ANALISIS *GRADIENT CONJUGATE METHOD* UNTUK MINIMISASI FUNGSI

Abdul Jalil
STKIP PGRI Jombang
abduljalil80.stkip@gmail.com

Abstrak

Fungsi persamaan linier panjang Ax = y dengan A dan y adalah matriks merupakan fungsi yang mudah untuk dipecahkan jika nilai dari variabel x nya diketahui. Namun lain halnya jika nilai dari variabel x pada fungsi tersebut tidak diketahui, maka akan menarik jika kita dapat menyelesaikan masalah tersebut karena pemecahannya akan cukup rumit, sehingga dibutuhkan suatu metode yang cukup mudah untuk menyelesaikan fungsi tersebut.

Permasalahan yang demikian ini dapat diselesaikan dengan cara minimisasi fungsi. Salah satu cara untuk minimisasi fungsi adalah dengan menggunakan metode gradien konjugat. Oleh karena itu dalam artikel ini penulis membahas tentang Gradient Conjugate Method atau metode grdien konjugat untuk minimisasi fungsi. Dalam penyelesaiannya meode gradien konjugat meggunakan langkah pemecahan iterasi dari fungsi persamaan linier Ax = y dengan matriks A adalah matriks simetris dan matriks positif definit.

Adapun hasil yang didapatkan terbukti bahwa metode gradien konjugat merupakan metode iterasi yang cukup mudah yang dapat digunakan untuk menyelesaikan persamaan Ax=y dengan nilai dari variabel x yang tidak diketahui.

Kata Kunci : *Minimisasi fungsi, metode Gradient conjugate, matriks*

Abstract

Functions of linear equations Ax = y with length A and y is a matrix is a function that is easy to solve if the value of the variable x in his mind. Yet another case when the value of the variable x in the function is not known, it will be interesting if we can solve these problems because the solution will be quite complicated, so it needs a method that is fairly easy to accomplish these functions. Such problems can be resolved by means of minimization of functions. One way to minimization of functions is to use the method of Conjugate gradient. Therefore, in this article the author discusses the Conjugate Gradient Method or grdien method for conjugate function minimization. Conjugate gradient solution in meode meggunakan step problem solving function equation of linear iterations of Ax = y with A symmetric matrix is a matrix and matrix of positive definit. As for the results that are obtained by Conjugate gradient methods proved that is a fairly easy iteration method that can be used to solve the equation Ax = y to the value of variable x is unknown.

Keywords: *minimization of functions, methods, conjugate Gradient matrix*

1. Pendahuluan

Metode gradien konjugat termasuk pada metode Subruang Krylov yang merujuk pada pemecahan iterasi dari persamaan linear dengan bentuk $Ax = y$. Topiknya adalah untuk membuat persamaan dari solusi perkiraan vektor kombinasi linear dengan tipe u, Au, A^2u, ... A^nu dalam metode ini matriks A adalah matriks simetris dan matriks positif definit.

Asumsikan bahwa $A \in R^{n\,x\,n}$ simetri dan merupakan matrik positif nyata yaitu $A^T = A$ dan $u^T Au > 0$, *untuk* $u \neq 0$. Dalam hal ini setiap nilai eigen (*eigen value*) harus positif dan matriks A dapat di inverskan (*invertible*). *Conjugate gradient* (*CG*) *method* digunakan untuk menentukan approximasi atau pendekatan persamaan x_j dari solusi persamaan $Ax = y$ dengan minimisasi fungsi. Pada artikel ini dibahas tentang persamaan kuadratik $\phi(x) = e^T\,Ae$. Untuk menyelesaikan persamaan kuadratik tersebut maka perlu diketahui tentang persamaan-persamaan berikut:

Jika $Ax = y$ maka didapatkan:

$$x_* = A^{-1}y \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(1)}$$

$$e = x_* - x \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(2)}$$

Dari kedua persamaan tersebut maka didapatkan:

$$\begin{aligned}
r &= y - Ax \\
&= y - A(x_* - e) \\
&= y - Ax_* - Ae \\
&= y - AA^{-1}y + Ae \\
&= y - y + Ae \\
&= Ae
\end{aligned}$$

Dengan x_* adalah solusi *exact*, A^{-1} adalah invers dari A, sedangkan e dan r adalah *error* dan sisa dari Approximasi x. Permasalahan fungsi kuadratik $\phi(x) = e^T\,Ae$ dapat dijabarkan sebagai berikut:

$$\begin{aligned}
\phi(x) &= e^T\,Ae \\
&= (x_* - x)(y - Ax) \\
&= (A^{-1}y - x)^T\,(y - Ax) \\
&= (y - Ax)^T\,A^{-1}(y - Ax) \\
&= r^T\,A^{-1}\,r
\end{aligned}$$

Dengan e dan r adalah error dan sisa dari Approximasi x. menghitung nilai dari fungsi ini untuk x yang diberikan pada solusi exact x_* atau alternatifnya A^{-1}. Oleh karena itu langkah pertama masalah minimisasi fungsi ini kita teliti dengan definisi A

$$\phi(x) = 0 = \min_{x \in R^n} \phi(x) \text{ jika dan hanya jika } x = x$$

Asumsikan kita punya x_1 awal dan jarak awal s_1, kemudian pandang fungsi minimumnya:

$$R \to R \qquad \alpha \to \phi(x_1 + \alpha s_1)$$

Sangat menarik jika kita dapat menyelesaikan masalah minimisasi tanpa mengetahui nilai dari ϕ. Berikutnya dibuktikan beberapa teorema yang mendukung untuk minimisasi fungsi tersebut.

2. Pembahasan

Sebelum membahas tentang pemecahan minimisasi dari fungsi kuadratik maka terlebih dahulu dibuktikan mengenai lemma yang terkait dengan permasalahan sebagai berikut:

Lemma 1. fungsi $\alpha \to \phi(x_1 + \alpha s_1)$ minimum pada

$$\alpha = \alpha_1 = \frac{s_1{}^T r_1}{s_1{}^T A s_1}$$

Dimana r_1 adalah sisa dari initial (taksiran) awal

$$r_1 = y - Ax$$

Bukti:

Sisa yang terhubung pada $x = x_1 + \alpha s_1$ adalah

$$\begin{aligned} y - Ax &= y - A(x_1 + a s_1) \\ &= y - Ax_1 - aAs_1 \\ &= r_1 - As_1 \end{aligned}$$

dan

$$\begin{aligned} \phi(x) &= r^T A^{-1} r \\ &= (r_1 - \alpha A s_1)^T A^{-1} (r_1 - \alpha A s_1) \\ &= \alpha^2 s_1{}^T A s_1 - 2\alpha s_1{}^T r_1 + r_1{}^T A^{-1} r_1 \end{aligned}$$

Selanjutnya diberikan rangkaian arah s_k ,sehingga kita akan mendapatkan rangkaian x_k dari solusi approximasi dengan:

$$x_{k+1} = x_k + \alpha_k s_k \,, \quad \alpha_k = \frac{s_k^T r_k}{s_k^T A s_k}$$

Dimana r_k adalah residu dari iterasi awal

$$r_k = y - Ax_k$$

Catat bahwa residu pada skema yang terbaru mengikuti bentuk:

$$\begin{aligned} r_{k+1} &= y - Ax_{k+1} \\ &= y - A(x_k + \alpha_k s_k) \\ &= y - Ax_k + \alpha_k A s_k \\ &= r_k - \alpha_k A s_k \end{aligned}$$

Prosedur ini dapat digunakan dengan beberapa arah s_k yang dicari, metode gradien konjugat dapat diketahui dengan memilih arah, berikutnya diberikan sebuah definisi:

Definisi 2. Vektor bebas linier $\{s_1, s_2, ..., s_k\}$ adalah konjugat dari A jika:

$$s_1{}^T A s_j = 0 \, \forall i \neq j$$

Dengan kata lain vektor-vektor adalah orthogonal dengan melihat hasil kali dalam yang didefinisikan dengan matriks *A*:

$$\langle u,v\rangle_A = \langle A^T u, v\rangle = \langle Au, v\rangle = \langle u, Av\rangle = u^T Av$$

Perhatikan bahwa jika diberikan himpunan vektor-vektor $\{u_1,...,u_k\}$ sebagai vektor yang bebas linier selalu mungkin untuk menemukan vektor konjugat A $v_j \in sp\{\mathrm{u}_1,...,\mathrm{u}_k\}$, $1 \le j \le k$ Oleh karena itu $sp\{\mathrm{u}_1,...,\mathrm{u}_k\} = sp\{\mathrm{v}_1,...,\mathrm{v}_k\}$, hal ini dapat dilakukan dengan metode *Gram Schmidt orthogonalization* dengan merespon hasil kali dalam $<.,.>_A$

Diberikan matriks $s_k = \{v_1,...,v_k\} \in R^{nxk}$ maka konjugasi A dari vektor $\{v_j\}$ ekuivalen dengan $\mathrm{s}_k{}^T As_k = D_k = Diag(d_1,...,d_k) \in R^{kxk}$ dengan $d_j \ne 0,\quad 1 \le j \le k$

Untuk memahmi signifikansi dari penggunaan konjugat arah A, pandang permasalahan umum berikut:

Diberikan matriks s: $\{s_1,...,s_k\}$ dengan kolom bebas linier, temukan minimisasi pemetaannya!

$$R_k \to R \qquad h \to \phi(x_1 + s_k h)$$

Dengan kata lain, untuk meminimalkan $\phi(k)$ tidak serta merta setelah diberikan arah, tetapi setelah melewati satu langkah subruang, hasil berikut mengikuti lemma 1.

Lemma 3. fungsi $h \to \phi(x_1 + s_k)$ minimum pada:

$$h = (s_k{}^T As_k)^{-1} s_k{}^T r_1$$

Bukti: Pertama kita perhatikan matriks $s_k{}^T As_k$ *invertible*, jika $s_k{}^T As_k = 0$ maka $x^T s_k{}^T As_k x = 0$, dan definisi positif dari A dan kolom bebas linier s_k mengakibatkan bahwa $x = 0$

Pada saat

$$\begin{aligned} r \quad &= y - A(x_1 + s_k h) \\ &= y - Ax_1 + As_k h \\ &= r_1 - As_k h \end{aligned}$$

Didapatkan

$$\begin{aligned} \phi(x_1 + s_k h) \quad &== r^T A^{-1} r \\ &= (r_1 - As_k h)^T A^{-1}(r_1 - As_k h) \\ &= h^T s_k{}^T As_k h - 2r_1{}^T s_k h + r_1{}^T A^{-1} r_1 \end{aligned}$$

Minimum dari fungsional kuadratik tersebut memenuhi:

$$s_k{}^T As_k h - s_k{}^T r_1 = 0$$

Jadi klaimnya terpenuhi

Secara komputasi untuk meminimalkan *h* menjadi trivial jika matriks $D_k = s_k{}^T As_k$ adalah matriks diagonal, tetapi tidak hanya keuntungan menggunakan arah A konjugat saja, asumsikan deret minimal $x_1,...,x_{k+1}$ telah dihitung dengan 2.2.2, pada saat $x_{k+1} \in x_1 + sp\{s_1,...,s_k\}$ maka didapatkan:

$$\phi(x_{k+1}) \ge \phi(x_1 + s_k h) \quad dengan\ \ h \in R\ \{s_1,...,s_k\}$$

Teorema 4. Asumsikan vektor $\{s_1,...,s_k\}$ bebaslinier dan konjugat A, maka $x_{k+1} = x_1 + s_k h$ Dengan kata lain rangkaian minimisasi ke (k+1) yang juga meminimalkan subruang merentang dengan arah s_j, $1 \le j \le k$

Bukti:

Diberikan $a_j = [a_1,..., aj]^T$ dengan notasi ini didapatkan

$$x_j = x_1 + s_{j-1} a_{j-1}$$

Sehingga residunya adalah:

$$\begin{aligned} r_j &= y - Ax_j \\ &= y - A(x_1 + s_{j-1}a_{j-1}) \\ &= y - Ax_1 - As_{j-1}a_{j-1} \\ &= r_1 - As_{j-1}a_{j-1} \end{aligned}$$

Kita perhatikan dengan konjugat A:

$$\begin{aligned} s_j^T r_j &= s_j^T (r_1 - As_{j-1}a_{j-1}) \\ &= s_j^T r_1 - s_j^T As_{j-1}a_{j-1} \\ &= s_j^T r_1 \end{aligned}$$

Oleh karena itu:

$$a_j = \frac{s_j^T r_j}{s_j^T As_j} = \frac{s_j^T r_1}{s_j^T As_j} = h_j$$

Dengan kata lain tersebut didapatkan $a_k = h$

Akibat 5. Jika vektor $\{s_1,...,s_k\}$ adalah A konjugat dan bebas linier, maka $r_{k+1} \perp sp\{s_1,...,s_k\}$

Bukti:

$$\begin{aligned} r_{k+1} &= y - Ax_{k+1} \\ &= y - A(x_1 + s_k h) \\ &= y - Ax_1 - As_k h \\ &= r_1 - As_k h \end{aligned}$$

Dan juga

$$r_{k+1}^T S_k = r_1^T S_k - h^T S_k^T AS_k = 0$$

Hasil diatas menunjukkan bahwa jika kita dapat memilih arah selanjutnya s_{k+1} sebagai konjugat A dengan hasil yang sebelumnya, pencarian untuk rangkaian minimum juga memberikan global minimum dari suatu subruang, maka pertanyaannya adalah bagaimanakah cara yang efisien untuk menentukan arah konjugat A, hal ini dikenal dengan polynomial orthogonal yang memenuhi tiga tahap yang harus digunakan.

Definisi 6. Diberikan $r_1 = y - Ax_1$ Subruang krylov dari A dengan inisial vektor r_1 yang didefinisikan sebagai:

$$K_k = K_k(A, r_1) = sp\{r_1, aAr_1..., A^{k-1}r_1\}, k \ge 1$$

Dimensi berapakah K_k ?

Jika r_1 adalah vektor eigen dari matriks A Maka $\dim(K_k) = 1, \forall k$ lebih umumnya jika $K \subset R^n$ subruang invarian dari A dan $\dim(K) = m$ maka $r_1 \in K$ mengakibatkan $K_k \subset K$ dan juga $\dim(K_k) \leq m$ implikasi ini akan disajikan kemudian.

Tujuan pembuktian tersebut adalah untuk membuat barisan atau rangkaian arah yang dicari secara induktif, asumsikan bahwa $r_1 \neq 0$ dan $x_1 = x_*$ untuk yang lainnya, kemudian diberikan $s_1 = r_1$.

proses selanjutnya membuktikan dengan induksi pada k, asumsikan bahwa $\forall\ k \geq 1$ didapatkan bentuk himpunan konjugat A $\{s_1,...,s_k\}$ yang arahnya bebas linier sehingga: $sp\{s_1,...,s_k\} = sp\{r_1,...,r_k\} = K_k$

Dengan pilihan s_1, maka benar untuk k=1 hasilnya adalah untuk mencari s_{k+1}. Oleh karena itu kondisi tersebut juga benar untuk *k+1*. Selanjutnya diberikan $r_{k+1} = y - Ax_{k+1} = r_k - a_k AS_k$.

Jika $r_{k+1} = 0$ didapatkan $x_k = x_*$ dan pencarian tersebut menguntungkan. Asumsikan terdapat $r_{k+1} = 0$, pada saat $r_k, s_k \in K_k$ dengan induksi maka $r_{k+1} \in K_{k+1}$ dan lainnya dengan akibat 5 yaitu $r_{k+1} \perp s_j\ \forall j, 1 \leq j \leq k$

$sp\{s_1,...s_k, s_{k+1}\} = sp\{r_1,...,r_{k+1}\} = K_{k+1}$

Untuk s_{k+1} mendapatkan konjugat A pada arah yang awal, kita dapatkan bentuk: $s_{k+1} = r_{k+1} + S_k\beta \in K_{k+1}, \beta \in R^k$

Koefisien vektor β dihitung dengan menggunakan kondisi konjugasi A.

${s_k}^T As_{k+1} = 0$ Yaitu: $D_k\beta = {s_k}^T As_k\beta = -{s_k}^T Ar_{k+1} = -(As_k)^T r_{k+1}$

Dari sini kita dapatkan

$$As_k = [As_{k-1}, As_k]$$

Kolom dari matriks As_{k-1} semua termasuk pada:

$$A(sp\{s_1,...,s_{k-1}\} = (Ak_{k-1}) \subset k_k = sp\{s_1,...,s_{k-1}\} \text{ dan } sp\{s_1,...,s_{k-1}\}$$

$$D_k\beta = \begin{bmatrix} 0 \\ \\ -s_k^T Ar_{k+1} \end{bmatrix}$$

Dengan kata lain $\beta_1 = ... = \beta_{k-1} = 0$ didapatkan

$$s_{k+1} = r_{k+1} + \beta_k s_k \qquad \beta_k = \frac{-s_k^T Ar_{k+1}}{s_k^T As_k}$$

Setelah didapatkan unsur-unsur yang cocok dari langkah minimisasi dapat dibuat modifikasi untuk mengupdate bentuk-bentuk yang membutuhkan stabilitas komputasi dari alghoritma, jika $r_{k+1} \perp r_k$ didapatkan:

$$s_k^T r_k = (r_k + \beta_{k-1}s_{k-1})^T r_k = \| r_k \|^2$$

Dengan kata lain formula dapat ditulis $a_k = \dfrac{\| r_k \|}{s_k^T As_k}$

Kemudian pada saat $r_k \in sp\{s_1,...,s_{k-1}\}$ didapatkan $r_{k+1} \perp r_k$ berakibat:

$$\| r_{k+1} \|^2 = r_{k+1}{}^T (r_k - a_k A s_k) = -\frac{\| r_k \|^2}{s_k{}^T A s_k} r_{k+1}^T A s_k$$

$$= \| r_k \|^2 \ \beta_k$$

Sehingga diperoleh:

$$\beta_k = \frac{\| r_{k+1} \|^2}{\| r_k \|^2}$$

Berikut ini langkah-langkah *conjugate gradient* (*CG*)*:*
Pilih x_1 himpunan $k=0$, $r_0=y-Ax_0, s_0=r_0$; Ulangi higga konvergen

$$a_k = \frac{\| r_k \|}{s_k^T A s_k}$$

$$x_{k+1} = x_k - a_k s_k$$

$$r_{k+1} = r_k - a_k A s_k$$

$$\beta_k = \frac{\| r_{k+1} \|^2}{\| r_k \|^2}$$

$$s_{k+1} = r_{k+1} + \beta s_k$$

$$k \leftarrow k+1$$

Pada saat konjugat arah bebas linier, maka algoritma gradien konjugat membutuhkan arah sebanyak-banyaknya n langkah untuk konvergensi, residu awal adalah subruang invariant k di A $\dim(k) = m < n$, maka alghoritma konvergensi membutuhkan m langkah, oleh karena itu jika menggunakan metode gradien konjugat untuk menyelesaikan invers problem, kita tidak harus megiterasi sampai residunya 0 sebagai gantinya iterasi diakhiri, norm dari residu lebih kecil atau sama dengan perkiraan nilai norm dari *noise*nya

Aplikasi Dari Conjugate Gradient Method

Contoh 1: Untuk menggambarkan metode gradien konjugat, kita akan menyelesaikan contoh sederhana.

Misal: Sistem linear $Ax = y$ diberikan oleh:

$$Ax = \begin{bmatrix} 4 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

dan

$$x_0 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Solusi

Langkah pertama adalah menghitung sisa vektor r_0 yang berhubungan dengan x_1.
Dengan rumus, $r_0 = y - Ax_0$ sehingga

$$r_0 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 4 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -8 \\ -3 \end{bmatrix}$$

Selanjutnya hitung skalar α_0 menggunakan hubungan:

$$\alpha_0 = \frac{s_0{}^T r_0}{s_0{}^T A s_0}$$

$$= \frac{\begin{bmatrix} -8 & -3 \end{bmatrix}\begin{bmatrix} 8 \\ 3 \end{bmatrix}}{\begin{bmatrix} -8 & -3 \end{bmatrix}\begin{bmatrix} 4 & 1 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 8 \\ 3 \end{bmatrix}} = \frac{73}{331}$$

Kemudian hitung x_{k+1} dengan:

$$x_1 = x_0 + a_0 s_0$$

$$= \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \frac{73}{331}\begin{bmatrix} -8 \\ -3 \end{bmatrix} = \begin{bmatrix} 0.2356 \\ 0.3384 \end{bmatrix}$$

Langkah selanjutnya shitung vektor r_1 residu berikutnya dengan:

$$r_1 = r_0 - \alpha_0 A s_0$$

$$= \begin{bmatrix} -8 \\ -3 \end{bmatrix} - \frac{73}{331}\begin{bmatrix} 4 & 1 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} -8 \\ -3 \end{bmatrix} = \begin{bmatrix} -0.2810 \\ 0.7492 \end{bmatrix}$$

Hitung β yang pada akhirnya akan digunakan untuk menentukan arah berikutnya:

$$\beta_0 = \frac{r_1^T r_1}{r_0^T r_0}$$

$$= \frac{\begin{bmatrix} -0.2810 & 0.7492 \end{bmatrix}\begin{bmatrix} -0.2810 \\ 0.7492 \end{bmatrix}}{\begin{bmatrix} -8 & -3 \end{bmatrix}\begin{bmatrix} -8 \\ -3 \end{bmatrix}}$$

$$= 0.0088$$

Kemudian dengan menggunakan skalar β_0, kita hitung arah berikutnya:

$$s_{k+1} = r_{k+1} + \beta_0 s_k$$

$$s_1 = r_1 + \beta_0 s_0$$

$$= \begin{bmatrix} -0.2810 \\ 0.7492 \end{bmatrix} + 0.0088\begin{bmatrix} -8 \\ -3 \end{bmatrix} = \begin{bmatrix} -0.3511 \\ 0.7229 \end{bmatrix}$$

Kemudian kembali menghitung skalar α_1

$$\alpha_1 = \frac{s_1^{\ T} r_1}{s_1^T A s_1}$$

$$= \frac{\begin{bmatrix}0.2810 & 0.7492\end{bmatrix}\begin{bmatrix}-0.2810 \\ 0.7492\end{bmatrix}}{\begin{bmatrix}0.3511 & 0.7229\end{bmatrix}\begin{bmatrix}4 & 1 \\ 1 & 3\end{bmatrix}\begin{bmatrix}0.3511 \\ 0.7229\end{bmatrix}}$$

$$= 0.4122$$

Akhirnya, kita menemukan $\mathbf{x}_2$ menggunakan metode yang sama seperti yang digunakan untuk mencari $\mathbf{x}_{1.}$

$$x_2 = x_1 + a_1 s_1$$

$$= \begin{bmatrix}0.2356 \\ 0.3384\end{bmatrix} + 0.4122\begin{bmatrix}-0.3511 \\ 0.7229\end{bmatrix} = \begin{bmatrix}0.0909 \\ 0.6364\end{bmatrix}$$

Daftar Pustaka

J. Nocedal and S. J. Wright, 2006. *Numerical Optimization, chapter 5.*

http://geodesy.gd.itb.ac.id/wedyanto/wp-content/uploads/2012/01/kuliah-ihg2-7c-ed.pdf diakses pada tanggal 29 Maret 2012