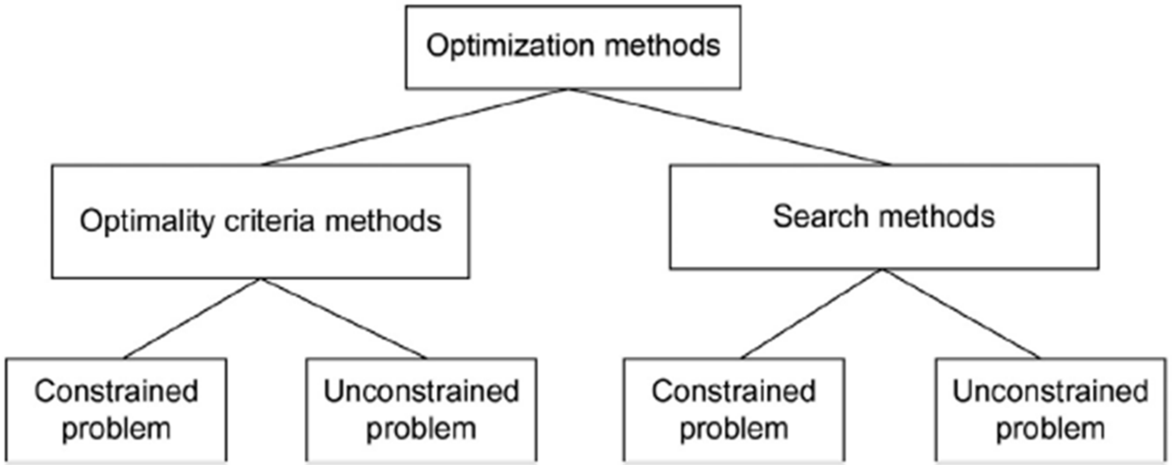
Contents

- Definitions of global and local minima
- Review of some basic calculus concepts
- Concepts of necessary and sufficient conditions
- Optimality conditions: unconstrained problem
- Necessary conditions: equality-constrained problem
- Necessary conditions for a general constrained problem
- Post-optimality analysis: the physical meaning of Lagrange multipliers
- Global optimality
- Second-order conditions for constrained optimization

Classification of Optimization Methods

- Optimality criteria methods (Ch.4~5)
  - Conditions a function must satisfy at its minimum point
  - Seeking solutions to the optimality conditions
- Search methods (Ch.6~13)
  - Numerically searching the design space: direct approach
  - Start with an estimate of the optimum design
  - Search the design space for optimum points



Minimum

- Global (absolute) minimum

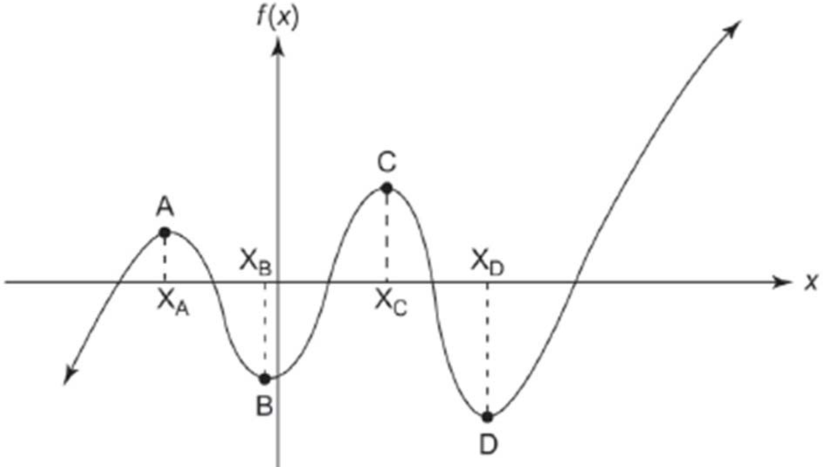
$$f\left(\mathbf{x}^*\right) \le f\left(\mathbf{x}\right) \text{ for all } \mathbf{x} \text{ in the feasible region (constraint set } S\text{)}$$

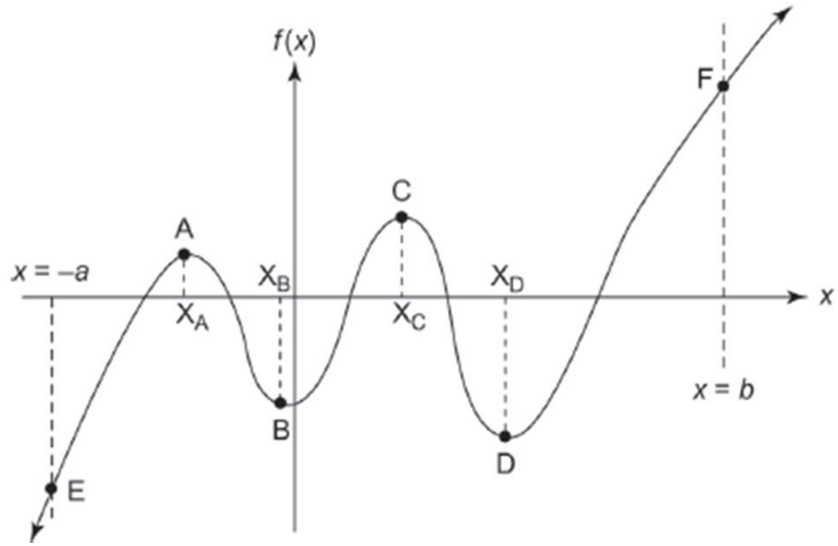
- Local (relative) minimum

$$f\left(\mathbf{x}^*\right) \le f\left(\mathbf{x}\right) \text{ for all } \mathbf{x} \text{ in a small } \textit{neighborhood } N \text{ of } \mathbf{x}^*$$

in the feasible region (constraint set S)

$$N = \left\{\mathbf{x} \middle| \mathbf{x} \in S \quad \text{with} \quad \left\|\mathbf{x} - \mathbf{x}^*\right\| < \delta\right\} \text{ for some small } \delta > 0$$





Example 4.1+2+3

- Find the local and global minima for the function *f* (*x*, *y*) using the graphical method

Minimize

$$f(x,y)=(x-4)^2+(y-6)^2 \quad \text{(a)}$$

subject to

$$g_1 = x+y-12 \le 0 \quad \text{(b)}$$

$$g_2 = x-8 \le 0 \quad \text{(c)}$$

$$g_3 = -x \le 0\,(x \ge 0) \quad \text{(d)}$$

$$g_4 = -y \le 0\,(y \ge 0) \quad \text{(e)}$$

Minimize

$$f(x,y)=(x-10)^2+(y-8)^2 \quad \text{(a)}$$

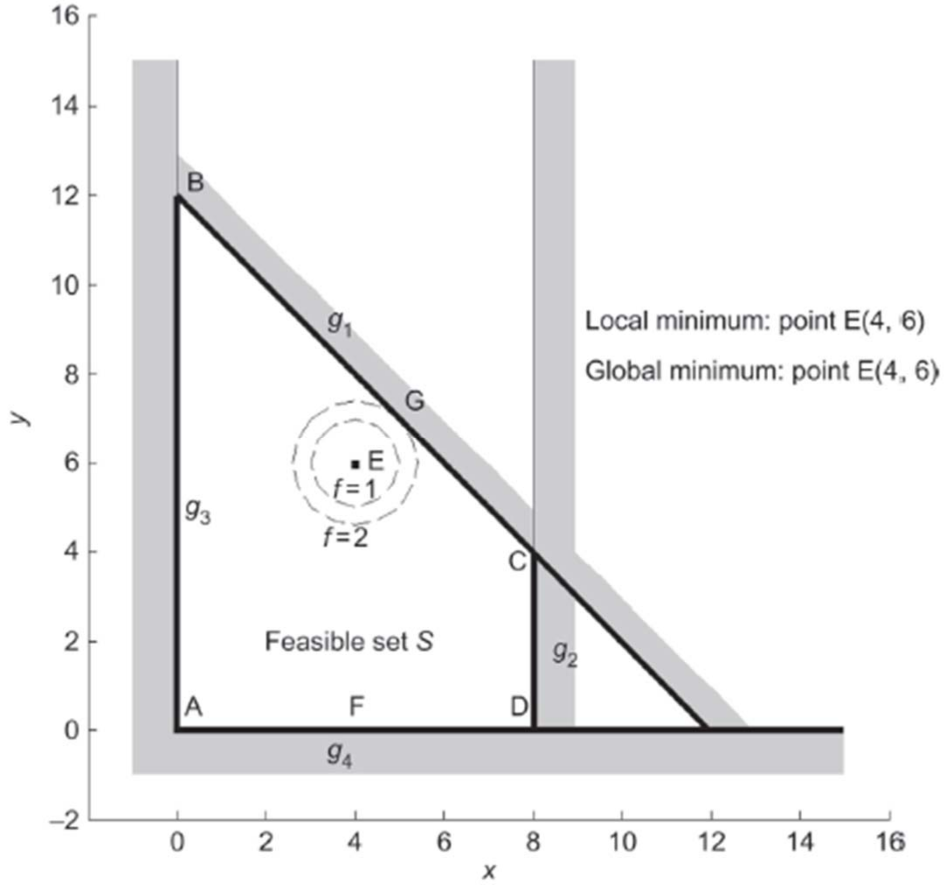
subject to the same constraints as in Example 4.1, Eqs. (b)–(e).

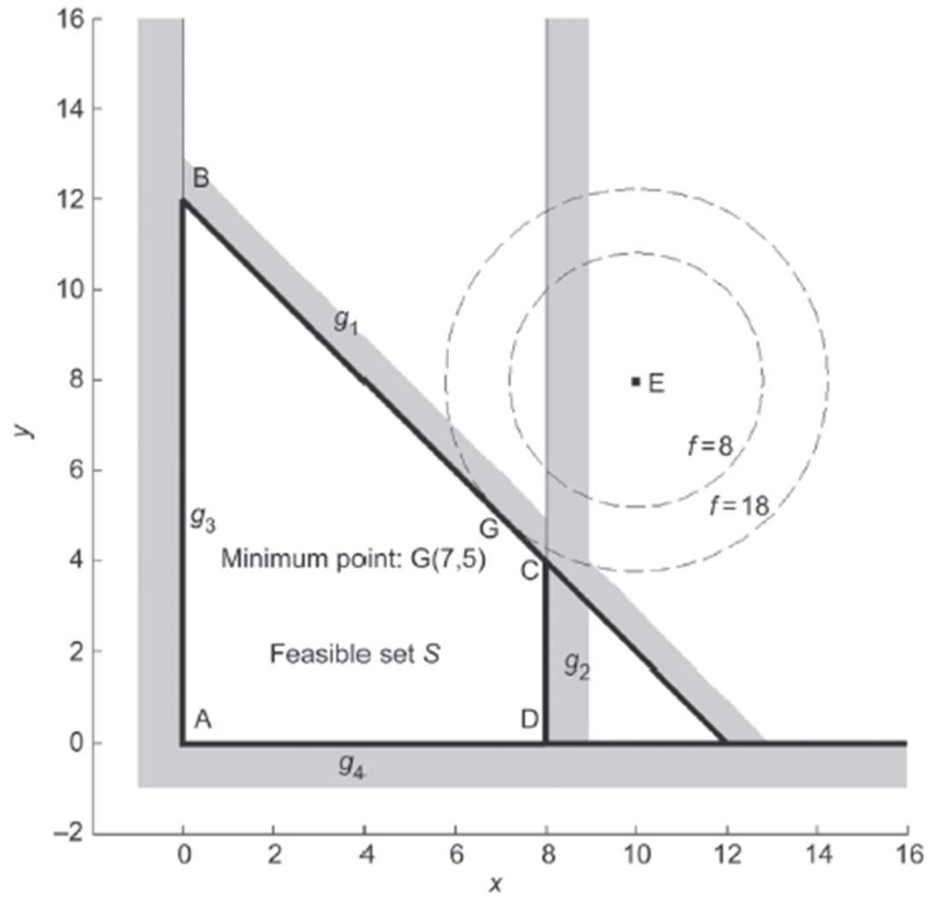
Maximize

$$f(x,y)=(x-4)^2+(y-6)^2 \quad \text{(a)}$$

subject to the same constraints as in Example 4.1, Eqs. (b)–(e).

Example 4.1+2+3





Weierstrass Theorem

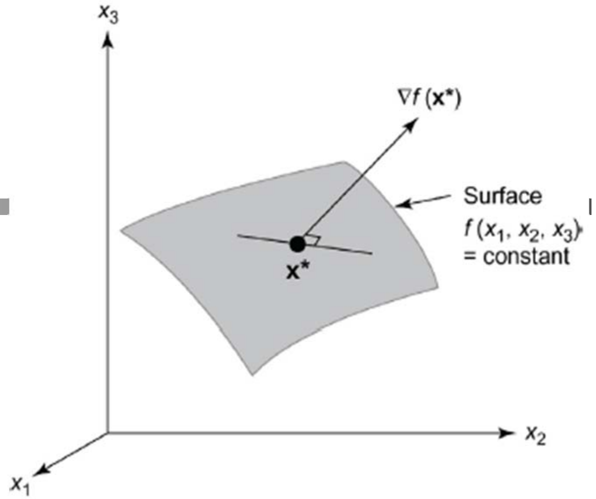
- Existence of global minimum
  - If *f(x)* is continuous on a nonempty feasible set *S* which is closed and bounded, then *f(x)* has a global minimum in *S.*
  - A set *S* is *closed* if it includes all its boundary points and every sequence of points has a subsequence that converges to a point in the set.
  - A set *S* is *bounded* if for any point $\mathbf{x} \in S$, $\mathbf{x}^T\mathbf{x} < c$, where *c* is a finite number.
  - The theorem does not rule out the possibility of a global minimum if its conditions are not met. (not an “if and only if” theorem)

$$e.g., \quad f(x) = -1/x$$

$$\begin{cases} S = \{x \mid 0 < x \le 1\} & : \text{ not closed} \\ S = \{x \mid 0 \le x \le 1\} & : \text{ closed and bounded, not continuous} \end{cases}$$

$$e.g., \quad f(x) = x^2 \text{ subject to } -1 < x < 1$$

Fundamentals



- Gradient vector
  - Normal to the tangent plane at the point **x***

$$\nabla f\left(\mathbf{x}^*\right) = \frac{\partial f\left(\mathbf{x}^*\right)}{\partial \mathbf{x}} = \text{grad } f\left(\mathbf{x}^*\right) = \begin{bmatrix} \dfrac{\partial f\left(\mathbf{x}^*\right)}{\partial x_1} & \cdots & \dfrac{\partial f\left(\mathbf{x}^*\right)}{\partial x_n} \end{bmatrix}^T$$

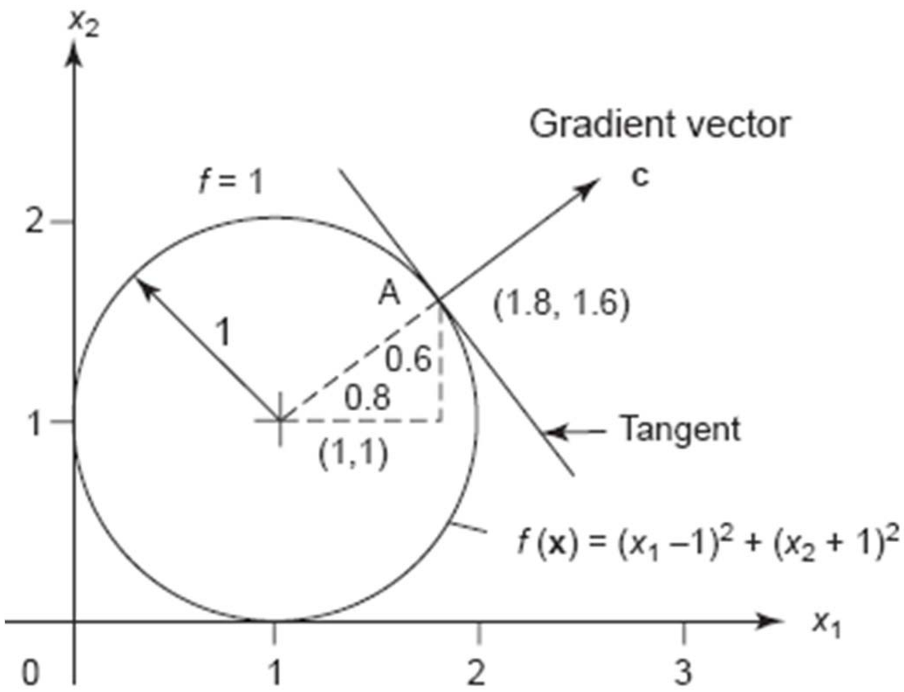
- Hessian matrix
  - Always a symmetric matrix

$$\mathbf{H} = \nabla^2 f\left(\mathbf{x}^*\right) = \frac{\partial^2 f\left(\mathbf{x}^*\right)}{\partial \mathbf{x} \partial \mathbf{x}} = \begin{bmatrix} \dfrac{\partial^2 f\left(\mathbf{x}^*\right)}{\partial {x_1}^2} & \cdots & \dfrac{\partial^2 f\left(\mathbf{x}^*\right)}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f\left(\mathbf{x}^*\right)}{\partial x_n \partial x_1} & \cdots & \dfrac{\partial^2 f\left(\mathbf{x}^*\right)}{\partial {x_n}^2} \end{bmatrix}^T = \left[\frac{\partial^2 f\left(\mathbf{x}^*\right)}{\partial x_i \partial x_j} \right]^T$$

Example

– 점 x* = (1.8, 1.6)에서 다음 함수의 경사도벡터(Gradient)를 구하라.

$$f(\boldsymbol{x}) = (x_1 - 1)^2 + (x_2 - 1)^2$$



Taylor Series Expansion (1)

– Polynomial approximation in a neighborhood of any point in terms of its value and derivatives

- Single variable

$$f(x) = f(x^*) + \frac{df(x^*)}{dx}(x - x^*) + \frac{1}{2}\frac{d^2 f(x^*)}{dx^2}(x - x^*)^2 + R$$

$$x - x^* = d \ : \text{small change in the point } x^*$$

$$f(x^* + d) = f(x^*) + \frac{df(x^*)}{dx}d + \frac{1}{2}\frac{d^2 f(x^*)}{dx^2}d^2 + R$$

- Two variables

$$f(x_1, x_2) = f(x_1^*, x_2^*) + \frac{\partial f}{\partial x_1}(x_1 - x_1^*) + \frac{\partial f}{\partial x_2}(x_2 - x_2^*)$$

$$+ \frac{1}{2}\left[\frac{\partial^2 f}{\partial x_1^2}(x_1 - x_1^*)^2 + 2\frac{\partial f}{\partial x_1 \partial x_2}(x_1 - x_1^*)(x_2 - x_2^*) + \frac{\partial^2 f}{\partial x_2^2}(x_2 - x_2^*)^2\right] + R$$

$$f(x_1, x_2) = f(x_1^*, x_2^*) + \sum_{i=1}^{2}\frac{\partial f}{\partial x_1}(x_i - x_i^*) + \frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2}\frac{\partial^2 f}{\partial x_i \partial x_j}(x_i - x_i^*)(x_j - x_j^*) + R$$

Taylor Series Expansion (2)

– Matrix notation

$$f(\boldsymbol{x}) = f(\boldsymbol{x}^*) + \nabla f^T(\boldsymbol{x}-\boldsymbol{x}^*) + \frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}^*)^T \boldsymbol{H}(\boldsymbol{x}-\boldsymbol{x}^*) + R$$

$$\boldsymbol{x}-\boldsymbol{x}^* = \boldsymbol{d}$$

$$f(\boldsymbol{x}^* + \boldsymbol{d}) = f(\boldsymbol{x}^*) + \nabla f^T \boldsymbol{d} + \frac{1}{2}\boldsymbol{d}^T \boldsymbol{H}\boldsymbol{d} + R$$

$$\Delta f = f(\boldsymbol{x}^* + \boldsymbol{d}) - f(\boldsymbol{x}^*) = \nabla f^T \boldsymbol{d} + \frac{1}{2}\boldsymbol{d}^T \boldsymbol{H}\boldsymbol{d} + R$$

first order change in $f(\boldsymbol{x})$ at $\boldsymbol{x}^*$

$$\delta f = \nabla f^T \delta\mathbf{x} \quad \text{where} \quad \delta\mathbf{x} = \boldsymbol{x}-\boldsymbol{x}^*$$

– examples

$$f(x) = \cos x \ @ \ x^* = 0$$

$$f(\mathbf{x}) = 3x_1^3 x_2 \ @ \ x^* = (1,1)$$

$$f(\mathbf{x}) = x_1^2 + x_2^2 - 4x_1 - 2x_2 + 4 \ @ \ x^* = (1,2)$$

Quadratic Form

– Special nonlinear function having only second-order terms

$$F(\boldsymbol{x}) = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} p_{ij}x_i x_j = \frac{1}{2}\sum_{i=1}^{n} x_i\left(\sum_{j=1}^{n} p_{ij}x_j\right) \xrightarrow{y_i = \sum_{j=1}^{n} p_{ij}x_j \to \boldsymbol{y}=\boldsymbol{Px}} F(\boldsymbol{x}) = \frac{1}{2}\sum_{i=1}^{n} x_i y_i$$

$$F(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{y} = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{Px}$$

$$= \frac{1}{2}\left\{\begin{aligned}&\left[p_{11}x_1^{\ 2} + \cdots + p_{nn}x_n^{\ 2}\right] + \left[(p_{12}+p_{21})x_1x_2 + \cdots + (p_{1n}+p_{n1})x_1x_n\right]\\ &+ \cdots + \left[(p_{n-1,n}+p_{n,n-1})x_{n-1}x_n\right]\end{aligned}\right\}$$

$$\xrightarrow{a_{ij}=\frac{1}{2}(p_{ij}+p_{ji}) \to a_{ij}+a_{ji}=p_{ij}+p_{ji}} F(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{Ax}$$

$$\nabla F(\boldsymbol{x}) = \boldsymbol{Ax}, \quad \frac{\partial^2 F(\boldsymbol{x})}{\partial x_j \partial x_i} = a_{ij}$$

Example

– 다음 이차형식에서 경사도벡터와 헷시안행렬을 계산하라.

$$F(x_1, x_2, x_3) = \frac{1}{2}\left(2x_1{}^2 + 2x_1x_2 + 4x_1x_3 - 6x_2{}^2 - 4x_2x_3 + 5x_3{}^2\right)$$

$$F(\boldsymbol{x}) = \frac{1}{2}\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}\begin{bmatrix} 2 & 2 & 4 \\ 0 & -6 & -4 \\ 0 & 0 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}\begin{bmatrix} 2 & 1 & 2 \\ 1 & -6 & -2 \\ 2 & -2 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Form of a Matrix

- Check the form of a matrix
  - Eigenvalue
  - Principal minors
    - If no two consecutive principal minors are zero

form	definition	eigenvalue	principal minors
positive definite	$x^T Ax > 0$	$\lambda_i > 0$	$M_k > 0\ (k = 1, \dots, n)$
positive semidefinite	$x^T Ax \geq 0$	$\lambda_i \geq 0$	$M_k > 0\ (k = 1, \dots, r)$
negative definite	$x^T Ax < 0$	$\lambda_i < 0$	$\left.\begin{matrix} M_k < 0\ (\text{odd } k) \\ M_k > 0\ (\text{even } k) \end{matrix}\right\} k = 1, \dots, n$
negative semidefinite	$x^T Ax \leq 0$	$\lambda_i \leq 0$	$\left.\begin{matrix} M_k < 0\ (\text{odd } k) \\ M_k > 0\ (\text{even } k) \end{matrix}\right\} k = 1, \dots, r$
indefinite	?	some $\lambda_i < 0$ other $\lambda_i > 0$	

Example

– 다음 행렬의 형태를 결정하라.

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \end{bmatrix} \rightarrow \text{positive definite}$$

$$B = \begin{bmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \rightarrow \text{negative semidefinite}$$

필요조건과 충분조건의 개념

- 최적점은 필요조건을 만족하여야 한다. 필요조건을 만족하는 점을 후보최적점(candidate optimum point)이라 한다. 필요조건을 만족하지 않는 점은 최적점이 될 수 없다.

- 필요조건을 만족한다고 해서 최적점인 것은 아니다. 즉, 비최적점도 필요조건을 만족시킬 수 있다.

- 충분조건을 만족하는 후보최적점은 실제로 최적점이다.

- 충분조건을 사용할 수 없거나 충분조건을 만족하지 않으면 후보점의 최적성 여부에 대하여 어떤 결론도 내릴 수 없게 된다.

비제약최적설계 문제

- 공학의 실제응용에서 자주 나타나는 문제는 아니지만, 제약문제의 최적성조건들은 비제약문제를 논리적으로 확장한 것이므로 개념이해가 중요
- 최적성조건(Optimality Conditions)을 이용하는 방법
  - 어떤 설계점이 주어지면 그 점의 후보최적점 여부 판정
  - 후보최적점을 계산
- 국부적 최적성조건
  - ***x**** 를 ***f(x)*** 의 국부적 최소점이라 하고 ***x*** 를 ***x**** 에 매우 가까운 점이라 하면, ***f(x)***는 ***x**** 에서 국부적 최소이므로 그 점으로부터 매우 작은 거리를 움직였을 때 함수값이 감소될 수 없다.

$$\mathbf{d} = \mathbf{x} - \mathbf{x}^*$$

$$\Delta f = f(\mathbf{x}) - f(\mathbf{x}^*) \geq 0$$

일변수함수의 최적성조건

- 일차 필요조건 (First-order necessary conditions)

$$f(x) = f(x^*) + f'(x^*)d + \frac{1}{2} f''(x^*)d^2 + R$$

$$\Delta f = f(x) - f(x^*) \cong f'(x^*)d \geq 0$$

$$\rightarrow f'(x^*) = 0$$

  – 국부적 최소 또는 최대점, 변곡점 : stationary point

- 충분조건 (Sufficient conditions)

$$\Delta f = \frac{1}{2} f''(x^*)d^2 + R \rightarrow f''(x^*) > 0$$ （최소점에서 함수의 곡률이 양수）

- 이차 필요조건 (Second-order necessary conditions)

$$f''(x^*) \geq 0$$

$$\text{if } f''(x^*) = 0,\ f'''(x^*) = 0 \text{ and } f^{(IV)}(x^*) > 0$$

Examples

- Local minimum points using first-order necessary conditions

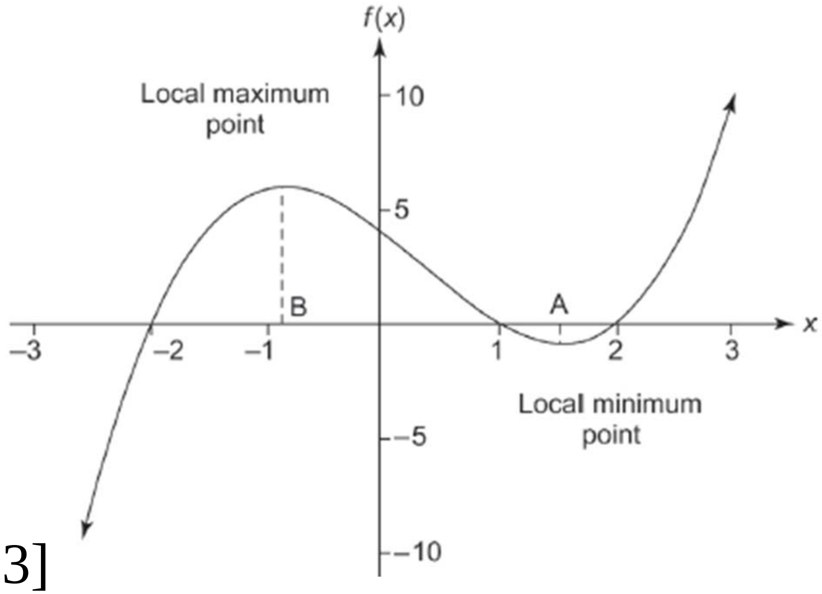
$$1.\ f(x) = \sin x$$

$$2.\ f(x) = x^2 - 4x + 4$$

$$3.\ f(x) = x^3 - x^2 - 4x + 4$$

$$4.\ f(x) = x^4$$

$$5.\ f(x) = ax + \frac{b}{x} \quad (a, b > 0) \text{[section 2.3]}$$



다변수함수의 최적성조건

- 일차 필요조건

$$f(\boldsymbol{x}) = f(\boldsymbol{x}^*) + \nabla f(\boldsymbol{x}^*)^T \boldsymbol{d} + \frac{1}{2}\boldsymbol{d}^T \boldsymbol{H}\boldsymbol{d} + R$$

$$\Delta f = f(\boldsymbol{x}) - f(\boldsymbol{x}^*) = \nabla f(\boldsymbol{x}^*)^T \boldsymbol{d} + \frac{1}{2}\boldsymbol{d}^T \boldsymbol{H}\boldsymbol{d} + R \geq \boldsymbol{0}$$

$$\rightarrow \nabla f(\boldsymbol{x}^*) = \boldsymbol{0}, \quad \frac{\partial f(\boldsymbol{x}^*)}{\partial x_i} = 0;\ i = 1, \ldots, n$$

- (이차)충분조건

$$\Delta f = \frac{1}{2}\boldsymbol{d}^T \boldsymbol{H}\boldsymbol{d} + R \rightarrow \boldsymbol{d}^T \boldsymbol{H}\boldsymbol{d} > 0$$ (헷시안행렬 H 가 positive definite)

- 이차 필요조건

$$\boldsymbol{d}^T \boldsymbol{H}\boldsymbol{d} \geq 0$$

Examples

1. $f(x) = x^2 - 2x + 2$ (effects of scaling or adding constants to a function)

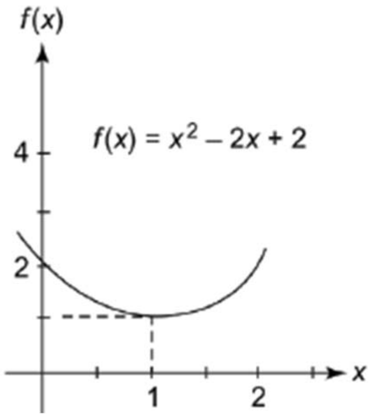
$\rightarrow [f(x)+1], [2f(x)], [-f(x)]$

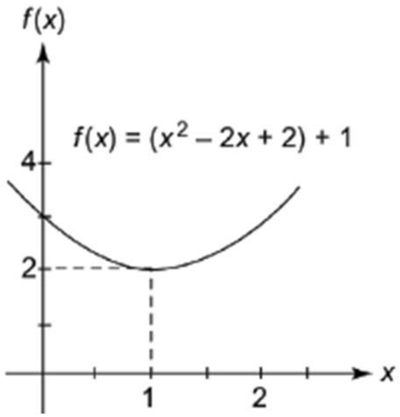
2. $f(\mathbf{x}) = x_1{}^2 + 2x_1x_2 + 2x_2{}^2 - 2x_1 + x_2 + 8$

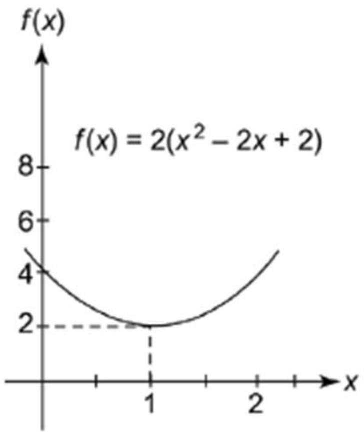
3. $\left.\begin{aligned} \bar{f} &= R^2 + RH \\ h &= \pi R^2 H - V = 0 \end{aligned}\right\} \rightarrow \bar{\bar{f}} = R^2 + \frac{V}{\pi R}$ [section 2.8]

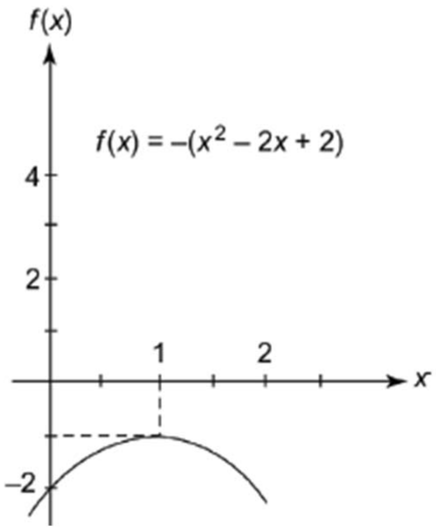
4. $f(x) = \frac{1}{3}x^2 + \cos x$

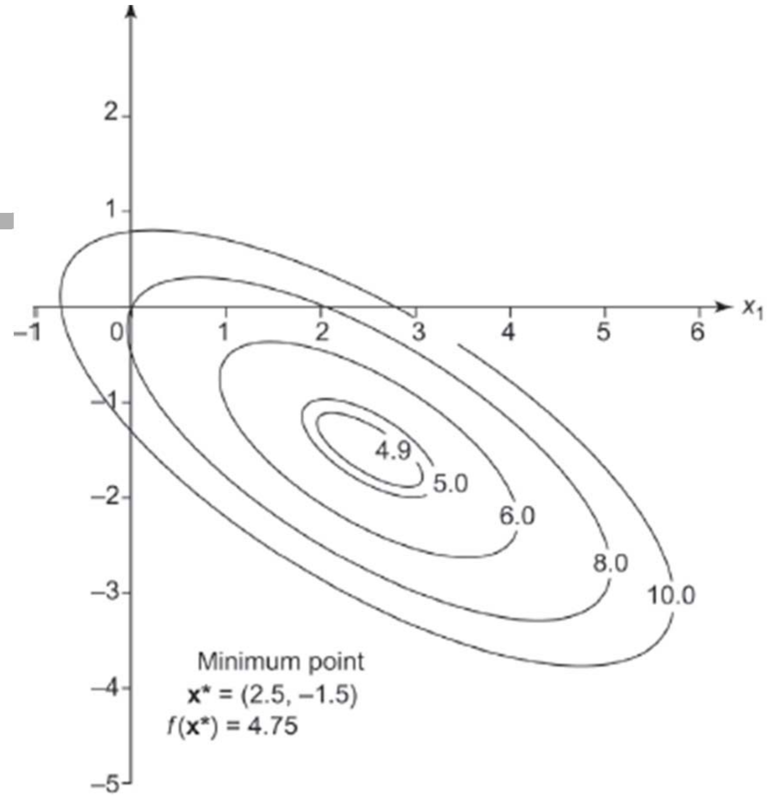
5. $f(\mathbf{x}) = x_1 + \frac{4.0E+06}{x_1 x_2} + 250x_2$

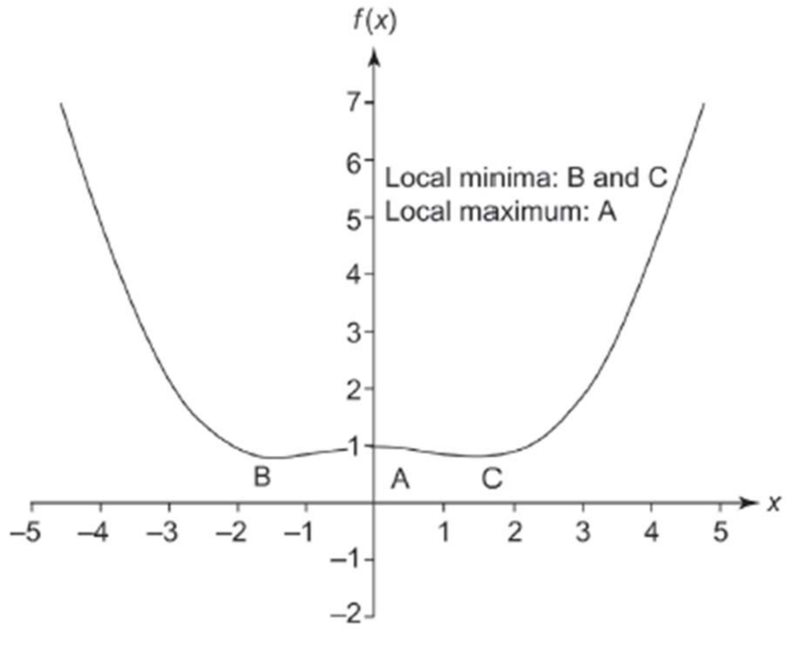

f(x)
f(x) = x^2 – 2x + 2
x


f(x)
f(x) = (x^2 – 2x + 2) + 1
x


f(x)
f(x) = 2(x^2 – 2x + 2)
x


f(x)
f(x) = –(x^2 – 2x + 2)
x


x1
4.9
5.0
6.0
8.0
10.0
Minimum point
x* = (2.5, –1.5)
f(x*) = 4.75


f(x)
Local minima: B and C
Local maximum: A
B
A
C
x

Optimality Conditions for Unconstrained Function

TABLE 4.1(a) Optimality Conditions for Unconstrained One Variable Problems

Problem: Find x to minimize $f(x)$

First-order necessary condition: $f' = 0$. Any point satisfying this condition is called a stationary point; it can be a local maximum, local minimum, or neither of the two (inflection point)

Second-order necessary condition for a local minimum: $f'' \geq 0$

Second-order necessary condition for a local maximum: $f'' \leq 0$

Second-order sufficient condition for a local minimum: $f'' > 0$

Second-order sufficient condition for a local maximum: $f'' < 0$

Higher-order necessary conditions for a local minimum or local maximum: calculate a higher-ordered derivative that is not 0; all odd-ordered derivatives below this one must be 0

Higher-order sufficient condition for a local minimum: highest nonzero derivative must be even-ordered and positive

TABLE 4.1(b) Optimality Conditions for Unconstrained Function of Several Variables

Problem: Find $\mathbf{x}$ to minimize $f(\mathbf{x})$

First-order necessary condition: $\nabla f = \mathbf{0}$. Any point satisfying this condition is called a stationary point; it can be a local minimum, local maximum, or neither of the two (inflection point)

Second-order necessary condition for a local minimum: **H** must be at least positive semidefinite

Second-order necessary condition for a local maximum: **H** must be at least negative semidefinite

Second-order sufficient condition for a local minimum: **H** must be positive definite

Second-order sufficient condition for a local maximum: **H** must be negative definite