Masalah Optimisasi tanpa Kendala

Dosen: Dr. Diah Chaerani, M.Si

Departemen Matematika FMIPA Universitas Padjadjaran

Minggu ke-2

Materi Kuliah Model Optimisasi

1. **Masalah Optimisasi secara Umum**

$$\min f_0(x)$$
$$s.t \;\; f_i(x) \leq b_i, \quad i = 1, \ldots, m$$

dimana $f_0(x)$ adalah fungsi objektif, $f_i(x)$ adalah fungsi kendala dan $x = (x_1, \ldots, x_m)$ adalah variabel keputusan. Solusi optimal x^* diperoleh bila nilai x^* ini dapat meminimumkan $f_0(x)$ diantara semua x yang ada.

2. **Masalah Optimisasi Tanpa Kendala**

$$\min f_0(x)$$
$$s.t \;\; x \in X$$

dimana $f_0(x) : R^n \rightarrow R$, $x \in X$ dan X adalah himpunan terbuka. Jika $x \in X$ maka x dikatakan solusi fisibel. Jika $x \in X$ dan meminimumkan $f_o(x)$ maka x dikatakan solusi optimal.

- **Masalah Optimisasi dengan Kendala**

$$\begin{aligned} \min\ & f(x) \\ s.t\ & h(x) = 0 \\ & g(x) \leq 0 \\ & x \in X \end{aligned}$$

Jika x memenuhi $h(x) = 0, g(x) \leq 0, x \in X$ maka x disebut solusi fisibel.

Definisi 1

Sebuah bola berpusat di titik $\overline{x}$ dengan jari-jari ϵ didefinisikan sebagai himpunan berikut:

$$B(\overline{x}, \epsilon) := \{x \ : \ \|x - \overline{x}\| \leq \epsilon\}.$$

Perhatikan masalah optimisasi atas himpunan $\mathcal{F}$ berikut ini:

$$\begin{aligned} \min_x \ & f(x) \qquad\qquad (P) \\ s.t \ & x \in \mathcal{F}. \end{aligned}$$

Definisi 2

Sebuah titik $x \in \mathcal{F}$ disebut titik minimum lokal dari (P) jika terdapat $\epsilon > 0$ sedemikian sehingga $f(x) \leq f(y)$ untuk semua $y \in B(\overline{x}, \epsilon) \cap \mathcal{F}$

Definisi 3
Sebuah titik $x \in \mathcal{F}$ disebut titik minimum global dari (P) jika terdapat $\epsilon > 0$ sedemikian sehingga $f(x) \leq f(y)$ untuk semua $y \in \mathcal{F}$.

Definisi 4
Sebuah titik $x \in \mathcal{F}$ disebut titik minimum lokal *strict* dari (P) jika terdapat $\epsilon > 0$ sedemikian sehingga $f(x) < f(y)$ untuk semua $y \in B(\overline{x}, \epsilon) \cap \mathcal{F}, y \neq x$.

Definisi 5
Sebuah titik $x \in \mathcal{F}$ disebut titik minimum global *strict* dari (P) jika terdapat $\epsilon > 0$ sedemikian sehingga $f(x) < f(y)$ untuk semua $y \in \mathcal{F}, y \neq x$.

Contoh

Perhatikan masalah optimisasi berikut

$$\begin{aligned} &\min\ x_1 + x_2 \\ &s.t\ \ x_1, x_2 \in \mathbb{R}^2. \end{aligned}$$

Tunjukkan bahwa $x^* = (0, 0)$ adalah

1. titik minimum lokal
2. titik minimum global.

Soal Latihan:

1. Untuk $f(x) = \frac{1}{2}x_1^2 + x_1x_2 + 2x_2^2 - 4x_1 - 4x_2 - x_2^2$ gunakan syarat cukup untuk optimalitas suatu titik untuk membuktikan bahwa $\overline{x} = (4, 0)$ adalah titik minimum lokal.
2. Misalkan diberikan $f(x) = x_1^3 + x_2^2$, tunjukkan bahwa titik $x = (0, 0)$ bukan titik minimum lokal.
3. Misalkan $f(x) = x_1^4 + x_2^2$ tunjukkan bahwa titik $x = (0, 0)$ merupakan titik minimum lokal.

Solusi:

1. Perhatikan bahwa

$$H(\overline{x}) = \begin{bmatrix} 1 & 1 \\ 1 & 4 \end{bmatrix}$$

adalah matriks positifdefinit hal ini dapat ditunjukkan dengan cara sebagai berikut. Misalkan $d = (d_1, d_2)$ maka diperoleh

$$d^T H d = d_1^2 + 2d_1 d_2 + 4d_2^2 = (d_1 + d_2)^2 + 3d_2^2 > 0, \forall d \neq 0.$$

Dengan demikian $\overline{x}$ memenuhi syarat cukup untuk lokal minimum.

Solusi:

1. Perhatikan bahwa

$$H(\overline{x}) = \begin{bmatrix} 1 & 1 \\ 1 & 4 \end{bmatrix}$$

adalah matriks positifdefinit hal ini dapat ditunjukkan dengan cara sebagai berikut. Misalkan $d = (d_1, d_2)$ maka diperoleh

$$d^T H d = d_1^2 + 2d_1 d_2 + 4d_2^2 = (d_1 + d_2)^2 + 3d_2^2 > 0, \forall d \neq 0.$$

Dengan demikian $\overline{x}$ memenuhi syarat cukup untuk lokal minimum.

Definisi 6

Misalkan $f(x) : X \to \mathbb{R}$ dimana X adalah himpunan buka, maka

1. $f(x)$ dikatakan *differensiabel* di titik $\overline{x} \in X$ jika terdapat vektor $\nabla f(x)$ (**gradien** $f(x)$ di titik $\overline{x}$) sedemikian sehingga untuk setiap $x \in X$ berlaku
$$f(x) = f(\overline{x}) + \nabla f(\overline{x})^t (x - \overline{x}) + \|x - \overline{x}\| \, \alpha(\overline{x}, x - \overline{x}) \qquad (1)$$
dan
$$\lim_{y \to 0} \alpha(\overline{x}, y) = 0. \qquad (2)$$
2. $f(x)$ dikatakan *differensiabel* di himpunan X jika $f(x)$ *differensiabel* di $\forall x \in X$.
3. Vektor gradien didefinisikan sebagai vektor dari turunan parsial
$$\nabla f(x) = \left(\frac{\partial f(\overline{x})}{\partial x_1}, \ldots, \frac{\partial f(\overline{x})}{\partial x_n} \right)^T . \qquad (3)$$

Contoh 1

Misalkan $f(x) = 3x_1^2x_2^3 + x_2^2x_3^3$ maka

$$\nabla f(x) = \left(6x_1x_2^3, 9x_1^2x_2^2 + 2x_2x_3^3, 3x_2^2x_3^2\right)^T$$

Definisi 7

Turunan berarah (*directional derivatives*) dari $f(x)$ dengan arah d adalah

$$\lim_{\lambda\to 0} \frac{f(\overline{x} + \lambda d) - f(\overline{x})}{\lambda} = \nabla f(x)^T d. \quad (4)$$

Definisi 8

1. Fungsi $f(x)$ dapat diturunkan dua kali (*twice differentiable*) pada $\overline{x} \in X$ jika terdapat vektor $\nabla f(\overline{x})$ dan matrik simetrik $H(\overline{x})$, yaitu matriks **Hessian** dari $f(x)$ di $\overline{x}$ sedemikian sehingga $\forall x \in X$ berlaku

$$f(x) = f(\overline{x}) + \nabla f(\overline{x})^t(x - \overline{x}) + \frac{1}{2}H(\overline{x})(x - \overline{x}) + \|x - \overline{x}\|^2 \alpha(\overline{x}, x - \overline{x}) \qquad (5)$$

dan

$$\lim_{y \to 0} \alpha(\overline{x}, y) = 0. \qquad (6)$$

2. Fungsi $f(x)$ dikatakan *twice differensiable* di himpunan X jika $f(x)$ *twice differensiable* di $\forall x \in X$.

3. **Matriks Hessian** adalah matriks dari turunan parsial tingkat dua (*the* second partial derivatives)

$$H(\overline{x})_{ij} = \frac{\partial^2 f(\overline{x})}{\partial x_i \partial x_j}. \qquad (7)$$

Contoh 2

Dari Contoh 1 diperoleh

$$H(x) = \left(\begin{array}{c|c|c} 6x_2^3 & 18x_1x_2^2 & 0 \\ \hline 18x_1x_2^2 & 18x_1^2x_2 + 2x_3^3 & 6x_2x_3^2 \\ \hline 0 & 6x_2x_3^2 & 6x_2^2x_3 \end{array}\right)$$

Definisi 9

Sebuah matriks M dengan ukuran $n \times n$ disebut

1. positif definit jika $x^T Mx > 0$ untuk $\forall x \in \mathbb{R}^n, x \neq 0$.
2. positif semidefinit jika $x^T Mx \geq 0$ untuk $\forall x \in \mathbb{R}^n$.
3. negatif definite jika $x^T Mx < 0$ untuk $\forall x \in \mathbb{R}^n, x \neq 0$.
4. negatif semidefinit jika $x^T Mx \leq 0$ untuk $\forall x \in \mathbb{R}^n, x \neq 0$.
5. indefinit jika terdapat $x, y \in \mathbb{R}^n$ sehingga berlaku $x^T Mx > 0$ dan $y^T My < 0$.

Definisi 10

Sebuah matriks M disebut simetrik positifdefinit (SPD) jika M adalah matriks simetrik dan positif definit. Dengan terminologi yang sama M disebut SPSD jika M adalah matriks symmetric dan positif semidefinit.

Contoh 3

Tunjukkan bahwa matriks M berikut adalah positif definit

$$M_1 = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}, \qquad M_2 = \begin{bmatrix} 8 & -1 \\ -1 & 1 \end{bmatrix}.$$

Jawab: Untuk menunjukkan hal ini, cukup ditunjukkan bahwa untuk setiap $x \in \mathbb{R}^2, x \neq 0$ berlaku

$$x^T M_1 x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^T \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2x_1^2 + 3x_2^2 > 0.$$

$$x^T M_2 x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^T \begin{bmatrix} 8 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= 8x_1^2 - 2x_1x_2 + x_2^2 = 7x_1^2 + (x_1 - x_2)^2 > 0.$$

Perhatikan bahwa fokus utama dari matakuliah Pemrograman Nonlinear ini adalah terkait dengan hal-hal berikut:

1. Eksistensi dari solusi optimal,
2. karakterisasi dari solusi optimal, dan
3. algoritma untuk menghitung solusi optimal.

Untuk melihat ilustrasi dari pertanyaan di atas perhatikan beberapa contoh masalah optimisasi berikut ini.

Contoh 4

$$\begin{aligned} \min \ & \frac{1+x}{2x} \\ s.t \ & x \geq 1. \end{aligned}$$

Perhatikan bahwa dalam masalah ini tidak ditemukan solusi optimal karena daerah fisibilitas tidak terbatas.

Contoh 5

$$\min \frac{1}{x}$$
$$s.t \;\; 1 < x < 2.$$

Dalam masalah ini tidak ditemukan solusi optimal karena daerah fisibilitas bukan himpunan tertutup.

Contoh 6

$$\min \; f(x)$$
$$s.t \;\; 1 \leq x \leq 2,$$
$$\text{dengan } f(x) = \begin{cases} \frac{1}{x}, & x < 2 \\ 1, & x = 2 \end{cases}$$

Tidak ditemukan solusi optimal karena $f(x)$ bukan fungsi *smooth*.

Theorema Weierstrass untuk Barisan

Misalkan $\{x_k\}, k \to \infty$ adalah barisan tak terbatas (*infinit*) dari titik-titik dari suatu himpunan *compact* F (yaitu himpunan yang tertutup dan terbatas). Maka sebagian sub-barisan infinit dari titik-titik x_{k_j} konvergen ke suatu titik di F.

Theorema Weierstrass untuk Fungsi

Misalkan $f(x)$ adalah fungsi bernilai riil dan kontinu pada suatu himpunan *compact* yang tidak kosong $F \subset \mathbb{R}^n$. Maka F memuat suatu titik yang dapat meminimumkan (atau memaksimumkan) $f(x)$ pada himpunan F.

Bukti Theorema Weierstrass untuk Barisan:

Karena Himpunan F terbatas, maka $f(x)$ terbatas di bawah pada F. Karena $F \neq \emptyset$, maka terdapat $v = \inf_{x \in F} f(x)$. Per definisi, untuk sebarang $\epsilon > 0$, maka Himpunan

$$F_\epsilon = \{x \ : \ v \leq f(x) \leq v + \epsilon\}$$

merupakan himpunan tidak kosong. Misalkan $\epsilon_k \to 0$ ketika $k \to \infty$, dan misalkan $x_k \in F_{\epsilon_k}$. Karena F terbatas, terdapat subbarisan $\{x_k\}$ yang konvergen menuju suatu $\overline{x} \in F$. Dengan menggunakan sifat kontinuitas dari $f(x)$, maka berlaku

$$f(\overline{x}) = \lim_{k \to \infty} f(x_k)$$

dan karena $v \leq f(v_k) \leq v + \epsilon_k$ maka berlaku

$$f(\overline{x}) = \lim_{k \to \infty} f(x_k) = v.$$

Perhatikan kembali masalah optimisasi tanpa fungsi kendala berikut:

$$\min \ f_0(x)$$
$$s.t \ \ x \in X$$

dimana $f_0(x) : \mathbb{R}^n \to \mathbb{R}$, $x \in X$ dan X adalah himpunan terbuka.

Definition

Suatu arah $\overline{d}$ disebut *descent direction* dari fungsi $f(x)$ di titik $x = \overline{x}$ jika

$$f(\overline{x} + \epsilon\overline{d}) < f(\overline{x}), \forall \epsilon > 0 \text{ dan cukup kecil.}$$

Theorem

Misalkan $f(x)$ diffensiabel di titik $\overline{x}$. Jika terdapat vektor d sedemikian sehingga $\nabla f(\overline{x})^T d < 0$ maka untuk semua $\lambda > 0$ yang cukup kecil berlaku $f(\overline{x} + \lambda\overline{d}) < f(\overline{x})$ dan dengan demikian $\overline{d}$ adalah descent direction dari $f(x)$ di titik $\overline{x}$.

Bukti: Perhatikan bahwa

$$f(\overline{x} + \lambda d) = f(\overline{x}) + \lambda \nabla f(\overline{x})^T d + \lambda \|d\| \alpha(\overline{x}, \lambda d),$$

dimana $\alpha(\overline{x}, \lambda d) \to 0$ ketika $\lambda \to 0$. Tulis kembali

$$\frac{f(\overline{x} + \lambda d) - f(\overline{x})}{\lambda} = \nabla f(\overline{x})^T d + \|d\| \alpha(\overline{x}, \lambda d).$$

Perhatikan bahwa $\nabla f(\overline{x})^T d < 0$ dan $\alpha(\overline{x}, \lambda d) \to 0$ ketika $\lambda \to 0$, maka $f(\overline{x} + \lambda d) - f(\overline{x}) < 0$ untuk $\forall \lambda > 0$ yang cukup kecil.

Corollary

First order necessary optimality condition: *Misalkan* $f(\overline{x})$ *differensiabel di* $\overline{x}$. *Jika* $\overline{x}$ *adalah lokal minimum maka* $\nabla f(\overline{x})^T d = 0$.

Bukti:

Andaikan $\nabla f(\overline{x})^T d \neq 0$ maka $d = -\nabla f(\overline{x})$ adalah arah **descent**, dimana bila hal ini terjadi maka $\overline{x}$ bukanlah titik lokal minimum.

Theorem

Misalkan bahwa $f(x)$ dapat diturunkan dua kali pada titik $\overline{x} \in X$. Jika $\overline{x}$ adalah titik lokal minimum maka $\nabla f(\overline{x}) = 0$ dan $H(\overline{x})$ adalah positif semidefinit.

Bukti:

Dari *first order necessary optimality condition* diketahui bahwa $\nabla f(\overline{x}) = 0$. Andaikan $H(\overline{x})$ bukan matriks positif semidefinit, maka terdapat d sedemikian sehingga $d^T H(\overline{x})d < 0$. Sehingga diperoleh

$$\begin{aligned} f(\overline{x} + \lambda d) &= f(\overline{x}) + \lambda \nabla f(\overline{x})^T d + \frac{1}{2}\lambda^2 d^T H(\overline{x})d + \lambda^2 \|d\|^2 \alpha(\overline{x}, \lambda d) \\ &= f(\overline{x}) + \frac{1}{2}\lambda^2 d^T H(\overline{x})d + \lambda^2 \|d\|^2 \alpha(\overline{x}, \lambda d). \end{aligned}$$

Dimana $\alpha(\overline{x}, \lambda d) \to 0$ ketika $\lambda \to 0$. Tulis kembali menjadi

$$\frac{f(\overline{x} + \lambda d) - f(\overline{x})}{\lambda^2} = \frac{1}{2}\lambda^2 d^T H(\overline{x})d + \lambda^2 \|d\|^2 \alpha(\overline{x}, \lambda d).$$

Karena $d^T H(\overline{x})d < 0$ dan $\alpha(\overline{x}, \lambda d) \to 0$ ketika $\lambda \to 0$, maka $f(\overline{x} + \lambda d) - f(\overline{x}) < 0$ untuk semua $\lambda > 0$ yang cukup kecil sehingga memberikan kondisi kontradiksi dengan asumsi.

Contoh: Diberikan fungsi

$$f(x) = \frac{1}{2}x_1^2 + x_1x_2 + 2x_2^2 - 4x_1 - 4x_2 - x_2^3.$$

Maka

$$\nabla f(x) = (x_1 + x_2 - 4, x_1 + 4x_2 - 4 - 3x_2^2)^T$$

dan

$$H(x) = \begin{bmatrix} 1 & 1 \\ 1 & 4 - 6x_2 \end{bmatrix}.$$

Untuk kondisi $\nabla f(x) = 0$ maka terdapat dua solusi exact yaitu $\overline{x} = (4, 0)$ dan $\overline{x} = (3, 1)$. Tetapi

$$H(x) = \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix}.$$

adalah matriks indefinit dengan demikian satu-satunya kandidat untuk titik minimum lokal adalah $\overline{x} = (4, 0)$.

Theorem

Misalkan $f(x)$ dapat diturunkan dua kali pada titik $\overline{x}$. Jika $\nabla f(\overline{x}) = 0$ dan $H(\overline{x})$ adalah positif definit, maka $\overline{x}$ adalah minimum lokal yang strict.

Bukti: Perhatikan

$$f(x) = f(\overline{x}) + \frac{1}{2}(x - \overline{x})^T H(\overline{x})(x - \overline{x}) + \|(x - \overline{x})\|^2 \alpha(\overline{x}, x - \overline{x}).$$

Misalkan $\overline{x}$ bukan titik minimum lokal yang strict. Maka terdapat barisan $x_k \to x$ sedemikian sehingga $x_k \neq x$ dan $f(x_k) - f(\overline{x})$ untuk setiap k. Definisikan $d_k = \frac{(x-\overline{x})}{\|(x-\overline{x})\|}$, maka

$$f(x_k) = f(\overline{x}) + \|(x - \overline{x})\|^2 (\frac{1}{2}d_k^T H(\overline{x})d_k + \alpha(\overline{x}, x_k - \overline{x})),$$

dengan demikian maka

$$\frac{1}{2}d_k^T H(\overline{x})d_k + \alpha(\overline{x}, x_k - \overline{x}) = \frac{f(x_k) - f(\overline{x})}{\|x_k - \overline{x}\|}.$$

Lanjutan Bukti: Sekarang perhatikan $\|d_k\| = 1$ untuk sembarang k, maka terdapat sub-barisan dari $\{d_k\}$ yang konvergen ke beberapa titik d sedemikian sehingga $\|d\| = 1$. Asumsikan tanpa menghilangkan perumuman bahwa $d_k \to d$ sehingga

$$0 \geq \lim_{k \to \infty} d^T H(\overline{s}) d_k + \alpha(\overline{x}, x_k - \overline{x}) = \frac{1}{2} d^T H(\overline{x}) d,$$

yang merupakan sebuah kontradiksi dengan sifat positif definit dari matriks $H(\overline{x})$.
Catat bahwa

1. Jika $\nabla f(\overline{x}) = 0$ dan $H(\overline{x})$ adalah negatif definit maka $\overline{x}$ adalah titik lokal minimum.
2. Jika $\nabla f(\overline{x}) = 0$ dan $H(\overline{x})$ adalah positif definit maka kita tidak dapat yakin apakah $\overline{x}$ adalah titik lokal minimum.

3 Exercises on Unconstrained Optimization

1. Find points satisfying necessary conditions for extrema (i.e., local minima or local maxima) of the function

$$f(x) = \frac{x_1 + x_2}{3 + x_1^2 + x_2^2 + x_1 x_2}.$$

   Try to establish the nature of these points by checking sufficient conditions.

2. Find minima of the function

$$f(x) = (x_2^2 - x_1)^2$$

   among all the points satisfying necessary conditions for an extremum.

3. Consider the problem to minimize $\|Ax - b\|^2$, where A is an $m \times n$ matrix and b is an m vector.

   a. Give a geometric interpretation of the problem.

   b. Write a necessary condition for optimality . Is this also a sufficient condition?

c. Is the optimal solution unique? Why or why not?

d. Can you give a closed form solution of the optimal solution? Specify any assumptions that you may need.

e. Solve the problem for A and b given below:

$$A = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \ b = \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \end{pmatrix}.$$

4. Let S be a nonempty set in $\Re^n$. Show that S is convex if and only if for each integer $k \geq 2$ the following holds true:

$$x^1, \ldots, x^k \in S \ \Rightarrow \ \sum_{j=1}^{k} \lambda_j x^j \in S$$

whenever $\lambda_1, \ldots, \lambda_k$ satisfy $\lambda_1, \ldots, \lambda_k \geq 0$ and $\sum_{j=1}^{k} \lambda_j = 1$.

Buktikan:

Theorema Weierstrass untuk Barisan

Misalkan $\{x_k\}, k \to \infty$ adalah barisan tak terbatas (*infinit*) dari titik-titik dari suatu himpunan *compact* F (yaitu himpunan yang tertutup dan terbatas). Maka sebagian sub-barisan infinit dari titik-titik x_{k_j} konvergen ke suatu titik di F.

Theorema Weierstrass untuk Fungsi

Misalkan $f(x)$ adalah fungsi bernilai riil dan kontinu pada suatu himpunan *compact* yang tidak kosong $F \subset \mathbb{R}^n$. Maka F memuat suatu titik yang dapat meminimumkan (atau memaksimumkan) $f(x)$ pada himpunan F.

Some references

1. Robert M. Freund, 2004., Introduction to Optimization, and Optimality Conditions for Unconstrained Problems, MIT Lecture Notes for Nonlinear Programming.
2. Luenberger, David G. Linear and Nonlinear Programming. 2nd ed. Reading, MA: Addison Wesley, 1984. ISBN: 0201157942.
3. Bazaraa, Mokhtar S., Hanif D. Sherali, and C. M. Shetty. Nonlinear Programming: Theory and Algorithms. New York: John Wiley and Sons, 1993. ISBN: 0471557935.
4. Rao, S.S. Optimization: Theory and Applications. Wiley Eastern Limited.1989 ISBN: 085226 756 8.