Metode Gradien (Bagian 1)

Kuliah 6 | Optimisasi Statistika

rahmaanisa@apps.ipb.ac.id

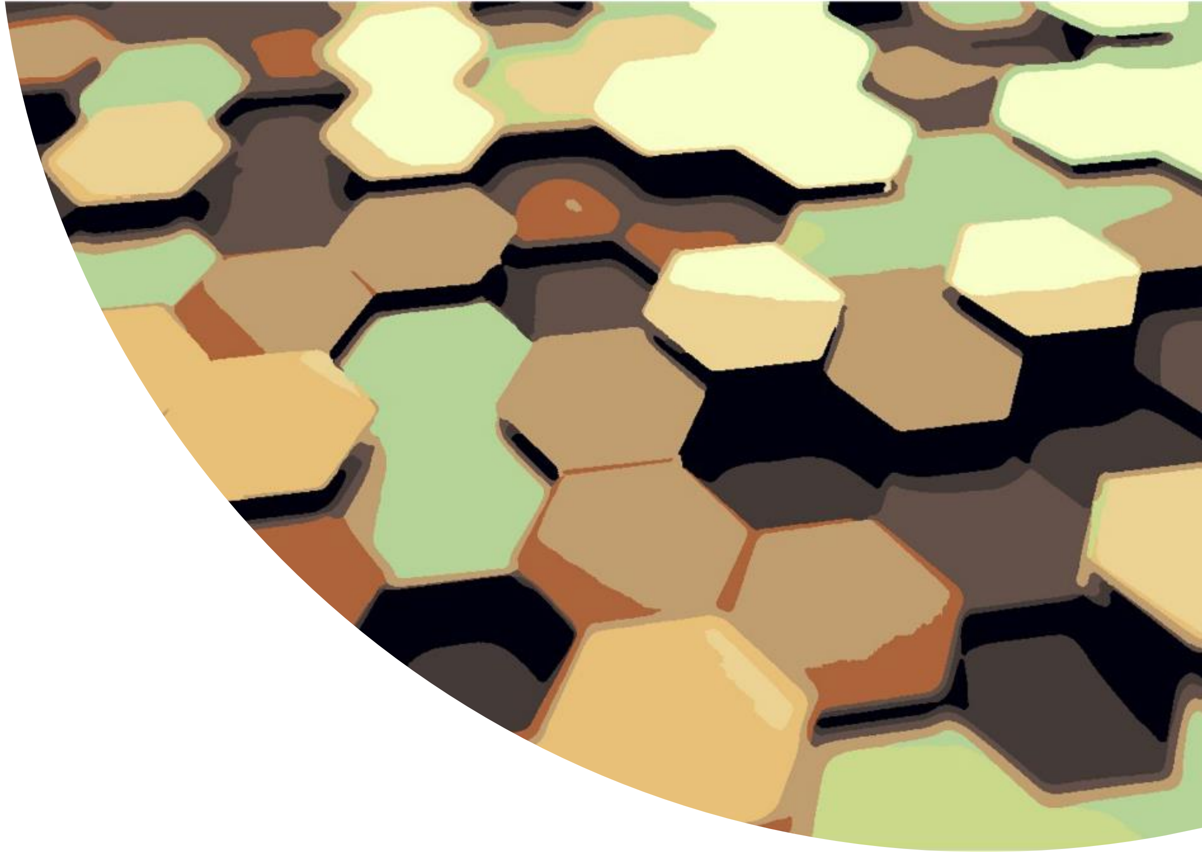
Outline | Metode Gradien

(Bagian 1)

- Pendahuluan tentang Metode Gradien
- Metode Steepest Descent
- Metode Gradien Konjugat

(Bagian 2)

- Metode Newton
- Metode Quasi Newton



PENDAHULUAN

Deskripsi Metode Gradien Secara Umum

Metode Gradien

- Metode ini diperkenalkan pertama kali oleh Louis Augustin Cauchy pada tahun 1847.
- Motivasi Cauchy adalah menghitung orbit benda langit tanpa menyelesaikan persamaan diferensial, melainkan menyelesaikan persamaan aljabar yang mewakili gerak benda tersebut.

(Lemaréchal, 2012)

Sumber gambar: https://www.sapaviva.com/augustin-louis-cauchy/

Metode Gradien

- Metode yang memerlukan evaluasi terhadap turunan dari fungsi dan nilai dari fungsi itu sendiri (Everitt, 1987).

- Metode untuk mencari solusi dari suatu optimisasi yang biasanya melalui proses iterasi, berpindah dari satu titik pada fungsi ke titik lainnya. Hal dasar yang harus ditentukan pada pendekatan ini adalah: arah atau jalur (p) dan panjang langkah (α) (untuk langkah $||\alpha p||$) (Gentle, 2006).

- Metode untuk meminimumkan fungsi $f(x)$ dengan cara mencari suatu nilai

$$x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)}$$

  dengan $k > 0$ dan $d^{(k)}$ dipilih berdasarkan gradien $\nabla f\left(x^{(k)}\right)$

  (Butenko & Parados, 2014).

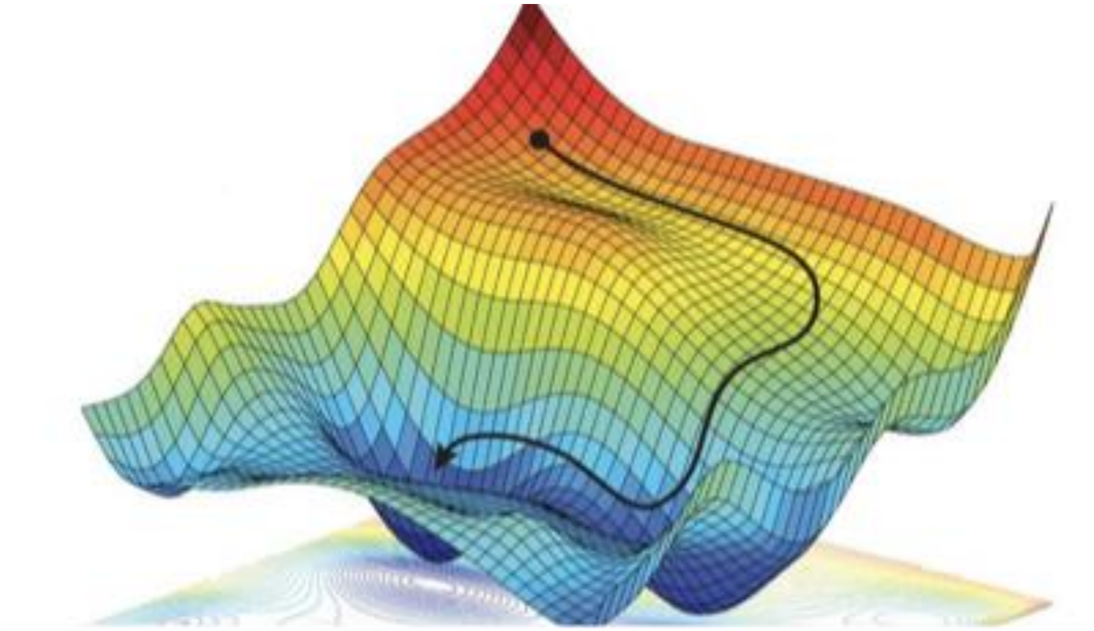
- Metode gradien merupakan salah satu pendekatan dalam menentukan solusi dari permasalahan optimisasi nonlinier tak berkendala. Metode ini dikenal juga sebagai metode “Indirect Search” (Rao, 2019).

Gradien dari Suatu Fungsi

Gradien dari suatu fungsi adalah suatu vektor yang terdiri dari n komponen berikut:

$$\nabla f = \begin{Bmatrix} \partial f / \partial x_1 \\ \partial f / \partial x_2 \\ \vdots \\ \partial f / \partial x_n \end{Bmatrix}$$

Jika kita bergerak sepanjang arah gradien dari titik manapun di dalam ruang berdimensi n, maka nilai fungsi tersebut akan naik (atau turun) dengan laju yang paling cepat.



Sumber gambar: https://medium.com/analytics-vidhya/gradient-descent-b0dc1af33517

Evaluasi Gradien

- Evaluasi akan memerlukan perhitungan turunan $\frac{\partial f}{\partial x_i}$
- Apa saja permasalahan yang mungkin kita temukan?
- Temukan jawabannya di Rao (2019) pada subbab 6.8.

Laju Perubahan dan Arah

$$X = X_i + \lambda S_i$$

laju perubahan arah

Jadi, jika λ^* meminimumkan suatu fungsi f dalam arah S_i, maka $\left.\frac{\partial f}{\partial \lambda}\right|_{\lambda=\lambda^*} = \nabla f|_{\lambda^*}^T S_i = 0$ pada titik $X_i + \lambda^* S_i$.

Steepest Descent (Metode Cauchy)

Algoritma *Steepest Descent* (Cauchy)

1. Menentukan nilai awal $\boldsymbol{X}_1$, dan nomor urut iterasi $i = 1$
2. Menentukan arah $\boldsymbol{S}_i$, yaitu: $\boldsymbol{S}_i = -\nabla f(\boldsymbol{X}_i)$
3. Menentukan panjang langkah optimal λ_i^* pada arah $\boldsymbol{S}_i$ dan
$$\boldsymbol{X}_{i+1} = \boldsymbol{X}_i + \lambda_i^* \boldsymbol{S}_i = \boldsymbol{X}_i - \lambda_i^* \nabla f_i$$
4. Memeriksa keoptimalan pada titik $\boldsymbol{X}_{i+1}$, jika sudah optimum maka proses berhenti di sini. Jika tidak, lanjutkan ke langkah 5.
5. Menetapkan nomor urut iterasi $i = i + 1$, dan kembali ke langkah 2.

Ilustrasi Perhitungan Manual

Contoh 6.8 di Rao (2019)

Minimumkan $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$, dengan nilai awal $\boldsymbol{X}_1 = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}$.

Latihan Soal (1)

Exercise 6.12 (Rao, 2020)

Solve the following equations using the steepest descent method (two iterations only) with the starting point, $\mathbf{X}_1 = \{0\ 0\ 0\}$:

$$2x_1 + x_2 = 4, \qquad x_1 + 2x_2 + x_3 = 8, \qquad x_2 + 3x_3 = 11$$

Latihan Soal (2)

Exercise 6.23 (Rao, 2020)

Perform two iterations of the steepest descent method to minimize the function below:

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

from the starting point $\begin{Bmatrix} -1.2 \\ 1.0 \end{Bmatrix}$.

Conjugate Gradient

(Metode Fletcher-Reeves)

Latar Belakang

- Karakteristik kekonvergenan metode *steepest descent* dapat diperbaiki dengan memodifikasinya menjadi metode gradien konjugat.
- Arah konjugat memiliki sifat konvergensi kuadratik. Hal ini menjamin bahwa metode gradien konjugat akan meminimumkan suatu fungsi kuadratik dalam n langkah atau kurang.
- Sehingga, konvergensi akan diperoleh dengan lebih cepat.

Algoritma Gradien Konjugat (Fletcher-Reeves)

1. Menentukan nilai awal $\boldsymbol{X}_1$, dan nilai urutan iterasi $i = 1$
2. Menentukan arah awal $\boldsymbol{S}_1 = -\nabla f(\boldsymbol{X}_1) = -\nabla f_1$
3. Menentukan nilai $\boldsymbol{X}_2$ berdasarkan persamaan berikut:
$$\boldsymbol{X}_2 = \boldsymbol{X}_1 + \lambda_1^* \boldsymbol{S}_1$$
dengan λ_1^* adalah panjang langkah optimal pada arah $\boldsymbol{S}_1$. Selanjutnya tetapkan nilai $i = 2$ dan lanjutkan ke langkah berikutnya.
4. Menentukan $\nabla f_i = \nabla f(\boldsymbol{X}_i)$, dan
$$\boldsymbol{S}_i = -\nabla f_i + \frac{|\nabla f_i|^2}{|\nabla f_{i-1}|^2} \boldsymbol{S}_{i-1}$$
5. Menghitung panjang langkah optimal λ_i^* pada arah $\boldsymbol{S}_i$, dan menentukan titik baru
$$\boldsymbol{X}_{i+1} = \boldsymbol{X}_i + \lambda_i^* \boldsymbol{S}_i$$
6. Memeriksa keoptimalan pada titik $\boldsymbol{X}_{i+1}$, jika sudah optimum maka proses berhenti di sini. Jika tidak, kembali ke langkah 4.

Ilustrasi Perhitungan Manual

Contoh 6.9 pada Rao (2019)

Minimumkan $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$, dengan nilai awal $\boldsymbol{X}_1 = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}$.

Latihan Soal (3)

Exercise 6.24 (Rao, 2020)

Perform two iterations of the Fletcher-Reeves method to minimize the function below:

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

from the starting point $\left\{\begin{matrix} -1.2 \\ 1.0 \end{matrix}\right\}$.

Latihan Soal (4)

Exercise 6.28 (Rao, 2020)

Prove that the search directions used in the Fletcher–Reeves method are $[A]$-conjugate while minimizing the function

$$f(x_1, x_2) = x_1^2 + 4x_2^2$$

Petunjuk:

- Gunakan titik awal $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$
- Arah $\boldsymbol{s}_1$ dan $\boldsymbol{s}_2$ dikatakan konjugat jika memenuhi $\boldsymbol{s}_1\ \boldsymbol{A}\ \boldsymbol{s}_2 = 0$ dimana $f(x_1, x_2) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}$

Referensi

Butenko,S., Pardalos, P.M. (2014). Numerical Methods and Optimization: An Introduction. CRC Press.

Everitt, B.S. (1987). Introduction to Optimization Methods and Their Application in Statistics. New York: Chapman and Hall.

Gentle, J.E. (2004). Optimization Methods for Applications in Statistics.

Lemaréchal, C. (2012). Cauchy and the gradient method. *Doc Math Extra*, *251*(254), 10.

Rao, S.S. (2020). Engineering Optimization. New Jersey: John Wiley & Sons.