti Malaysia Terengganu (UMT), Malaysia
[1] mus@umt.edu.my, [3]ismailmd@umt.edu.my

.

Mohd Asrul Hery Bin Ibrahim[2]
School of Applied Science and Foundation
Kuala Lumpur Infrastructure University College (KLIUC), Malaysia
[2] asrulhery@yahoo.com.my

Abstrak: Kaedah quasi-Newton merupakan kaedah cukup popular dan berkesan bagi menyelesaikan masalah pengoptimuman tak berkekangan. Kemaskini BFGS merupakan suatu kaedah yang terbaik dalam kaedah quasi-Newton. Dalam kajian ini diperkenalkan suatu algoritma alternatif bagi kaedah BFGS iaitu dengan mengubah saiz langkah dan juga arah carian. Di akhir kajian ini, suatu perbandingan keputusan berangka akan dipaparkan dengan merujuk kepada bilangan lelaran dan bilangan kiraan fungsi untuk membandingkan keberkesanan algoritma yang dicadangkan.

Katakunci: Arah carian, bilangan lelaran, pengoptimuman tak berkekangan, saiz langkah, superlinear.

Abstract: The quasi-Newton methods are well known and effective method in solving unconstrained optimization problems. The most popular method in quasi-Newton methods is BFGS method. In this research, we establish an alternative algorithm for BFGS method which is modifying the step size and search direction. The numerical results based on number of iterations, number of gradient evaluations and number of function evaluations shown in last section to show the effectiveness of our algorithms.

Keywords: Search direction, number of iterations, unconstrained optimization, step size, superlinear.

1. Pengenalan

Pertimbangkan masalah pengoptimuman tak berkekangan

$$\min_{x \in R^n} f(x) \tag{1.1}$$

dengan f adalah fungsi selanjar dan boleh beza dua kali dari R^n kepada R. Kemaskini BFGS merupakan satu kaedah berlelaran yang mana pada lelaran ke-$(k+1)$

$$x_{k+1} = x_k + \alpha_k d_k \tag{1.2}$$

dengan α_k merupakan saiz langkah (Cauchy,1847; Curry,1944)

$$\alpha_k = \arg \min_{\alpha > 0} f(x + \alpha d_k) \tag{1.3}$$

dan d_k merupakan arah carian

$$d_k = -B_k^{-1} g_k \tag{1.4}$$

yang mana g_k adalah kecerunan bagi f. Seterusnya B_k dalam (1.4) merupakan penghampiran bagi matriks Hessian dan ianya dikemaskini melalui rumus

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{s_k^T y_k}, \tag{1.5}$$

dengan $s_k = x_{k+1} - x_k$ dan $y_k = g_{k+1} - g_k$ (Byrd et al., 1987) Penghampiran matriks Hessian (1.5) perlu memenuhi persamaan $B_{k+1} s_k = y_k$.

Artikel ini disusun seperti berikut. Dalam Bahagian 2, perbincangan tentang penentuan saiz langkah. Kemudian diikuti dengan penerangan secara mendalam tentang pengiraan arah carian kacukan BFGS dan penurunan tercuram diberi dalam Bahagian 3. Keputusan berangka serta perbincangan diperihalkan dalam Bahagian 4. Kesimpulan daripada kajian dipaparkan dalam bahagian akhir iaitu Bahagian 5.

2. Saiz Langkah

Dalam kaedah quasi-Newton, penentuan bagi saiz langkah adalah sangat penting untuk mendapatkan penumpuan secara superlinear. Kebanyakan penyelidik sebelum ini hanya menggunakan satu jenis penentuan saiz langkah sahaja seperti dalam Cauchy (1847), Han dan Newmann (2003), Mustafa et al. (2004) dan banyak lagi. Tetapi Yuan (2006) telah mempopularkan kaedah penurunan tercuram yang menggunakan dua jenis saiz langkah. Dalam kajiannya, Yuan (2006) telah mencadangkan kombinasi saiz langkah antara carian garis tepat α_k (Cauchy, 1847) dan juga saiz langkah λ_k dengan

$$\lambda_k := \frac{2}{\sqrt{\left(\frac{1}{\alpha_{k-1}^*} - \frac{1}{\alpha_k^*}\right)^2 + \frac{4\|g_k\|_2^2}{\|s_{k-1}\|_2^2}} + \frac{1}{\alpha_{k-1}^*} + \frac{1}{\alpha_k^*}} \tag{2.1}$$

Penggunaan kedua-dua jenis saiz langkah berikut adalah secara selang-seli yang mana pada lelaran ganjil, carian garis tepat digunakan dan pada lelaran genap, saiz langkah Yuan (2006) digunakan. Hal ini boleh digambarkan sebagai

$$\left.\begin{aligned} x_2 &= x_1 + \alpha_1^* d_1 \\ x_3 &= x_2 + \lambda_2 d_2 \\ x_4 &= x_3 + \alpha_3^* d_3 \\ &\vdots \\ x_{k+1}^* &= x_k + \alpha_k^* d_k . \end{aligned}\right\} \tag{2.2}$$

Namun, dalam kajian ini pengubahsuaian dan kombinasi kepada kaedah BFGS adalah berbeza dengan apa yang dilakukan oleh Yuan (2006). Kami menggunakan carian garis tepat dan juga carian garis

Armijo. Tetapi penentuan saiz langkah dengan menggunakan carian garis Armijo hanya digunakan apabila nilai saiz langkah yang dicari melalui carian garis tepat adalah lebih kecil dari 0.1 ($\alpha_k < 0.1$). Carian garis Armijo boleh dilihat dalam Armijo (1966) dan juga Shi (2006). Carian garis Armijo boleh diringkaskan seperti berikut.

Diberi $s > 0, \beta \in (0,1), \sigma \in (0,1)$ dan $\lambda_k = \text{maks}\{s, s\beta, s\beta^2, \ldots\}$ apabila

$$f(x) - f(x_k + \lambda_k d_k) \geq -\sigma \lambda_k g_k^T d_k \tag{2.3}$$

dengan $k = 0,1,2,\ldots$. Walau bagaimanapun, dalam kaedah yang dicadangkan, penggunaan carian garis Armijo hanya digunakan apabila $k > 2$. Penentuan saiz langkah secara alternatif ini digambarkan dalam Algoritma 3.2 pada bahagian berikut. Dalam Algoritma 3.3, kami menggunakan kaedah hibrid arah carian BFGS dan penurunan tercuram dan akan dibincangkan dalam bahagian seterusnya.

3. Arah Carian

Pengubahsuaian terhadap kaedah quasi-Newton yang dilakukan oleh banyak penyelidik dalam bidang pengoptimuman telah mencetuskan idea kepada Ludwig (2007) untuk mengkombinasikan kaedah quasi-Newton dengan kaedah Gauss-Siedel bagi menyelesaikan sistem persamaan linear. Kemudiannya, Luo et al. (2008) mencadangkan kaedah hibrid quasi-Newton dengan 'pengoptimuman chaos' bagi menyelesaikan sistem persamaan tak linear. Selain daripada itu, Mustafa et al. (2006, 2007) pula mencadangkan satu algoritma baru yang dinamai kaedah quasi-Newton-SD. Kaedah kacukan antara kaedah quasi-Newton dan juga kaedah penurunan tercuram telah meningkatkan lagi keberkesanan bagi menyelesaikan masalah pengoptimuman tak berkekangan dari segi bilangan lelaran dan bilangan kiraan fungsi. Penumpuan arah carian baru ini tetap menumpu secara superlinear (Han dan Newman, 2003; Mustafa et al., 2009).

Perbezaan kaedah biasa dengan kaedah hibrid BFGS-SD adalah pada arah carian. Bagi kaedah hibrid BFGS-SD, arah cariannya dirumuskan seperti

$$d_k = -\eta_k B_k^{-1} g_k - \delta_k g_k \tag{3.1}$$

yang mana parameter $\eta_k > 0$ dan $\delta_k > 0$. Oleh yang demikian, lelaran ke-$(k+1)$ diperoleh melalui

$$x_{k+1} = x_k + \alpha_k(-\eta_k B_k^{-1} g_k - \delta_k g_k). \tag{3.2}$$

Algoritma 3.1, 3.2 dan 3.3 merupakan algoritma bagi kaedah BFGS, kaedah BFGS bersyarat dan juga kaedah hibrid BFGS-SD bersyarat dijelaskan seperti berikut.

Algoritma 3.1: Kaedah BFGS

Data $f \in C^2,\ x \in D$

1. $k := 1$
2. $x_k := x$
3. $g_k := \nabla f(x_k)$
4. $B_k := I_n$

5. $d_k := -B_k^{-1} g_k$
6. Kirakan α_k^* berdasarkan (1.4)
7. $x_{k+1} := x_k + \alpha_k^* d_k$
8. jika $\|g_{k+1}\| \le \varepsilon$, maka berhenti
9. $s_k := \alpha_k^* d_k$
10. $y_k := g_{k+1} - g_k$
11. Kirakan B_{k+1}^{-1} melalui (1.5)
12. $k := k+1$

Algoritma 3.2: Kaedah BFGS Bersyarat

Data $f \in C^2,\ x \in D$

1. $k := 1$
2. $x_k := x$
3. $g_k := \nabla f(x_k)$
4. $B_k := I_n$
5. $d_k := -B_k^{-1} g_k$
6. Kirakan α_k^* berdasarkan (1.4)
7. $x_{k+1} := x_k + \alpha_k^* d_k$
8. Jika $\|g_{k+1}\| \le \varepsilon$ maka berhenti.
9. Kirakan B_{k+1}^{-1} melalui (1.5)
10. $k := k+1$
11. Bagi $k > 2$ jika $\alpha_{k-1}^* < 0.1$

    Gantikan α_k^* dengan λ_k

    11.1 λ_k dikira melalui (2.2)

    11.2. $x_{k+1} := x_k + \lambda_k d_k$

Algoritma 3.3: Kaedah BFGS-SD Bersyarat

Data $f \in C^2,\ x \in D$

1. $k := 1$
2. $x_k := x$
3. $g_k := \nabla f(x_k)$
4. $B_k := I_n$
5. $d_k = -\eta_k B_k^{-1} g_k - \delta_k g_k$
6. Kirakan α_k^* berdasarkan (1.4)
7. $x_{k+1} := x_k + \alpha_k^* d_k$
8. Jika $\|g_{k+1}\| \le \varepsilon$ maka berhenti.
9. Kirakan B_{k+1}^{-1} melalui (1.5)

10. $k := k+1$
11. Bagi $k > 2$ jika $\alpha_{k-1}^* < 0.1$

    Gantikan α_k^* dengan λ_k

    11.1 λ_k dikira melalui (2.2)

    11.2. $x_{k+1} := x_k + \lambda_k d_k$

Ketiga-tiga algoritma di atas diuji keberkesanannya dengan menggunakan beberapa masalah ujian piawai dan dibincangkan dengan lebih lanjut pada Bahagian 4. Perbandingan berangka juga diberikan. Berikut adalah beberapa andaian, teorem dan lemma yang menunjukkan bahawa hibrid arah carian yang digunakan adalah menumpu secara superlinear.

Andaian 3.1

(1). Fungsi objektif bagi f adalah pembezaan kedua selanjar.

(2). Set bagi L adalah cembung**.** Tambahan pula, wujud pemalar-pemalar positif iaitu c_1 dan c_2 yang memenuhi

$$c_1 \|z\|^2 \le z^T G(x) z \le c_2 \|z\|^2$$

bagi semua $z \in R^n$ dan $x \in L,$ yang mana $G(x)$ adalah matriks Hessian bagi f.

Andaian 3.2

Matriks Hessian adalah Lipschitz selanjar pada titik x^*, yang mana wujud pemalar positif, c_3 yang memenuhi

$$\|G(x) - G(x^*)\| \le c_3 \|x - x^*\|$$

bagi semua x dalam kejiranan bagi x^*.

Teorem 3.1 (Penumpuan Sejagat)

Andaikan bahawa Andaian 3.1 dipenuhi, maka

$$\operatorname*{had}_{k\to\infty} \|g_k\| = 0.$$

Bukti: Rujuk (Mustafa et al., 2007)

Teorem 3.2 (Penumpuan Superlinear)

Anggapkan bahawa Andaian 3.2 dipenuhi. Kemudian, andaikan juga $\{x_k\} \to x^*$ dan jujukan $\{\|B_k\|\}$ dan $\{\|H_k\|\}$ adalah terbatas. Jika $x_{k+1} = x_k + \alpha_k d_k$ memenuhi semua nilai k yang cukup besar dan jika

$$\operatorname*{had}_{k\to\infty} \frac{\|[B_k - H(x^*)d_k]\|}{\|d_k\|} = 0$$

maka

$$\underset{k\to\infty}{\text{had}}\frac{\|x_{k+1}-x^*\|}{\|x_{k+1}-x_k\|}=0.$$

dengan $B_k^{-1}=H_k$.

Bukti: Berdasarkan Han dan Newman (2003), kita dapati

$$B_k d_k = \eta_k g_k - \delta_k B_k g_k .$$

Dengan demikian,

$$\left(B_k - G(x^*)\right) d_k = -(\eta_k I + \delta_k B_k) g_{k+1} + \left(\eta_k I + \delta_k B_k\right) G(x^*) + o\left(\|d_k\|\right)$$

Oleh kerana $\{\|B_k\|\}$ merupakan jujukan terbatas, maka

$$\underset{k\to\infty}{\text{had}}\frac{\|B_k - G\left(x^*\right) d_k\|}{\|d_k\|} = \eta_k \underset{k\to\infty}{\text{had}}\frac{\|g_{k+1}\|}{\|x_{k+1}-x_k\|}.$$

Kemudian, melalui proses algebra mudah akan diperolehi

$$\underset{k\to\infty}{\text{had}}\frac{\|x_{k+1}-x^*\|}{\|x_{k+1}-x_k\|}=0. \quad \blacksquare$$

Teorem ini menunjukkan bahawa arah carian yang digunakan menumpu secara superlinear.

Lemma 3.1
Jika Andaian 3.1 dan 3.2 dipenuhi, maka wujud jujukan-jujukan bagi $\{\varepsilon_k\}$ dengan

$$\frac{\|y_k - G(x^*) s_k\|}{\|s_k\|} \le \varepsilon_k \text{ dan } \sum_{k=1}^{\infty} \varepsilon_k < \infty.$$

yang mana $s_k = x_{k+1} - x_k$ dan $y_k = g_{k+1} - g_k$.

Bukti: Dengan menggunakan pengembangan Taylor, maka dapati

$$\begin{aligned} y_k &= \int_0^1 G(x_x + \tau s_k) s_k d\tau \\ &= G(x^*) s_k. \end{aligned}$$

Oleh yang demikian

$$g_{k+1} - g_k = G(x_k) s_k .$$

Seterusnya didapati

$$\frac{\|y_k - G(x^*)s_k\|}{\|s_k\|} \le \gamma \text{maks}\left\{\|x_{k+1} - x^*\|, \|x_k - x^*\|\right\}.$$

dengan γ merupakan parameter positif. Takrifkan

$$\varepsilon_k = \text{maks}\left\{\|x_{k+1} - x^*\|, \|x_k - x^*\|\right\},$$

maka dengan Andaian 3.1 dan Teorem 3.1 didapati,

$$\sum_{k=1}^{\infty} \varepsilon_k < \infty \quad \blacksquare$$

Selanjutnya boleh dirujuk daripada Byrd et al. (1987) serta Han dan Newman (2003).

4. Keputusan Berangka

Dalam bahagian ini, dipamerkan keputusan berangka untuk membandingkan ke berkesanan ketiga-tiga algoritma yang diterangkan di bahagian sebelum ini. Tujuh ujian piawai masalah pengoptimuman tak berkekangan telah dipilih dan dijalankan ke atas kaedah BFGS, BFGS Bersyarat dan BFGS-SD Bersyarat. Setiap fungsi yang diuji, dengan memilih titik permulaan yang munasabah. Masalah ujian piawaian yang digunakan adalah seperti berikut:-

Masalah 1. (Fungsi Cube, $n = 2$)

$$f(x_1, x_2) = 100(x_2 - x_1^3)^2 + (1 - x_1)^2.$$

Titik Permulaan: $(-1.2, 1.6)$, Penyelesaian: $(1,1)$

Masalah 2. (Fungsi Rosenbrock, $n = 2$)

$$f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2.$$

Titik Permulaan: $(x_1 = 1.6, x_2 = -2.5)$, Penyelesaian: $(1,1)$

Masalah 3. (Fungsi Rosenbrock, $n = 4$)

$$f(x_1, x_2, x_3, x_4) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2 + 100(x_4 - x_3^2)^2 + (1 - x_3)^2.$$

Titik Permulaan: $(-3.2, -3.3, -3.2, -3.3)$, Penyelesaian: $(1,1,1,1)$

Masalah 4. (Fungsi Shalow, $n = 2$)

$$f(x_1, x_2) = (x_2 - x_1^2)^2 + (1 - x_1)^2.$$

Titik Permulaan: $(50, 100)$, Penyelesaian: $(1,1)$

Masalah 5. (Fungsi Shalow, $n = 4$)

$$f(x_1, x_2, x_3, x_4) = (x_2 - x_1^2)^2 + (1 - x_1)^2 + (x_4 - x_3^2)^2 + (1 - x_3)^2.$$

Titik Permulaan: $(x_1 = 20., x_2 = 40., x_3 = 20., x_4 = 40.)$, Penyelesaian: $(1,1,1,1)$

Masalah 6. (Fungsi Strait, $n = 2$)

$$f(x_1, x_2) = (x_2 - x_1^2)^2 + 100(1 - x_1)^2.$$

Titik Permulaan: $(30., 50.)$, Penyelesaian: $(1,1)$

Masalah 7. (Fungsi Strait, $n = 4$)

$$f(x_1, x_2, x_3, x_4) = (x_2 - x_1^2)^2 + 100(1 - x_1)^2 + (x_4 - x_3^2)^2 + 100(1 - x_3)^2.$$

Titik Permulaan: $(30., 50., 30., 50.)$, Penyelesaian: $(1,1,1,1)$

Jadual 4.1: Keputusan berangka bagi Algoritma 3.1, Algoritma 3.2 dan juga Algoritma 3.3

Masalah diuji.	BFGS		BFGS Bersyarat		BFGS-SD Bersyarat	
	Bilangan Lelaran, n_i	Bilangan pengiraan fungsi, n_f	Bilangan Lelaran, n_i	Bilangan pengiraan fungsi, n_f	Bilangan Lelaran, n_i	Bilangan pengiraan fungsi, n_f
1	17	103	10	56	**9**	**50**
2	19	77	17	63	**15**	**55**
3	23	93	21	72	**17**	**63**
4	17	69	17	69	**13**	**53**
5	14	57	14	57	**11**	**41**
6	9	37	9	37	**7**	**29**
7	10	41	10	41	**9**	**37**

Keputusan berangka menunjukkan bilangan lelaran dan bilangan kiraan fungsi yang diperlukan untuk menyelesaikan masalah pengoptimuman tak berkekangan bagi ketiga-tiga algoritma diberi dalam Jadual 4.1. Nombor-nombor yang ditebalkan adalah merupakan bilangan lelaran dan juga bilangan kiraan fungsi yang paling kecil. Oleh itu, Algoritma BFGS-SD Bersyarat adalah merupakan kaedah alternatif yang cukup baik dan berkesan dalam menyelesaikan masalah pengoptimuman tak berkekangan jika dibandingkan dengan Algoritma BFGS dan Algoritma BFGS Bersyarat. Berdasarkan Jadual 4.1, Algoritma BFGS Bersyarat adalah lebih baik daripada Algoritma BFGS. Hal ini menunjukkan bahawa pengubahsuaian pada algoritma asal kaedah BFGS yang hanya menggunakan carian garis tepat kepada algoritma alternatif seperti Algoritma BFGS Bersyarat dan juga BFGS-SD Bersyarat, memberi keputusan yang cukup memberangsangkan.

5. Kesimpulan

Dalam kajian ini, kami mencadangkan algoritma-algoritma BFGS, BFGS Bersyarat dan BFGS-SD Bersyarat yang mana hasil pengubahsuaian terhadap penentuan saiz langkah dan juga arah carian terhadap kemaskini BFGS. Hasil kajian menunjukkan algoritma-algoritma tersebut yang diperkenalkan merupakan suatu algoritma yang baik dan berkesan dalam menyelesaikan masalah pengoptimuman tak berkekangan terhadap beberapa fungsi ujian piawai. Keberkesanan algoritma-algoritma tersebut dapat digambarkan dalam Jadual 4.1 jika dibandingkan dengan kaedah BFGS (Algoritma 3.1). Dengan yang demikian, berkemungkinan banyak lagi pengubahsuaian yang dapat dilakukan terhadap kemaskini BFGS terutamanya pada bahagian-bahagian yang penting seperti penentuan saiz langkah, arah carian dan juga penghampiran kepada Hessian. Mungkin juga, kombinasi beberapa lagi saiz langkah dapat membantu mempertingkatkan lagi sifat asal kemaskini BFGS. Kini kami sedang mengusahakan kajian melibatkan kombinasi arah carian BFGS dan kecerunan konjugat. Dengan mengambil kira bilangan lelaran dan bilangan kiraan fungsi, kami dapati ada kemungkinan kajian selanjutnya boleh diperluaskan kepada kemaskini Broyden.

References

Armijo, L. (1966). Minimization functions having Lipshitz continuous first partial derivatives, *PacificJ. Mathematics.*16:1-3.

Byrd, R. H., Nocedal, J., & Yuan, Y. H. (1987). Global convergence of a class of quasi-Newton method on convex problems, *SIAM Journal of Numerical Analysis* 24: 1171-1190.

Cauchy, A. (1847). General methods in Analysis for the resolution of linear equation, *C. R. Acad. Sci. Paris 25*:536-538.

Curry, H. B. (1944). The method of steepest descent for nonlinear minimization problems, *Quart. Appl. Math.,* 2:258-261.

Han, L. & Newman, M. (2003). Combining quasi-Newton and Cauchy directions, *Int. Journal of Applied Mathematics* 12(2):167-171.

Ludwig, A. (2007). The Gauss-Siedel-quasi-newton method: A hybrid algorithm for solving dynamic economic models, *Journal of Economic Dynamics and Control,* 31: 1610-1632.

Luo, Y. Z., Tang, G. J. & Zhou, L. N. (2008). Hybrid approach for solving systems of nonlinear equations using chaos optimization and quasi-Newton method, *Applied soft Computing* , 8: 1068-1073.

Mustafa Mamat, Yosza Dasril & Ismail Mohd (2004). Kaedah quasi-Newton untuk pengoptimuman tak berkekangan, *Prosiding Simposium Kebangsaan Sains Matematik ke 12, UIA, Gombak Selangor*: 23-25.

Mustafa Mamat, Yosza Dasril & Ismail Mohd (2006). Analisis awal kepeutusan berangka kaedah BFGS-SD bagi pengoptimuman tak berkekangan, *Prosiding Simposium Kebangsaan Sains Matematik ke 14,* UM di PNB Darby Park: 259-262.

Mustafa Mamat, Yosza Dasril & Ismail Mohd (2007). Kaedah PSB-SD bagi pengoptimuman tak berkekangan, *Prosiding Simposium Kebangsaan Sains Matematik ke 15,* Hotel Concorde Shah Alam: 565-569.

Mustafa Mamat, Yosza Dasril, Leong Wah June & Ismail Mohd (2009). Hybrid Broyden method for unconstrained optimization. *Int. Journal of Numerical Methods and Applications,* 1 (2) : 121-129.

Shi, Z.J. (2006). Convergence of quasi-Newton method with new inexact line search, *J. Math. Anal. Appl.* 315: 120-131.

Yuan, Y.X. (2006). A new stepsize for the stepest descent method, *Journal of Computational Mathematics,* 24(2): 149-156.