Matematika, 1999, Jilid 15, bil. 1, hlm. 31–36

Penumpuan Kaedah Kuasi-Newton dengan Gelintaran Garis Armijo Untuk Meminimumkan Fungsi Kuasicembung

Malik Hj. Abu Hassan, Mansor b. Monsi & Leong Wah June
Jabatan Matematik
Fakulti Sains dan Pengajian Alam Sekitar
Universiti Putra Malaysia
43400 Serdang, Selangor, Malaysia

Abstrak Untuk minimumkan suatu fungsi kuasicembung yang selanjar dan boleh beza $f : \Re^n \to \Re$, kaedah kuasi-Newton Armijo menjana suatu jujukan

$$x^{k+1} = x^k - t_k H^k g^k,$$

dengan H^k hampiran bagi songsangan Hessian, $g^k = \nabla f(x^k)$ dan $t_k > 0$ panjang langkah yang memenuhi syarat Armijo. Kita menubuhkan sifat penumpuan kuat untuk kaedah ini di bawah andaian longgar: sama ada $x^k \to \bar{x}$, sehinggakan $\nabla f(\bar{x}) = 0$; atau huj minim $f = \emptyset$, $\|x^k\| \to \infty$, dan $f(x^k) \downarrow \inf f$.

Katakunci Kaedah Kuasi-Newton, Fungsi Kuasicembung, Gelintaran Garis Armijo.

Abstract *To minimize a continuously differentiable quasiconvex function* $f : \Re^n \to \Re$, *Armijo's quasi-Newton method generates a sequence*

$$x^{k+1} = x^k - t_k H^k g^k,$$

where H^k *is an approximation for inverse Hessian,* $g^k = \nabla f(x^k)$ *and* $t_k > 0$ *stepsize which satisfies the Armijo's condition. We establish strong convergence properties of this method under mild assumption: either* $x^k \to \bar{x}$, *such that* $\nabla f(\bar{x}) = 0$; *or arg min* $f = \emptyset$, $\|x^k\| \to \infty$, *and* $f(x^k) \downarrow \inf f$.

Keywords *Quasi-Newton Method, Quasiconvex Function, Armijo Line Search*

1 Pengenalan

Untuk minimumkan suatu fungsi kuasicembung yang selanjar dan boleh beza $f : \Re^n \to \Re$, kaedah kuasi-Newton [2] dengan panjang langkah Armijo klasik [1] menjana suatu jujukan

$\{x^k\}$ dengan

$$x^{k+1} = x^k - t_k H^k g^k, \quad k = 0, 1, ... \tag{1}$$

dengan H^k hampiran bagi songsangan Hessian, $g^k = \nabla f(x^k)$ dan t_k panjang langkah yang memenuhi

$$t_k = \text{ huj maks } \{t : f(x^k - t_k H^k g^k) \leq f(x^k) - \alpha t {g^k}^T H^k g^k, \quad t = 2^{-i}, \quad i = 0, 1, ...\}, \tag{2}$$

dengan $\alpha \in (0, 1)$. Kita akan buktikan dalam Seksyen 2 keputusan penumpuan kuat berikut.

Teorem 1 Penumpuan Sejagat *Sama ada* $x^k \to \bar{x} \in \mathbf{X} := \{x : \nabla f(x) = 0\}$, *atau* $\hat{\mathbf{X}} :=$ *huj minim* $f = \emptyset$, $\|x^k\| \to \infty$, *dan* $f(x^k) \downarrow \inf f$.

Suatu keputusan penumpuan untuk kaedah penurunan tercuram Cauchy muncul dalam [5]. Kita lanjutkan bukti penumpuannya untuk kaedah kuasi-Newton.

2 Penumpuan Sejagat Kaedah Kuasi-Newton

Kita akan membuat andaian berikut.

Andaian 1 *Katakan* $\phi : \Re_+ \to \Re_+$ *adalah fungsi sedemikian hingga:*

(*A*1) $\exists \alpha \in (0, 1), \tau_\alpha > 0, \forall t \in (0, \tau_\alpha] : \phi(t) \leq \alpha t,$

(*A*2) $\exists \beta > 0, \tau_\beta \in (0, \infty), \forall t \in (0, \tau_\beta] \cap \Re : \phi(t) \geq \beta t^2,$

(*A*3) $\forall k, f(x^k - t_k H^k g^k) \leq f(x^k) - \phi(t) {g^k}^T H^k g^k \text{dan} 0 < t_k < \tau_\beta$ dalam (1)

(*A*4) $\exists \gamma > 1, \tau_\gamma > 0, \forall k : t_k \geq \tau_\gamma \text{atau}$
$[\exists t_k \in [t_k, \gamma t_k] : f(x^k - t_k H^k g^k) \geq f(x^k) - \phi(t_k) {g^k}^T H^k g^k].$

(*A*5) *nombor syarat* κ^k *dalam norma asli (nisbah* $\lambda^k_{maks} / \lambda^k_{\text{minim}}$, iaitu *hasilbahagi nilai eigen terbesar dengan terkecil) bagi* H^k *adalah terbatas atas untuk semua* k, *atau*
$\exists \xi_1, \xi_2 > 0, \lambda^k_{\text{maks}} \leq \xi_1$ *dan* $\lambda^k_{\text{minim}} \geq \xi_2, \forall k$ *s.h.*
$\|H^k g^k\| \leq \lambda^k_{\text{maks}} \|gk\| \leq \xi_1 \|g^k\|$ *dan*
${g^k}^T H^k g^k \geq \lambda^k_{\text{minim}} \|gk\|^2 \geq \xi_2 \|g^k\|^2, \forall k.$

Perhatikan bahawa (A1)-(A4) adalah syarat am Armijo (2), dan (2) setara dengan

$$\phi(t) = \alpha t, \quad \beta = \alpha, \quad \gamma = 2, \quad \tau_\alpha = \tau_\beta = \tau_\gamma = 1.$$

Seperti dalam [4], kita mula dengan mempertimbangkan syarat

$$f(x^k) \geq f(\tilde{x}), \quad \text{untuk sesuatu } \tilde{x} \text{ yang tetap dan semua } k, \tag{3}$$

adalah benar jika $\hat{\mathbf{X}} \neq \emptyset$ atau $\tilde{x}$ ialah titik gugusan bagi $\{x^k\}$.

Lema 1 *Jika (3) benar, maka*

$$\sum_{k=0}^{\infty} t_k^2 \|g^k\|^2 \leq [f(x^0) - f(\tilde{x})] / (\xi_2 \beta),$$

dan justeru itu,

$$\sum_{k=0}^{\infty} t_k \|g^k\| < \infty. \tag{4}$$

Lebih-lebih lagi, $x^k \to \bar{x}$ untuk sesuatu $\bar{x}$.

Bukti Daripada (A2)-(A3),

$$\beta t_k^2 g^{k^T} H^k g^k \leq \phi(t) g^{k^T} H^k g^k \leq f(x^k) - f(x^{k+1});$$

menambah semua ketaksamaan di atas dan juga daripada (3), kita perolehi

$$\sum_{k=0}^{\infty} t_k^2 g^{k^T} H^k g^k \leq [f(x^0) - f(\tilde{x})]/\beta;$$

daripada (A5), kita ada

$$\sum_{k=0}^{\infty} t_k^2 \|g^k\|^2 \leq [f(x^0) - f(\tilde{x})]/(\xi_2 \beta).$$

Adalah mudah untuk menunjukkan bahawa

$$\operatorname*{had}_{n\to\infty} \sum_{k=0}^{\infty} t_k \|g^k\| < \operatorname*{had}_{n\to\infty} n\, [f(x^0) - f(\tilde{x}) + 1]/(\xi_2 \beta),$$

yang memberikan (4). Seterusnya, oleh kerana $x^k - x^{k+1} = t_k H^k g^k$, kita deduksikan yang

$$\begin{aligned} \|\tilde{x} - x^{k+1}\| &\leq \|\tilde{x} - x^k\| + \|x^k - x^{k+1}\| \\ &= \|\tilde{x} - x^k\| + t_k \|H^k g^k\| \\ &\leq \|\tilde{x} - x^k\| + \xi_1 t_k \|g^k\|, \quad \text{[daripada (A5)]} \end{aligned}$$

supaya

$$\|\tilde{x} - x^l\| \leq \|\tilde{x} - x^k\| + \xi_1 \sum_{j=k}^{l} t_j \|g^j\| < \infty,$$

jika $l > k$. Justeru itu, $\{x^k\}$ adalah terbatas dan mempunyai titik gugusan $\bar{x}$. Jadi kita boleh setkan $\tilde{x} = \bar{x}$ yang dideduksikan dari (4) untuk sebarang $\epsilon > 0$, wujud k sedemikian hingga $\|\bar{x} - x^k\| \leq \epsilon/2$ dan

$$\sum_{j=k}^{l} t_j \|g^j\| < \epsilon/(2\xi_1);$$

oleh yang demikian, $\|\bar{x} - x^l\| < \epsilon$ untuk semua $l > k$, iaitu, $x^k \to \bar{x}$.

Lema 2 *Jika $\bar{x}$ ialah titik gugusan bagi $\{x^k\}$, maka $\bar{x} \in \hat{\mathbf{X}}$, iaitu, $\nabla f(\bar{x}) = 0$.*

Bukti Andaikan untuk sesuatu $k \in \mathbf{K}$, set indeks, $x^k \xrightarrow{K} \bar{x}$, tetapi $\bar{g} := \nabla f(\bar{x}) \neq 0$. Maka, $t_k \xrightarrow{K} 0$ daripada (A2)-(A3) dan (A5)

$$0 \leq \beta\xi_2 t_k^2 \|g^k\|^2 \leq \beta t_k^2 g^{k^T} H^k g^k \leq f(x^k) - f(x^{k+1}) \xrightarrow{K} 0,$$

dengan $g^k \xrightarrow{K} \bar{g} \neq 0$ dan dari keselanjaran $f(x^k) \downarrow f(\bar{x})$. Jadi, untuk semua nilai besar $k \in \mathbf{K}$, kita dapat ketaksamaan

$$f(x^k - t_k H^k g^k) - f(x^k) \geq -\phi(t_k) g^{k^T} H^k g^k \geq -\alpha t_k g^{k^T} H^k g^k, \tag{5}$$

daripada (A4) dan (A1), di mana sebelah kiri (5) sama dengan $-\tilde{t}_k \langle H^k g^k, \nabla f(x^k - t_k H^k g^k)\rangle$ untuk sesuatu $\tilde{t}_k \in [0, t_k]$ dengan menggunakan teorem nilai min, dan juga daripada (A4), $0 \leq \tilde{t}_k \leq \gamma t_k \xrightarrow{K} 0$.

Justeru itu dengan membahagi (5) dengan $\tilde{t}_k$ dan biarkan $k \xrightarrow{K} \infty$, kita dapat

$$g^{k^T} H^k g^k \leq \alpha g^{k^T} H^k g^k$$

iaitu suatu percanggahan dengan $\alpha < 1$ [daripada (A1)].

Sekarang, kita boleh buktikan Teorem 1 di bawah Andaian 1.

Bukti bagi Teorem 1. Jika (3) benar, i.i. $\hat{\mathbf{X}} \neq \emptyset$ atau $\{x^k\}$ mempunyai titik gugusan, maka kita akan dapat $x^k \to \bar{x} \in \hat{\mathbf{X}}$ (dari Lema 1). Jika $\|x^k\| \to \infty$, maka $\{x^k\}$ mempunyai titik gugusan. Jika $\underset{k\to\infty}{\text{had}} f(x^k) > \inf f$, maka (3) benar.

3 Contoh Berangka dan Perbincangan

Terdapat dua kelebihan dalam teorem 1 di atas ini. Pertama, syarat gelintaran garis Armijo tidak memerlukan interpolasi fungsi matlamat (kadang-kadang proses ekstrapolasi yang lebih rumit) yang biasanya diperlu semasa menggunakan syarat gelintaran garis Wolfe [8,9] dan Goldstein [3] dalam mencari panjang langkah. Kedua, jika syarat (A5) dalam pembuktian dipenuhi, iaitu κ^k terbatas, maka syarat ujian sudut yang diperlukan dalam kebanyakan pembuktian penumpuan algorithma juga akan dipatuhi (Sebagai contoh, lihat pembuktian penumpuan mengguna syarat gelintaran garis Wolfe dan syarat gelintaran garis Goldstein). Perhatikan bahawa kuasi kecembungan [7] bagi f adalah perlu untuk Lema 2, dan seterusnya Teorem 1. Sebagai contoh, pertimbangkan 3 masalah berikut:

$$f(x_1, x_2) = exp(x_1) - x_2^2,\ x^0 = (0, 0)^T, \tag{6}$$

$$f(x_1, x_2) = (x_1^2 + x_2^2 - 1)^2 + (x_1 + x_2 - 1)^2,\ x^0 = (-1, 1)^T, \tag{7}$$

$$f(x) = x^6 - 15x^4 + 27x^2 + 250(\text{fungsi Goldstein}),\ \forall x \in (1, \infty),\ x^0 = 1.1 \tag{8}$$

Ketiga-tiga masalah (6)–(8) diselesaikan dengan kaedah kuasi-Newton Armijo menggunakan rumus kemaskini songsangan Hessian BFGS terusik bernombor syarat terbatas dalam norma surihan yang diberi oleh Malik, Mansor dan Leong [6]. α dalam (2) disetkan sebagai 0.9. Keputusan masalah (6) dipamerkan dalam Jadual 1 dan masalah (7) serta (8) dalam Jadual 2.

Jadual 1: Keputusan Kaedah kuasi-Newton BFGS Terusik dengan Panjang Langkah Armijo untuk Masalah (6)

Lelaran, k	$x^k = (x_1^k, x_2^k)^T$	$f(x^k)$
0	$(0.0, 0.0)^T$	1.000000
10	$(-2.529007, 0.0)^T$	7.973816×10^{-2}
50	$(-3.994234, 0.0)^T$	1.877060×10^{-2}
100	$(-4.640620, 0.0)^T$	9.651681×10^{-3}
500	$(-6.223374, 0.0)^T$	1.982544×10^{-3}
1000	$(-6.912495, 0.0)^T$	9.952718×10^{-4}
1500	$(-7.316516, 0.0)^T$	6.644733×10^{-4}
2000	$(-7.603448, 0.0)^T$	4.987289×10^{-4}
3000	$(-8.008131, 0.0)^T$	3.327460×10^{-4}
4000	$(-8.295410, 0.0)^T$	2.496601×10^{-4}
5000	$(-8.518306, 0.0)^T$	1.997776×10^{-4}

Jadual 2: Keputusan Kaedah kuasi-Newton BFGS Terusik dengan Panjang Langkah Armijo untuk Masalah (7) dan (8)

Masalah	Lelaran, k	x^0	x^k	$x^{\star}_{\text{tepat}}$
(3.2)	59	(-1, 1)	$(-3.454580 \times 10^{-5}, 1.0)$	(0, 1)
(3.3)	456	1.1	2.999962	3

Perhatikan fungsi matlamat f dalam masalah (6) adalah cembung tetapi bukan kuasi-cembung. Maka kaedah di atas menjana

$$x^k = (x_1^k, 0)^T, \text{ dengan } x_1^k \downarrow -\infty, f(x^k) \downarrow 0 \text{ dan } \nabla f(x^k) \to (0,0)^T,$$

iaitu suatu minimum setempat manakala inf $f = -\infty$.

Masalah (7) pula mempunyai fungsi matlamat yang cembung mutlak di mana minimum setempat juga merupakan minimum sejagat manakala masalah (8) mempunyai fungsi matlamat yang kuasi-cembung dalam selang $(1, \infty)$ dan $(-\infty, -1)$ (walaupun bukan kuasi-cembung untuk sebarang x). Masalah (8) mempunyai 3 minimum iaitu minimum sejagat di $x = 3$ dan -3 dan minimum setempat di $x = 0$. Kaedah diberi di atas berjaya menumpu kepada salah satu minimum sejagat iaitu di $x^{\star} = 3$ apabila $x^0 = 1.1$ diambil. Ini bermakna kalau f bukan kuasicembung, penumpuan sejagat tidak terjamin. Kebanyakan algoritma kuasi-Newton yang wujud tidak menunjukkan yang (A5) benar, iaitu κ^k tidak terbatas. Maka dengan bukti penumpuan di atas, algoritma boleh diubahsuaikan untuk rumus kemaskini H^k supaya κ^k terbatas.

Rujukan

[1] L. Armijo, *Minimization of Functions Having Lipschitz Continuous Partial Derivatives*, Pacific Journal of Mathematics, **Vol. 16**(1966), 1-3.

[2] W. C. Davidon, *Variable Metric Method for Minimization*, AEC Res. And Dev. Report ANL-5990 (Revised),1959.

[3] A. A. Goldstein, *On Steepest Descent*, SIAM J. Control, **Vol. 3**(1965), 147-151.

[4] K. C. Kiwiel, *An Aggregate Subgradient Method for Nonsmooth Convex Minimization*, Mathematical Programming, **Vol. 27**(1983), 320-341.

[5] K. C. Kiwiel and K. Murty, *Convergence of the Steepest Descent Method for Minimizing Quasiconvex Functions*, Journal of Optimization Theory and Applications, **Vol. 89, No.1**(1996), 221-226.

[6] Malik Hj.Abu Hassan, Mansor Monsi and W. J. Leong, *Analisis Pengusikan ke atas Rumus Kemaskini Quasi-Newton Pangkat-dua yang Simetri dan Tentu Positif*, MATEMATIKA, **Jilid 14**(1998), 7-16.

[7] O. L. Mangasarian, *Nonlinear Programming*, Mc-Graw-Hill, New York, 1969.

[8] P. Wolfe, *Convergence Conditions for Ascent Methods*, SIAM Review, **Vol. 11**(1969), 226-235.

[9] P. Wolfe, *Convergence Conditions for Ascent Methods II: Some Corrections* SIAM Review, **Vol. 11**(1969), 226-235., SIAM Review, **Vol. 13**(1971), 185-188.