Sains & Matematika
ISSN 2302-7290
Vol. 2 No. 2, April 2014

Konvergensi Global Metode *Spectral Conjugate Descent* yang Baru Menggunakan Pencarian Garis Armijo yang Termodifikasi

Global Convergence of the New Spectral Conjugate Descent Method with the Modified Armijo-type Line Search

Dahliatul Hasanah
Fakultas MIPA Universitas Negeri Malang
Jalan Semarang 5, Malang 65145

ABSTRAK

Metode *nonlinear spectral conjugate descent* yang baru merupakan gabungan dari metode *conjugate descent* dan metode *spectral conjugate gradient*. Metode ini menghasilkan d_k yang berorientasi menurun (*descent direction*). Dengan menggunakan pencarian garis *strong Wolfe Rule*, metode ini telah dibuktikan konvergen global. Tulisan ini menganalisis tentang konvergensi global metode tersebut dengan menggunakan pencarian garis Armijo yang termodifikasi. Analisis ini menunjukkan bahwa menggunakan pencarian garis Armijo yang termodifikasi, metode *nonlinear spectral conjugate descent* yang baru konvergen global.

Kata kunci: metode *nonlinear spectral conjugate descent*, konvergen global, pencarian garis Armijo yang termodifikasi

ABSTRACT

The new nonlinear spectral conjugate descent method is constructed by combining the spectral conjugate gradient method and conjugate descent method. This method generates descent direction d_k. Under the strong Wolfe rule this method is globally convergent. This paper analyzed global convergence of this method using the modified Armijo line search. The result shows that the new nonlinear spectral conjugate descent method is globally convergent under the modified Armijo line search.

Key words: *nonlinear spectral conjugate descent method, global convergence, modified Armijo line search*

PENDAHULUAN

Beberapa tahun terakhir banyak bermunculan metode-metode baru dan yang merupakan modifikasi dari metode-metode lama untuk menyelesaikan masalah optimasi tanpa kendala. Hal ini dikarenakan masalah optimasi tanpa kendala banyak digunakan di berbagai bidang seperti transportasi, eksplorasi bahan bakar minyak, *aerospace,* dan sebagainya (Liu *et al*., 2012). Semakin besar skala atau variabel yang terlibat dalam masalah optimasi tanpa kendala, maka semakin besar pula kebutuhan akan adanya suatu metode yang mampu menyelesaikan masalah optimasi tanpa kendala dalam skala besar.

Masalah optimasi nonlinier tanpa kendala merupakan pencarian nilai minimum dari fungsi bernilai $f(x)$ yaitu

$$\min f(x),\ x \in R^n \quad (1.1)$$

f adalah fungsi nonlinier yang kontinu dan terdiferensialkan dengan gradien $g(x) = \nabla f(x)$. Masalah tersebut dapat diselesaikan secara numerik, yaitu dengan cara iteratif. Iterasi yang dijalankan dinotasikan dengan x_k, k=1,2,.... Setiap iterasi diberikan oleh

$$x_{k+1} = x_k + \alpha_k d_k \quad (1.2)$$

Alamat Korespondensi:
e-mail: iyut_dlia@yahoo.com

α_k adalah panjang langkah (*step length*) yang ditentukan oleh suatu pencarian garis, d_k adalah arah pencarian dan x_1 adalah titik awal.

Salah satu metode yang terkenal untuk menyelesaikan masalah ini adalah metode *conjugate gradient*. Metode ini membutuhkan memori yang sedikit dan mempunyai sifat kekonvergenan lokal maupun global. Namun arah pencarian yang dihasilkan dari metode ini tidak dijamin arah yang menurun. Di lain pihak, metode *spectral gradient* menggunakan hanya gradien sebagai arahnya di setiap pencarian garisnya. Metode ini mempunyai performa yang lebih bagus daripada metode *conjugate gradient* dalam berbagai masalah (Birgin & Martinez, 2001).

Birgin & Martinez (2001) mengembangkan metode *spectral conjugate gradient* yang merupakan gabungan dari metode *spectral gradient* dan metode *conjugate gradient*. Metode ini terbukti sangat efektif untuk menyelesaikan masalah optimasi tanpa kendala nonlinier. Akan tetapi metode ini menghasilkan arah yang belum tentu menurun.

Metode lain yang terkenal adalah metode *conjugate descent* dengan menggunakan pencarian garis *strong Wolfe*. Metode ini dikembangkan oleh Fletcher (1987). Salah satu sifat penting dari metode ini adalah arah pencarian yang dihasilkan merupakan arah pencarian yang menurun (*descent direction*).

Termotivasi oleh Birgin & Martinez, Liu *et al.* (2012) mengembangkan metode yang menggabungkan metode *conjugate descent* dan metode *spectral conjugate gradient* yang disebut dengan metode *spectral conjugate descent*. Metode ini menghasilkan arah pencarian yang menurun. Liu *et al.* (2012) telah menunjukkan bahwa metode ini konvergen global menggunakan pencarian garis *strong Wolfe*. Di lain pihak, untuk pencarian garis yang lain belum ditunjukkan kekonvergenan globalnya.

Dalam makalah ini dibahas kekonvergenan global metode *spectral conjugate descent* dengan menggunakan pencarian garis Armijo yang termodifikasi. Hal ini termotivasi oleh Kim *et al.* (2008) yang menunjukkan bahwa pencarian garis Armijo yang termodifikasi lebih efisien dalam mencari nilai minimum lokal walaupun membutuhkan perhitungan yang lebih banyak dibandingkan pencarian garis lainnya.

METODE PENELITIAN

Kekonvergenan metode *spectral conjugate descent* dengan menggunakan pencarian garis Armijo yang termodifikasi dianalisis dengan menggunakan pendekatan secara numerik.

Metode *Conjugate Gradient*

Pada umumnya metode *conjugate gradient* dideskripsikan oleh iterasi berikut ini:

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k$$

$\alpha k > 0$ adalah panjang langkah (*steplength*) yang ditentukan oleh suatu pencarian garis, d_k didefinisikan oleh:

$$\boldsymbol{d}_k = \begin{cases} -\boldsymbol{g}_k & \text{if } k = 1 \\ -\boldsymbol{g}_k + \beta_k \boldsymbol{d}_{k-1}, & \text{if } k > 1 \end{cases} \quad \text{..........................(2.1)}$$

dengan $g_k = g(x_k)$ adalah gradien dari f di titik x_k, β_k adalah parameter yang jika digunakan untuk meminimumkan fungsi kuadratik yang konveks ketat, maka arah pencarian d_k dan d_{k-1} merupakan konjugat berdasarkan Hessian dari fungsi objektif (Zhang *et al.*, 2006). Dai *et al.* (2004) menyatakan bahwa jika fungsi objektifnya adalah fungsi kuadratik yang konveks ketat dan menggunakan pencarian garis eksak, maka metode *conjugate gradient* menghasilkan arah-arah pencarian yang konjugat satu dengan lainnya. Terdapat banyak formula untuk β_k yang terkenal, di antaranya adalah Fletcher-Reeves (FR), Polak-Ribiere-Polyak (PRP) dan *conjugate descent* (CD). Formula-formula tersebut diberikan berikut ini:

$$\beta_k^{FR} = \frac{\|\boldsymbol{g}_k\|^2}{\|\boldsymbol{g}_{k-1}\|^2}, \quad \text{..(2.2)}$$

$$\beta_k^{PRP} = \frac{{\boldsymbol{g}_k}^T \boldsymbol{y}_{k-1}}{\|\boldsymbol{g}_{k-1}\|^2} \quad \text{..(2.3)}$$

dan

$$\beta_k^{CD} = -\frac{\|\boldsymbol{g}_k\|^2}{{\boldsymbol{d}_{k-1}}^T \boldsymbol{g}_{k-1}}. \quad \text{..(2.4)}$$

dengan $\boldsymbol{y}_{k-1} = \boldsymbol{g}_k - \boldsymbol{g}_{k-1}$ dan $\|.\|$ menyatakan Euclidean norm.

Metode *Spectral Conjugate Gradient*

Metode *spectral gradient* awalnya dikembangkan oleh Barzilai dan Borwein pada tahun 1988, kemudian Raydan mengembangkan metode ini lebih jauh untuk menyelesaikan masalah optimasi tanpa kendala dalam skala besar. Birgin & Martinez (2001) mengembangkan tiga macam metode *spectral conjugate gradient* yang merupakan gabungan dari metode *spectral gradient* dan metode *conjugate gradient*. Arah pencarian diberikan oleh:

$$\boldsymbol{d}_k = -\theta_k \boldsymbol{g}_k + \beta_k \boldsymbol{s}_{k-1}$$

$\boldsymbol{s}_{k-1} = \boldsymbol{x}_k - \boldsymbol{x}_{k-1}$ dan parameter β_k ditentukan oleh:

$$\beta_k = \frac{(\theta_k \boldsymbol{y}_{k-1} - \boldsymbol{s}_{k-1})^T \boldsymbol{g}_k}{\boldsymbol{s}_{k-1}{}^T \boldsymbol{y}_{k-1}},$$

$$\beta_k = \frac{\theta_k \boldsymbol{y}_{k-1}{}^T \boldsymbol{g}_k}{\alpha_{k-1}\theta_{k-1}\boldsymbol{g}_{k-1}{}^T \boldsymbol{g}_{k-1}}$$

dan

$$\beta_k = \frac{\theta_k \boldsymbol{g}_k{}^T \boldsymbol{g}_k}{\alpha_{k-1}\theta_{k-1}\boldsymbol{g}_{k-1}{}^T \boldsymbol{g}_{k-1}}.$$

sedangkan θ_k adalah *spectral gradient* yang dapat ditentukan oleh:

$$\theta_k = \frac{\boldsymbol{s}_{k-1}{}^T \boldsymbol{s}_{k-1}}{\boldsymbol{s}_{k-1}{}^T \boldsymbol{y}_{k-1}}, \quad \text{.......} (2.5)$$

$\boldsymbol{y}_{k-1} = \boldsymbol{g}_k - \boldsymbol{g}_{k-1}$. Secara numerik metode ini sangat efisien, akan tetapi arah pencarian yang dihasilkan belum tentu arah yang menurun (*descent direction*).

Berdasarkan metode *spectral conjugate gradient*, Zhang *et al.* (2006) memodifikasi metode Flethcer-Reeves sehingga arah pencarian yang dihasilkan adalah arah pencarian yang menurun. Metode Fletcher-Reeves yang termodifikasi diberikan oleh:

$$\boldsymbol{d}_k = \begin{cases} -\boldsymbol{g}_k & \text{if } k = 1, \\ -\theta_k \boldsymbol{g}_k + \beta_k^{FR} \boldsymbol{d}_{k-1}, & \text{if } k > 1, \end{cases}$$

β_k^{FR} diberikan oleh (2.2).

$$\theta_k = \frac{\boldsymbol{d}_{k-1}{}^T \boldsymbol{y}_{k-1}}{\boldsymbol{g}_{k-1}{}^T \boldsymbol{g}_{k-1}}.$$

Jika pencarian garis yang digunakan adalah pencarian garis eksak, maka metode ini kembali ke metode Fletcer-Reeves yang standar.

Metode *Spectral Conjugate Descent*

Untuk mendapatkan arah pencarian yang menurun, Liu *et al.* (2012) mengkombinasi metode *spectral conjugate gradient* dan metode *conjugate descent*. Arah pencarian dari metode *conjugate gradient descent* diberikan oleh:

$$\boldsymbol{d}_k = \begin{cases} -\boldsymbol{g}_k & \text{if } k = 1, \\ -\theta_k \boldsymbol{g}_k + \beta_k \boldsymbol{d}_{k-1}, & \text{if } k > 1, \end{cases} \quad \text{.......} (2.6)$$

β_k adalah

$$\beta_k = \begin{cases} \beta_k^{CD}, \text{ if } \boldsymbol{g}_k{}^T \boldsymbol{d}_{k-1} \le 0 \\ 0, \quad \text{else} \end{cases} \quad \text{.......} (2.7)$$

$$\theta_k = 1 - \frac{\boldsymbol{g}_k{}^T \boldsymbol{d}_{k-1}}{\boldsymbol{g}_{k-1}{}^T \boldsymbol{d}_{k-1}}. \quad \text{.......} (2.8)$$

Dengan menggunakan pencarian garis Armijo yang termodifikasi akan ditunjukkan bahwa metode ini konvergen global serta menghasilkan arah pencarian yang menurun.

Pencarian Garis Armijo Yang Termodifikasi

Pencarian garis merupakan salah satu bagian yang penting dalam metode optimisasi. Terdapat dua jenis pencarian garis, yakni pencarian garis eksak dan tidak eksak. Namun dalam praktiknya, pencarian garis yang tidak eksak lebih dipilih daripada pencarian garis yang eksak karena sifat pencarian garis eksak yang terlalu mahal dan tidak efisien. Di lain pihak, bagi sebagian besar metode-metode optimisasi, sebagai contoh metode Newton dan metode Quasi-Newton, kekonvergenan metodenya tidak bergantung pada pencarian garis eksak yang digunakan.

Secara umum, pencarian garis tidak eksak mencari panjang langkah (*steplength*) yang sesuai sehingga fungsi objektif merupakan fungsi yang menurun di setiap iterasinya. Salah satu pencarian garis yang tidak eksak adalah pencarian garis Armijo yang dideskripsikan sebagai berikut: diberikan $\beta \in (0,1), \rho \in \left(0, \frac{1}{2}\right), \tau > 0$, terdapat bilangan bulat nonnegatif paling kecil m_k sedemikian sehingga

$$f(\boldsymbol{x}_k) - f(\boldsymbol{x}_k + \beta^m \tau \boldsymbol{d}_k) \ge -\rho \beta^m \tau \boldsymbol{g}_k{}^T \boldsymbol{d}_k.$$

Modifikasi pencarian garis Armijo berdasarkan fungsi

$$f_k = f + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_k)^T B_k (\boldsymbol{x} - \boldsymbol{x}_k)$$

B_k adalah matriks simetris dan definit positif yang dikembangkan oleh Wei *et al.* (2000). Untuk memudahkan penggunaan, modifikasi ini dapat digunakan dengan mengubah B_k menjadi μI untuk $\mu > 0$. Untuk metode *conjugate descent*, pencarian garis Armijo yang termodifikasi dideskripsikan sebagai berikut: diberikan $c_0 \in (0, \frac{1}{2})$, $\rho \in (0,1)$, $c \in (0,1)$, $\mu > 0$ dan $\varphi_k > 0$, misal didefinisikan $\alpha_k = \rho^{j_k} \varphi_k$, pencarian garis Armijo yang termodifikasi adalah mencari α_k di mana j_k adalah bilangan bulat nonnegatif paling kecil sehingga

$$f(\boldsymbol{x}_k + \rho^j \varphi_k \boldsymbol{d}_k) \le f(\boldsymbol{x}_k) + c_0 \rho^j \varphi_k \boldsymbol{g}_k^T \boldsymbol{d}_k - \frac{\mu}{2}(\rho^j \varphi_k)^2 \|\boldsymbol{d}_k\|^2$$

dan

$$\boldsymbol{g}(\boldsymbol{x}_k + \rho^j \varphi_k \boldsymbol{d}_k)^T \boldsymbol{d}_k \le -(1-c)\boldsymbol{g}_k^T \boldsymbol{d}_k. \quad \text{.......} (2.10)$$

Misal η>0 dan ε>0 bilangan riil yang sangat kecil, maka φk yang paling sesuai ditentukan oleh

$$\varphi_k = \begin{cases} \frac{-g_k^T d_k}{d_k^T z_k}, \text{ if } \frac{-g_k^T d_k}{d_k^T z_k} \geq \eta, \\ 1, \qquad \text{else,} \end{cases} \qquad (2.11)$$

dengan $z_k = \frac{g(x_k + \epsilon d_k) - g_k}{\epsilon}$.

HASIL DAN PEMBAHASAN

Berikut adalah algoritma metode *spectral conjugate descent* dengan menggunakan pencarian garis Armijo yang termodifikasi.

Algoritma SCD-MA

Langkah 1 : diberikan $c_0 \in (0, \frac{1}{2})$, $\rho \in (0,1)$, $c \in (0,1)$, $\mu > 0$. Pilih titik awal $x_1 \in \boldsymbol{R}^n, \varepsilon > 0$. Set $\boldsymbol{d}_1 = -\boldsymbol{g}_1, k = 1$. Jika $\|\boldsymbol{g}_1\| \leq \varepsilon$, maka stop.
Langkah 2 : menentukan φ_k pada (2.11) dan cari α_k yang memenuhi (2.9) dan (2.10).
Langkah 3 : misal $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k$ dan $\boldsymbol{g}_{k+1} = \boldsymbol{g}(\boldsymbol{x}_{k+1})$. Jika $\boldsymbol{g}_{k+1} = 0$ maka stop.
Langkah 4 : tentukan β_k oleh (2.7) dan $\varphi_k$oleh (2.8), maka tentukan $\boldsymbol{d}_k$ oleh (2.6).
Langkah 5 : set k=k+1, kembali ke langkah 2.

Kekonvergenan global Algoritma SCD-MA akan dibuktikan dengan menggunakan asumsi-asumsi berikut.

Asumsi SCD-MA

1. Level set $\Omega = \{\boldsymbol{x} \in \boldsymbol{R}^n : f(\boldsymbol{x}) \leq f(\boldsymbol{x}_1)\}$ terbatas; x_1 adalah titik awal.
2. Gradien dari fungsi objektif memenuhi kondisi Lipschitz, yaitu terdapat $L > 0$ sedemikian sehingga untuk semua $\boldsymbol{x}, \boldsymbol{y} \in \Omega$ berlaku

$$\|\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{y})\| \leq L\|\boldsymbol{x} - \boldsymbol{y}\|. \qquad (3.1)$$

Berikut adalah teorema yang menunjukkan bahwa Algoritma SCD-MA menghasilkan arah pencarian yang menurun.

Teorema 3.1. Misal $\{g_k\}$ dan $\{\boldsymbol{d}_k\}$ dihasilkan oleh Algoritma SCD-MA, maka

$$\boldsymbol{g}_k{}^T \boldsymbol{d}_k \leq -\|\boldsymbol{g}_k\|^2. \qquad (3.2)$$

Bukti. Dengan menggunakan induksi, kita akan tunjukkan bahwa kesimpulan benar. Untuk k=1 berlaku d_1=g_1, sehingga berlaku $\boldsymbol{g}_1{}^T \boldsymbol{d}_1 = -\|\boldsymbol{g}_1\|^2$. Dengan demikian, (3.2) benar untuk k=1. Kita asumsikan benar untuk k-1 dan $\boldsymbol{g}_k \neq 0$, yaitu $\boldsymbol{g}_{k-1}{}^T \boldsymbol{d}_{k-1} \leq -\|\boldsymbol{g}_{k-1}\|^2$. Jika $\boldsymbol{g}_k{}^T \boldsymbol{d}_{k-1} \leq 0$, maka $\beta_k = \beta_k^{CD}$. Dari (2.4), (2.6) dan (2.8) didapatkan

$$\boldsymbol{g}_k{}^T \boldsymbol{d}_k = -\left(1 - \frac{\boldsymbol{g}_k{}^T \boldsymbol{d}_{k-1}}{\boldsymbol{g}_{k-1}{}^T \boldsymbol{d}_{k-1}}\right)\|\boldsymbol{g}_k\|^2 - \frac{\boldsymbol{d}_{k-1}{}^T \boldsymbol{g}_k}{\boldsymbol{d}_{k-1}{}^T \boldsymbol{g}_{k-1}}\|\boldsymbol{g}_k\|^2$$

$$= -\|\boldsymbol{g}_k\|^2.$$

Jika $\boldsymbol{g}_k{}^T \boldsymbol{d}_{k-1} > 0$, maka β_k=0. Dengan demikian dari (2.6), (2.8) dan asumsi bahwa didapatkan $\boldsymbol{g}_{k-1}{}^T \boldsymbol{d}_{k-1} < 0$

$$\boldsymbol{g}_k{}^T \boldsymbol{d}_k = -\left(1 - \frac{\boldsymbol{g}_k{}^T \boldsymbol{d}_{k-1}}{\boldsymbol{g}_{k-1}{}^T \boldsymbol{d}_{k-1}}\right)\|\boldsymbol{g}_k\|^2$$

$$= -\|\boldsymbol{g}_k\|^2 + \frac{\boldsymbol{g}_k{}^T \boldsymbol{d}_{k-1}}{\boldsymbol{g}_{k-1}{}^T \boldsymbol{d}_{k-1}}\|\boldsymbol{g}_k\|^2 < -\|\boldsymbol{g}_k\|^2.$$

Jadi kesimpulan benar untuk k. Dengan demikian terbukti bahwa metode *spectral conjugate descent* menghasilkan arah pencarian yang menurun, yaitu $\boldsymbol{g}_k{}^T \boldsymbol{d}_k \leq -\|\boldsymbol{g}_k\|^2$ untuk setiap iterasinya.

Dengan menggunakan pencarian garis Armijo yang termodifikasi, $\{f(x_k)\}$ adalah barisan yang menurun. Hal ini mengakibatkan $\{x_k\}$ yang dihasilkan oleh Algoritma SCD-MA termuat di Ω. Akibatnya terdapat suatu konstan f^* sedemikian sehingga

$$\lim_{k\to\infty} f(\boldsymbol{x}_k) = f^* \qquad (3.3)$$

Asumsi SCD-MA juga mengakibatkan terdapat suatu konstan M>0 sedemikian sehingga untuk semua k berlaku

$$\|\boldsymbol{g}_k\| \leq M. \qquad (3.4)$$

Lemma 3.2. Misalkan Asumsi SCD-MA berlaku. Maka

$$\lim_{k\to\infty} \alpha_k \|\boldsymbol{d}_k\| = 0 \text{ dan } \lim_{k\to\infty} -\alpha_k \boldsymbol{g}_k{}^T \boldsymbol{d}_k = 0 \qquad (3.5)$$

Bukti. Dari (3.3) kita dapatkan

$$\sum_{k=1}^{\infty}\left(f(\boldsymbol{x}_k) - f(\boldsymbol{x}_{k+1})\right) = \lim_{N\to\infty}\sum_{k=1}^{N}\left(f(\boldsymbol{x}_k) - f(\boldsymbol{x}_{k+1})\right)$$

$$= \lim_{N\to\infty}\left(f(\boldsymbol{x}_1) - f(\boldsymbol{x}_{N+1})\right)$$

$$= f(\boldsymbol{x}_1) - f^*.$$

Dengan demikian $\sum_{k=1}^{\infty}\left(f(\boldsymbol{x}_k) - f(\boldsymbol{x}_{k+1})\right) < \infty$., sedangkan dari (2.9) didapatkan

$$\sum_{k=1}^{\infty}\left(-c_0\alpha_k \boldsymbol{g}_k^T \boldsymbol{d}_k + \frac{\mu}{2}(\alpha_k)^2\|\boldsymbol{d}_k\|^2\right)$$

$$\leq \sum_{k=1}^{\infty}\left(f(\boldsymbol{x}_k) - f(\boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k)\right) < \infty.$$

Ini mengakibatkan $\sum_{k=1}^{\infty}(\alpha_k)^2\|\boldsymbol{d}_k\|^2 < \infty$ dan $\sum_{k=1}^{\infty} -\alpha_k \boldsymbol{g}_k^T \boldsymbol{d}_k < \infty$. Berdasarkan (3.2), kedua deret tersebut adalah deret nonnegatif sehingga deret tersebut adalah deret yang menurun. Dengan demikian $\lim_{k\to\infty} \alpha_k \|\boldsymbol{d}_k\| = 0$ dan $\lim_{k\to\infty} -\alpha_k \boldsymbol{g}_k{}^T \boldsymbol{d}_k = 0$..

Kedua sifat ini sangat penting untuk menunjukkan kekonvergenan global Algoritma SCD-MA.

Lemma 3.3. Misal Asumsi SCD-MA berlaku. Jika ada suatu konstan $\varepsilon>0$ sedemikian sehingga untuk semua $k>1$ berlaku $\|\boldsymbol{g}_k\| \geq \varepsilon$, maka terdapat suatu konstan $M_1>0$ sedemikian sehingga untuk semua k berlaku

$\|\boldsymbol{d}_k\| \leq M_1$. ....(3.6)

Bukti. Untuk k=1 kita peroleh d_1=g_1 sehingga $\|\boldsymbol{d}_1\| = \|\boldsymbol{g}_1\|$. Untuk $k>2$, jika $\boldsymbol{g}_k{}^T \boldsymbol{d}_{k-1} \leq 0$, maka dengan menggunakan (2.4), (2.6), (2.8) dan ketaksamaan segitiga didapatkan

$$\|\boldsymbol{d}_k\| \leq \|\boldsymbol{\theta}_k\| \|\boldsymbol{g}_k\| + \left\|\beta_k^{CD}\right\| \|\boldsymbol{d}_{k-1}\|$$
$$= \frac{\|\boldsymbol{d}_{k-1}\| \|\boldsymbol{g}_k - \boldsymbol{g}_{k-1}\| \|\boldsymbol{g}_k\|}{\|\boldsymbol{g}_{k-1}{}^T \boldsymbol{d}_{k-1}\|} + \frac{\|\boldsymbol{g}_k\|^2 \|\boldsymbol{d}_{k-1}\|}{\|\boldsymbol{g}_{k-1}{}^T \boldsymbol{d}_{k-1}\|}.$$

Karena g memenuhi kondisi Lipschitz dan $\|\boldsymbol{x}_k - \boldsymbol{x}_{k-1}\| = \alpha_{k-1}\|\boldsymbol{d}_{k-1}\|$, maka

$$\|\boldsymbol{d}_k\| \leq \frac{\|\boldsymbol{d}_{k-1}\| \|\boldsymbol{g}_k\|}{\|\boldsymbol{g}_{k-1}{}^T \boldsymbol{d}_{k-1}\|} (L\alpha_{k-1}\|\boldsymbol{d}_{k-1}\| + \|\boldsymbol{g}_k\|)$$
$$\leq \frac{\|\boldsymbol{g}_k\|}{\|\boldsymbol{g}_{k-1}\|} (L\alpha_{k-1}\|\boldsymbol{d}_{k-1}\| + \|\boldsymbol{g}_k\|)$$

Dari (3.4) dan $\|\boldsymbol{g}_k\| \geq \varepsilon$ untuk semua k, maka

$$\|\boldsymbol{d}_k\| \leq \frac{M}{\varepsilon} (L\alpha_{k-1}\|\boldsymbol{d}_{k-1}\| + M).$$
$$= \frac{M^2}{\varepsilon} + \frac{M}{\varepsilon} L\alpha_{k-1}\|\boldsymbol{d}_{k-1}\|.$$

Karena $\lim_{k\to\infty} \alpha_k \|\boldsymbol{d}_k\| = 0$, maka terdapat suatu konstan $q\in(0,1)$ dan bilangan bulat K sedemikian sehingga untuk semua $k>K$,

$$\frac{M}{\varepsilon} L\alpha_{k-1}\|\boldsymbol{d}_{k-1}\| \leq q.$$

Dengan demikian, untuk semua $k>K$ berlaku

$$\|\boldsymbol{d}_k\| \leq \frac{M^2}{\varepsilon} + q.$$

Kita definisikan

$$M_1 = \max\left\{\|\boldsymbol{d}_1\|, \|\boldsymbol{d}_2\|, \ldots, \|\boldsymbol{d}_K\|, \frac{M^2}{\varepsilon} + q\right\}$$

Oleh karena itu, dapat disimpulkan bahwa $\|\boldsymbol{d}_k\| \leq M_1$ untuk semua k.

Lemma 3.4. Misal asumsi SCD-MA berlaku, $\{x_k\}$ dan $\{g_k\}$ dihasilkan oleh algoritma SCD-MA. Oleh karena itu, terdapat suatu konstan $M_2>0$ sedemikian sehingga untuk semua k berlaku

$\alpha_k \geq M_2 \frac{\|\boldsymbol{g}_k\|^2}{\|\boldsymbol{d}_k\|^2}$ ....(3.7)

Bukti. (3.7) dapat dibuktikan pada beberapa kasus berikut.

Kasus 1. α_k=1. Berdasarkan (3.2) kita peroleh

$$\|\boldsymbol{g}_k\|^2 \leq -\boldsymbol{g}_k{}^T \boldsymbol{d}_k = |\boldsymbol{g}_k{}^T \boldsymbol{d}_k| \leq \|\boldsymbol{g}_k\| \|\boldsymbol{d}_k\|$$

sehingga $\|\boldsymbol{g}_k\| \leq \|\boldsymbol{d}_k\|$.

Hal ini mengakibatkan $\|\boldsymbol{g}_k\|^2 \leq \|\boldsymbol{d}_k\|^2 = \alpha_k \|\boldsymbol{d}_k\|^2$.

Dengan demikian $\alpha_k \geq \frac{\|\boldsymbol{g}_k\|^2}{\|\boldsymbol{d}_k\|^2}$ untuk $\alpha_k = 1$.

Kasus 2. α_k<1. Jika α_k<1 maka j_k-1 adalah bilangan nonnegatif. Berdasarkan definisi α_k dalam pencarian garis Armijo yang termodofikasi, $\frac{\alpha_k}{\rho} = \rho^{j_k - 1}\varphi_k$ tidak mungkin memenuhi (2.9) dan (2.10) secara bersamaan.

Jika $\frac{\alpha_k}{\rho}$ tidak memenuhi (2.9), maka

$$f\left(\boldsymbol{x}_k + \left(\frac{\alpha_k}{\rho}\right)\boldsymbol{d}_k\right) - f(\boldsymbol{x}_k) > c_0 \left(\frac{\alpha_k}{\rho}\right) \boldsymbol{g}_k^T \boldsymbol{d}_k$$
$$-\frac{\mu}{2}\left(\frac{\alpha_k}{\rho}\right)^2 \|\boldsymbol{d}_k\|^2.$$

Berdasarkan Teorema Nilai Rata-rata terhadap turunan, terdapat $\gamma_k \in (0,1)$ sedemikian sehingga

$$f\left(\boldsymbol{x}_k + \left(\frac{\alpha_k}{\rho}\right)\boldsymbol{d}_k\right) - f(\boldsymbol{x}_k) = \left(\frac{\alpha_k}{\rho}\right) \boldsymbol{d}_k^T \boldsymbol{g}\left(\boldsymbol{x}_k + \gamma_k \left(\frac{\alpha_k}{\rho}\right)\boldsymbol{d}_k\right).$$

Sehingga

$$\left(\frac{\alpha_k}{\rho}\right) \boldsymbol{d}_k^T \boldsymbol{g}\left(\boldsymbol{x}_k + \gamma_k \left(\frac{\alpha_k}{\rho}\right)\boldsymbol{d}_k\right) > c_0 \left(\frac{\alpha_k}{\rho}\right) \boldsymbol{g}_k^T \boldsymbol{d}_k$$
$$-\frac{\mu}{2}\left(\frac{\alpha_k}{\rho}\right)^2 \|\boldsymbol{d}_k\|^2.$$

Kedua ruas dibagi dengan $\frac{\alpha_k}{\rho}$ kemudian dikurangkan dengan $\boldsymbol{g}_k^T \boldsymbol{d}_k$ didapatkan

$$\boldsymbol{d}_k^T \left(\boldsymbol{g}\left(\boldsymbol{x}_k + \gamma_k \left(\frac{\alpha_k}{\rho}\right)\boldsymbol{d}_k\right) - \boldsymbol{g}_k\right) > -(1 - c_0)\boldsymbol{g}_k^T \boldsymbol{d}_k$$
$$-\frac{\mu}{2}\left(\frac{\alpha_k}{\rho}\right) \|\boldsymbol{d}_k\|^2.$$

Karena

$$\boldsymbol{d}_k^T\left(\boldsymbol{g}\left(\boldsymbol{x}_k+\gamma_k\left(\frac{\alpha_k}{\rho}\right)\boldsymbol{d}_k\right)-\boldsymbol{g}_k\right)\le\|\boldsymbol{d}_k\|\left\|\boldsymbol{g}\left(\boldsymbol{x}_k+\gamma_k\left(\frac{\alpha_k}{\rho}\right)\boldsymbol{d}_k\right)-\boldsymbol{g}_k\right\|,$$

maka

$$\boldsymbol{d}_k^T\left(\boldsymbol{g}\left(\boldsymbol{x}_k+\gamma_k\left(\frac{\alpha_k}{\rho}\right)\boldsymbol{d}_k\right)-\boldsymbol{g}_k\right)\le L\gamma_k\left(\frac{\alpha_k}{\rho}\right)\|\boldsymbol{d}_k\|^2.$$

Sehingga

$$-(1-c_0)\boldsymbol{g}_k^T\boldsymbol{d}_k-\frac{\mu}{2}\left(\frac{\alpha_k}{\rho}\right)\|\boldsymbol{d}_k\|^2<L\gamma_k\left(\frac{\alpha_k}{\rho}\right)\|\boldsymbol{d}_k\|^2.$$

Kita dapatkan

$$\alpha_k>\frac{(1-c_0)\rho(-\boldsymbol{g}_k^T\boldsymbol{d}_k)}{\left(L\gamma_k+\frac{\mu}{2}\right)\|\boldsymbol{d}_k\|^2}.$$

Dari (3.2) diperoleh

$$\alpha_k>\frac{(1-c_0)\rho\|\boldsymbol{g}_k\|^2}{\left(L\gamma_k+\frac{\mu}{2}\right)\|\boldsymbol{d}_k\|^2}.$$

Jika $\frac{\alpha_k}{\rho}$ tidak memenuhi (2.10), maka

$$\boldsymbol{g}\left(\boldsymbol{x}_k+\frac{\alpha_k}{\rho}\boldsymbol{d}_k\right)^T\boldsymbol{d}_k>-(1-c)\boldsymbol{g}_k^T\boldsymbol{d}_k.$$

Kedua ruas dikurangi oleh $\boldsymbol{g}_k^T\boldsymbol{d}_k$ maka didapatkan

$$\left(\boldsymbol{g}\left(\boldsymbol{x}_k+\frac{\alpha_k}{\rho}\boldsymbol{d}_k\right)-\boldsymbol{g}_k\right)^T\boldsymbol{d}_k>-(2-c)\boldsymbol{g}_k^T\boldsymbol{d}_k$$
$$\ge(2-c)\|\boldsymbol{g}_k\|^2.$$

Berdasarkan asumsi SCD-MA didapatkan

$$(2-c)\|\boldsymbol{g}_k\|^2<\left(\boldsymbol{g}\left(\boldsymbol{x}_k+\left(\frac{\alpha_k}{\rho}\right)\boldsymbol{d}_k\right)-\boldsymbol{g}_k\right)^T\boldsymbol{d}_k$$
$$\le\left\|\boldsymbol{g}\left(\boldsymbol{x}_k+\left(\frac{\alpha_k}{ñ}\right)\boldsymbol{d}_k\right)-\boldsymbol{g}_k\right\|\|\boldsymbol{d}_k\|$$
$$\le\left(\frac{\alpha_k}{\rho}\right)L\|\boldsymbol{d}_k\|^2.$$

Sehingga

$$\alpha_k>\frac{(2-c)\rho\|\boldsymbol{g}_k\|^2}{L\|\boldsymbol{d}_k\|^2}.$$

Didefinisikan

$$M_2=\min\left\{1,\frac{(1-c_0)\rho}{\left(L\gamma_k+\frac{\mu}{2}\right)},\frac{(2-c)\rho}{L}\right\},$$

maka didapatkan $\alpha_k\ge M_2\frac{\|\boldsymbol{g}_k\|^2}{\|\boldsymbol{d}_k\|^2}$. .

Teorema 3.5. Misalkan Asumsi SCD-MA berlaku, $\{x_k\}$ dan $\{g_k\}$ dihasilkan oleh algoritma SCD-MA. Maka

$$\lim_{k\to\infty}\inf\|\boldsymbol{g}_k\|=0.$$

Bukti. Andaikan tidak benar bahwa

$$\lim_{k\to\infty}\inf\|\boldsymbol{g}_k\|=0.$$

Maka terdapat suatu konstan ε>0 sedemikian sehingga untuk setiap k berlaku $\|\boldsymbol{g}_k\|\ge\varepsilon$. Berdasarkan lemma 3.3, terdapat konstan M_1>0 sedemikian sehingga untuk semua k berlaku $\|\boldsymbol{d}_k\|\le M_1$. Di lain pihak, $\alpha_k\ge M_2\frac{\|\boldsymbol{g}_k\|^2}{\|\boldsymbol{d}_k\|^2}$ untuk suatu konstan M_2>0.

Sehingga kita dapatkan

$$\|\boldsymbol{g}_k\|^2\le\frac{M_1}{M_2}\alpha_k\|\boldsymbol{d}_k\|.$$

Untuk $k\to\infty$ berlaku $\lim_{k\to\infty}\alpha_k\|\boldsymbol{d}_k\|=0$.

Hal ini kontradiksi dengan $\|\boldsymbol{g}_k\|\ge\varepsilon$ untuk semua k dan untuk suatu ε>0. Sehingga berlaku $\lim_{k\to\infty}\inf\|\boldsymbol{g}_k\|=0$.

Dengan demikian metode *spectral conjugate descent* menghasilkan arah pencarian d_k yang menurun. Dengan menggunakan pencarian garis Armijo yang termodifikasi, metode ini konvergen global.

SIMPULAN DAN SARAN

Metode *nonlinear spectral conjugate descent* baru yang merupakan gabungan dari metode conjugate descent dan metode spectral conjugate gradient terbukti mempunyai arah pencarian yang menurun (*descent direction*). Dengan menggunakan pencarian garis Strong Wolfe, metode ini konvergen global. Dengan menggunakan pencarian garis Armijo yang termodifikasi yang dilaporkan merupakan pencarian garis yang efisien, metode *spectal conjugate descent* memberikan hasil yang konvergen global. Selain pencarian garis Armijo yang termodifikasi, masih terdapat dua pencarian garis tidak eksak yang termodifikasi yang perlu diteliti kekonvergenan metode *nonlinear spectral conjugate descent* dengan pencarian garis tersebut sehingga dapat digunakan sebagai bahan penelitian selanjutnya.

DAFTAR PUSTAKA

Birgin EG & Martinez JM, 2001. A Spectral Conjugate Gradient Method for Unconstrained Optimization. *Appl. Math. Optim.*; 43: 117-128.

Dai YH, Liao LZ & Li D, 2004. On Restart Procedures for the Conjugate Gradient Method. *Numerical Algorithms*; 35:249-260.

Kim M, Kwon S & Oh S, 2008. The Performance of A Modified Armijo Line Search Rule in BFGS Optimization Method. *Journal of The Chungcheong Mathematical Society*; 21(1):117-127.

Liu J & Jiang Y, 2012. Global Convergence of A Spectral Conjugate Gradient Method for Unconstrained Optimization. *Abstract and Applied Analysis*; Vol. 2012. doi: 10.1155/2012/758287

Wei Z, Qi L, Ito S, 2000. New step-size rules for optimization problems. Guangxi: Department of Mathematics and Information Science, Guangxi University.

Wei XW, Li GY & Qi LQ, 2008. Global Convergence of The Polak-Ribiere-Polyak Conjugate Gradient Method with an Armijo-Type Inexact Line Search for Nonconvex Unconstrained Optimization Problems. *Mathematics of Computation*; 77(264):2173-2193.

Zhang L, Zhou W & Li D, 2006. Global Convergence of a Modified Fletcher-Reeves Conjugate Gradient Method with Armijo-Type Line Search. *Numer. Math*; 104:561-572.