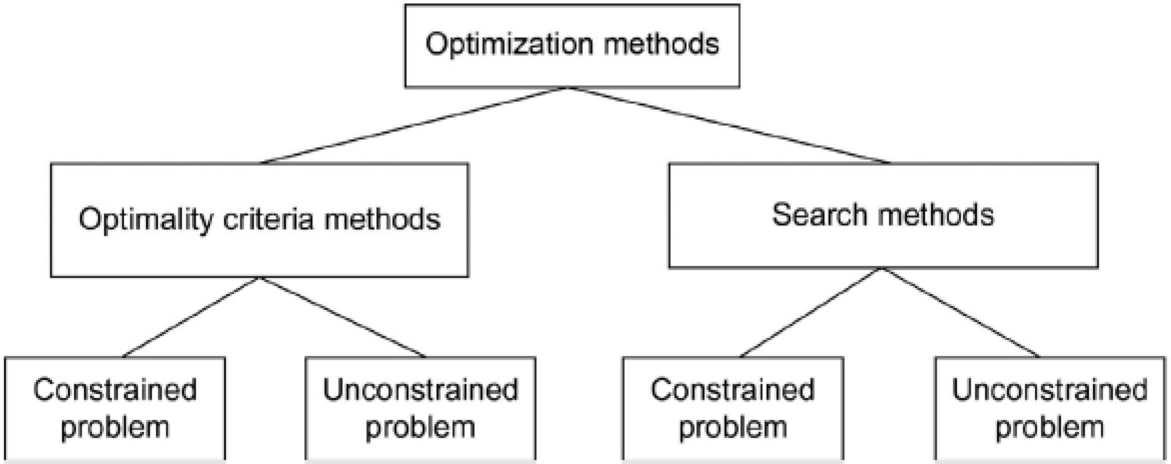
Contents

- 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
- Duality in nonlinear programming

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



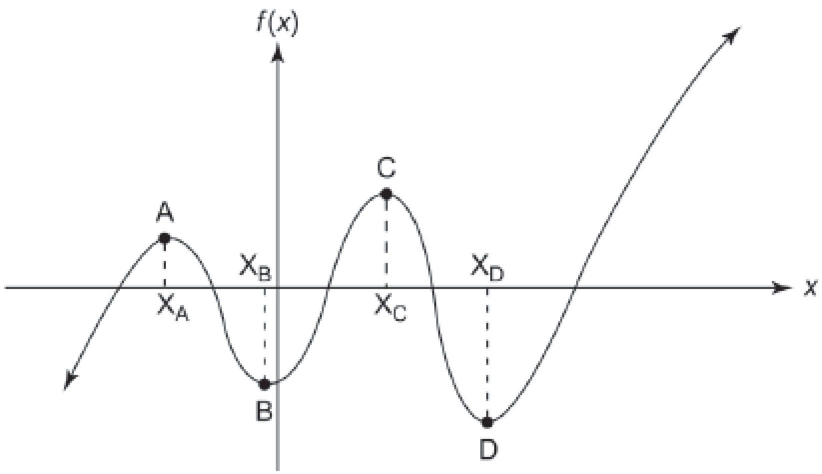
Minimum

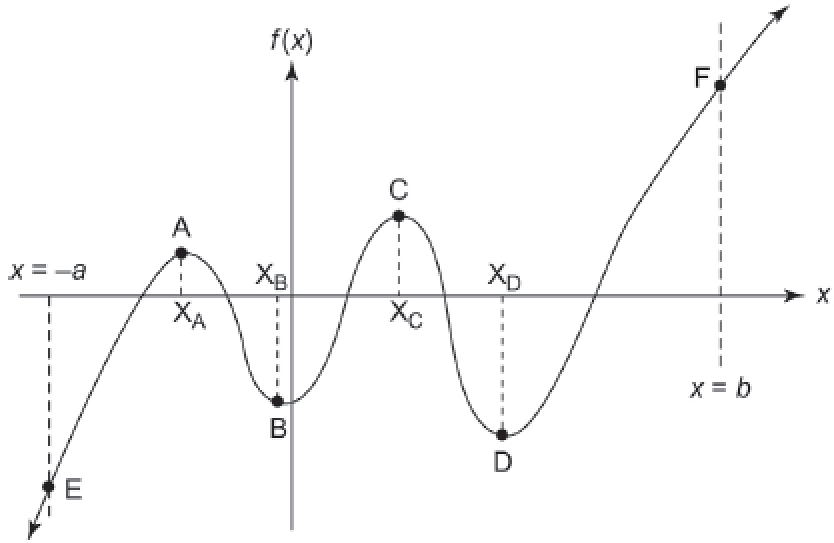
- Global (absolute) minimum

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

- Local (relative) minimum

$$f(\mathbf{x}^*) \le f(\mathbf{x}) \text{ for all } \mathbf{x} \text{ in a small } \textit{neighborhood } N \text{ of } \mathbf{x}^*$$
$$\text{in the feasible region (constraint set } S\text{)}$$
$$N = \{\mathbf{x} \mid \mathbf{x} \in S \quad \text{with} \quad \|\mathbf{x} - \mathbf{x}^*\| < \delta\} \text{ for some small } \delta > 0$$





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

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

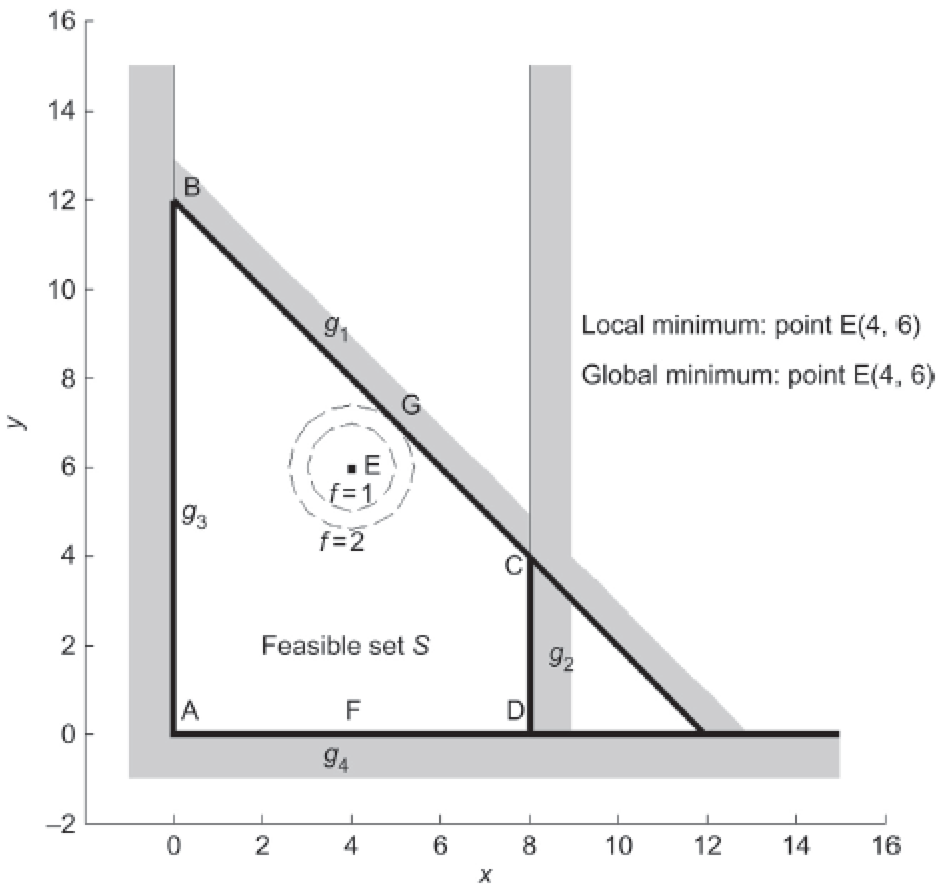
subject to the same constraints as in Example 4.1, Eqs. (b)–(e).

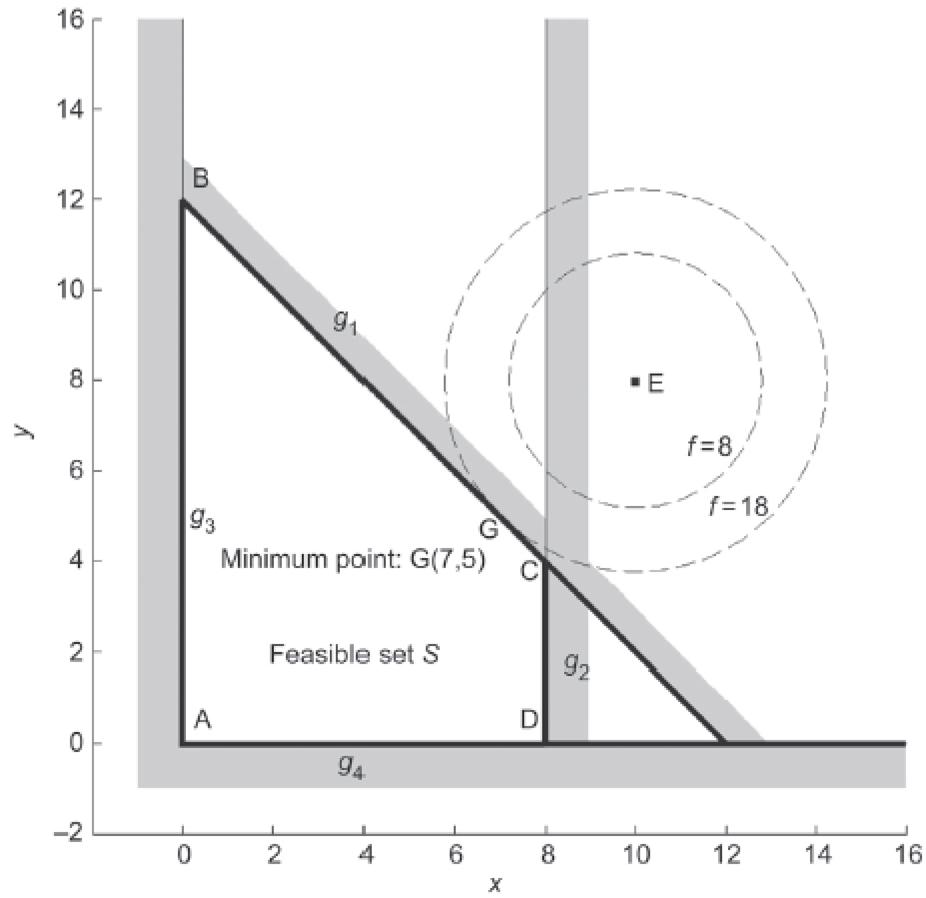
Maximize

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

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

Example 4.1+2+3





Weierstrass Theorem

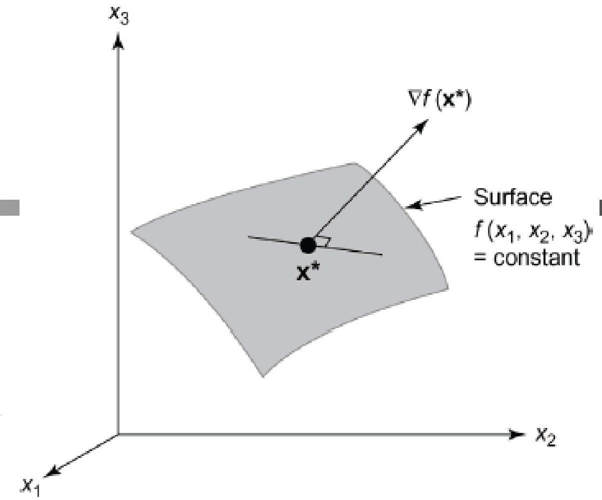
- Existence of global minimum
  - If *f(x)* is continuous on a nonempty feasible set *S* which is closed and bounded, then *f(x)* has a global minimum in *S.*
  - A set *S* is *closed* if it includes all its boundary points and every sequence of points has a subsequence that converges to a point in the 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***

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

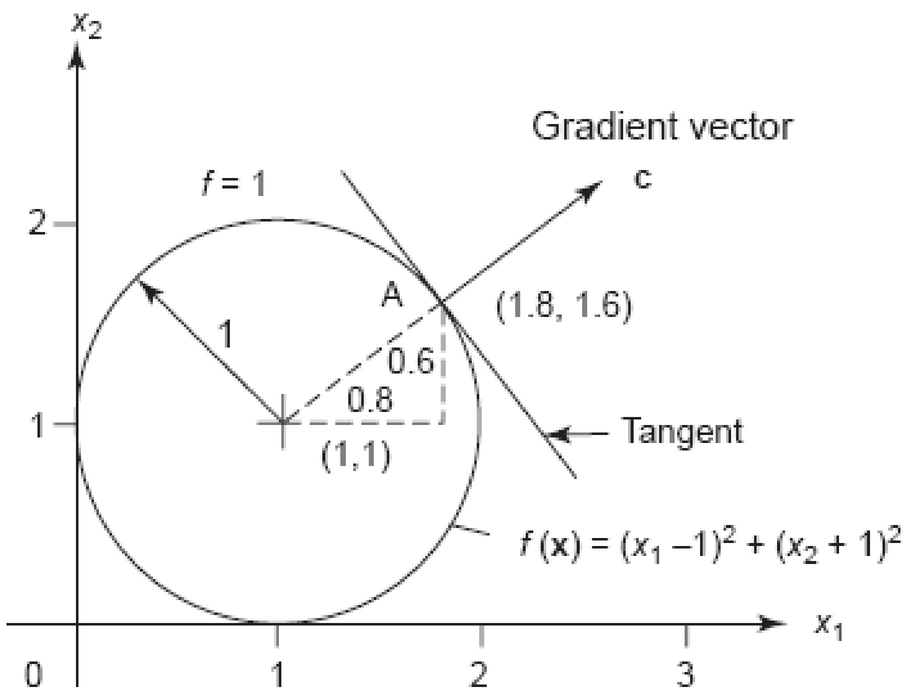
- Hessian matrix
  - Always a symmetric matrix

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

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

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

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

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

Quadratic Form

– Special nonlinear function having only second-order terms

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

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

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

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

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

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	$x^T Ax > 0$	$\lambda_i > 0$	$M_k > 0\,(k = 1, \ldots, n)$
positive semidefinite	$x^T Ax \geq 0$	$\lambda_i \geq 0$	$M_k > 0\,(k = 1, \ldots, r)$
negative definite	$x^T Ax < 0$	$\lambda_i < 0$	$\left.\begin{matrix} M_k < 0\,(\text{odd } k) \\ M_k > 0\,(\text{even } k) \end{matrix}\right\} k = 1, \ldots, n$
negative semidefinite	$x^T Ax \leq 0$	$\lambda_i \leq 0$	$\left.\begin{matrix} M_k < 0\,(\text{odd } k) \\ M_k > 0\,(\text{even } k) \end{matrix}\right\} k = 1, \ldots, r$
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
  - $\boldsymbol{x}^*$ 를 $\boldsymbol{f(x)}$ 의 국부적 최소점이라 하고 $\boldsymbol{x}$ 를 $\boldsymbol{x}^*$ 에 매우 가까운 점이라 하면, $\boldsymbol{f(x)}$는 $\boldsymbol{x}^*$ 에서 국부적 최소이므로 그 점으로부터 매우 작은 거리를 움직였을 때 함수값이 감소될 수 없다.

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

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

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

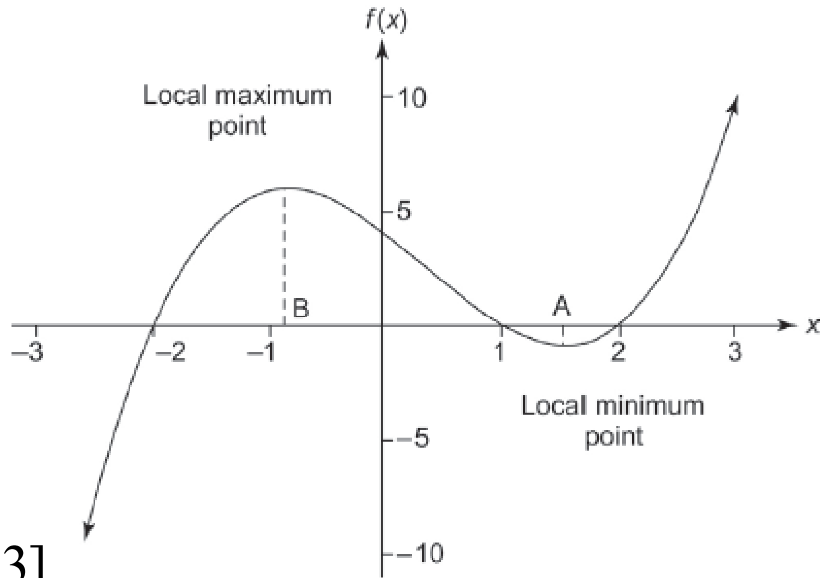
$$1.\ f(x) = \sin x$$

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

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

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

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



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)

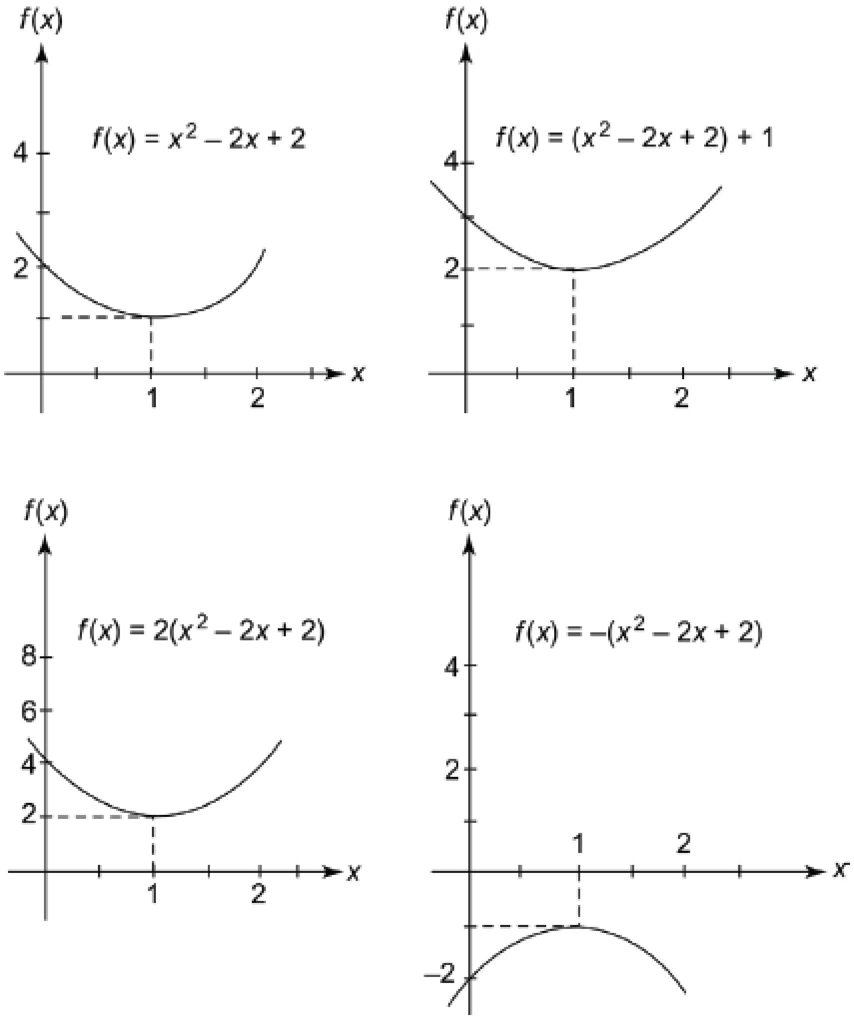
$\rightarrow \left[f(x)+1\right], \left[2f(x)\right], \left[-f(x)\right]$

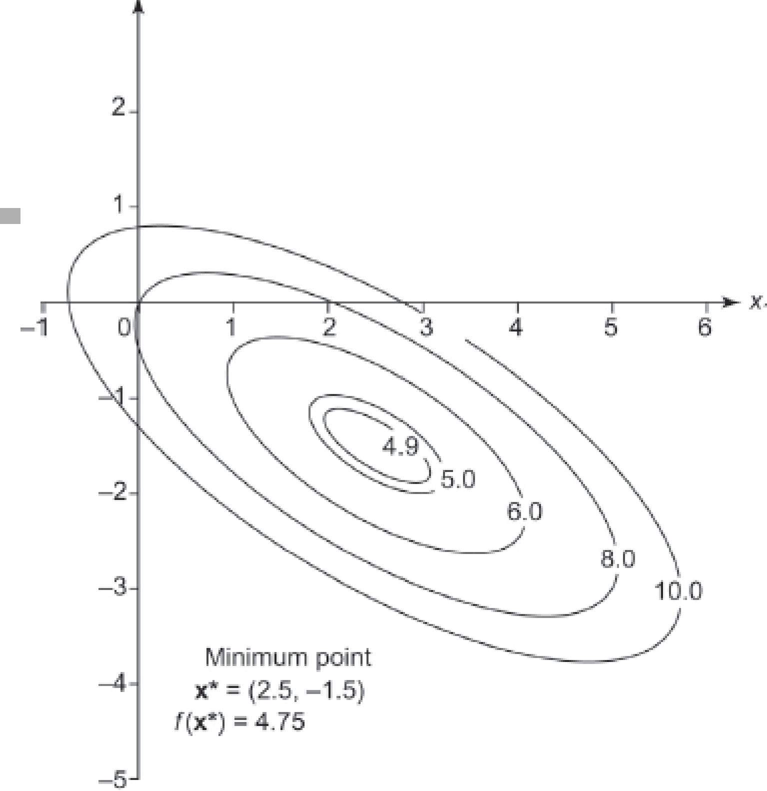
$2.\ f(\mathbf{x}) = x_1^{\ 2} + 2x_1x_2 + 2x_2^{\ 2} - 2x_1 + x_2 + 8$

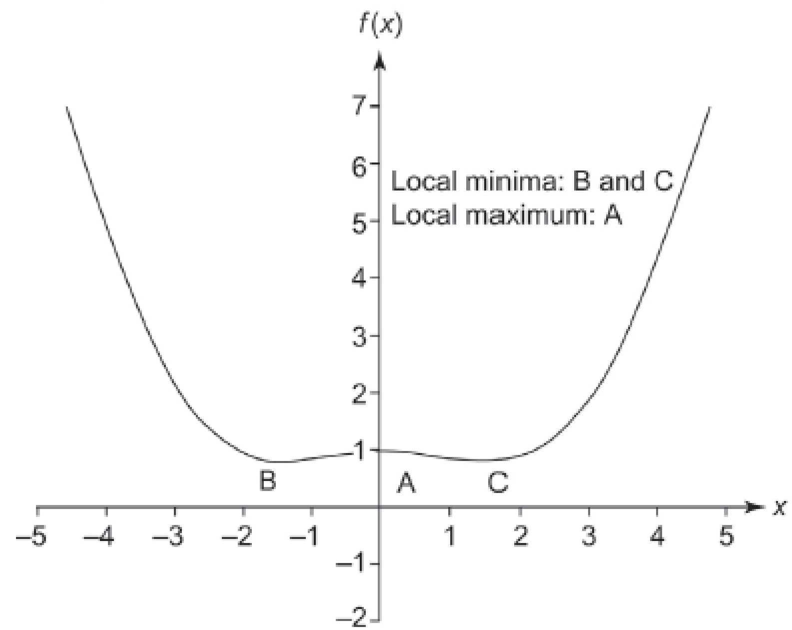
$$3.\ \left.\begin{array}{l} \bar{f} = R^2 + RH \\ h = \pi R^2 H - V = 0 \end{array}\right\} \rightarrow \bar{\bar{f}} = R^2 + \frac{V}{\pi R} \quad \text{[section 2.8]}$$

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

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


$f(x) = x^2 - 2x + 2$
$f(x) = (x^2 - 2x + 2) + 1$
$f(x) = 2(x^2 - 2x + 2)$
$f(x) = -(x^2 - 2x + 2)$


Minimum point
$\mathbf{x}^* = (2.5, -1.5)$
$f(\mathbf{x}^*) = 4.75$
4.9
5.0
6.0
8.0
10.0


Local minima: B and C
Local maximum: A
B
A
C

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: $\mathbf{H}$ must be at least positive semidefinite

Second-order necessary condition for a local maximum: $\mathbf{H}$ must be at least negative semidefinite

Second-order sufficient condition for a local minimum: $\mathbf{H}$ must be positive definite

Second-order sufficient condition for a local maximum: $\mathbf{H}$ must be negative definite

제약최적설계 문제

- 등호제약조건과 부등호제약조건을 만족하면서 목적함수를 최소화하는 설계변수벡터를 찾는 것

$$
\begin{aligned}
& f(\boldsymbol{x}) = f(x_1, \ldots, x_n) \\
& h_j(\boldsymbol{x}) = 0; \quad j = 1, \ldots, p \\
& g_i(\boldsymbol{x}) \le 0; \quad i = 1, \ldots, m
\end{aligned}
$$

- 제약함수들이 최적해를 구하는데 결정적인 역할
  - 해가 존재하지 않을 수도 있음: 과제약 (overconstrained)
- Equality constraints are always active for any feasible design, whereas an inequality constraint may not be active at a feasible point

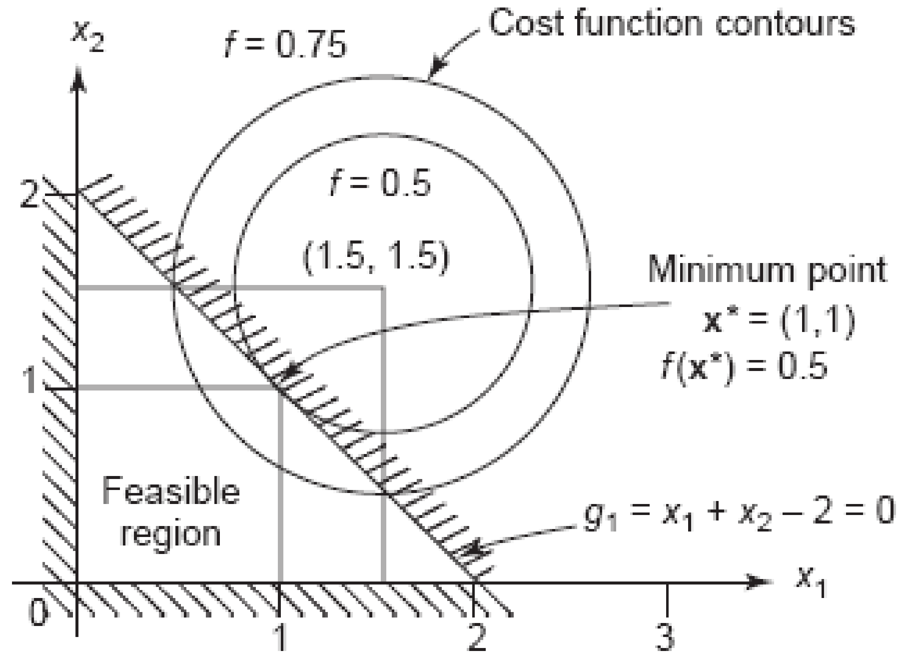
Example 4.24

- 다음 함수를 최소화하라.

$$f(\boldsymbol{x}) = (x_1 - 1.5)^2 + (x_2 - 1.5)^2$$

여기서 제약조건은 다음과 같다.

$$g_1(\boldsymbol{x}) = x_1 + x_2 - 2 \le 0$$
$$g_2(\boldsymbol{x}) = -x_1 \le 0$$
$$g_3(\boldsymbol{x}) = -x_2 \le 0$$



필요조건: 등호제약조건

- 정칙점 (regular point)
  - 등호제약조건을 만족하고 모든 제약함수의 경사도벡터들이 일차독립
  - 일차독립: 두개의 경사도벡터가 서로 평행하지 않고 어떤 경사도벡터도 다른 경사도벡터들의 선형결합으로 표현할 수 없다는 것을 의미

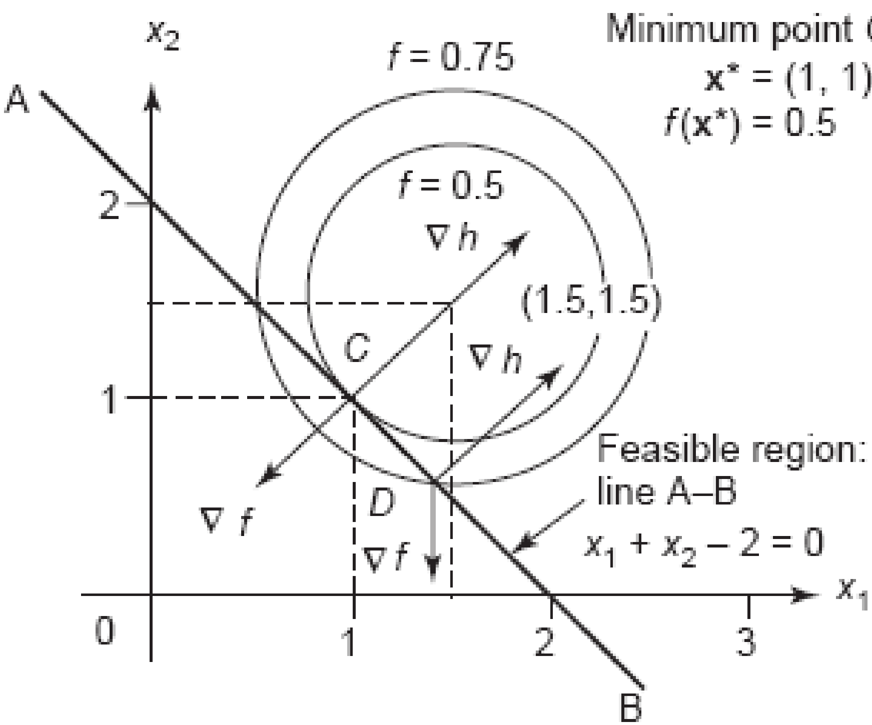
- 라그란지승수 (Lagrange multiplier)
  - 각각의 제약조건에 대응하는 승수 (scalar multiplier)
  - 목적함수나 제약함수의 형태에 따라 좌우

Example 4.27

- 다음 함수를 최소화하는 x_1 과 x_2 를 구하라.

$$f(x_1, x_2) = (x_1 - 1.5)^2 + (x_2 - 1.5)^2$$
$$h(x_1, x_2) = x_1 + x_2 - 2 = 0$$

$$\textit{Minimize} \quad f(x_1, x_2)$$
$$\textit{subject to} \quad h(x_1, x_2) = 0 \rightarrow x_2 = \phi(x_1)$$



등호제약조건이 양함수가 아닌 경우

$$\frac{df(x_1,x_2)}{dx_1} = \frac{\partial f(x_1,x_2)}{\partial x_1} + \frac{\partial f(x_1,x_2)}{\partial x_2}\frac{dx_2}{dx_1} = 0$$

$$\rightarrow \frac{\partial f(x_1^*,x_2^*)}{\partial x_1} + \frac{\partial f(x_1^*,x_2^*)}{\partial x_2}\frac{d\phi}{dx_1} = 0 \quad @\,\text{optimum}$$

$$\frac{dh(x_1^*,x_2^*)}{dx_1} = \frac{\partial h(x_1^*,x_2^*)}{\partial x_1} + \frac{\partial h(x_1^*,x_2^*)}{\partial x_2}\frac{d\phi}{dx_1} = 0 \rightarrow \frac{d\phi}{dx_1} = -\frac{\partial h(x_1^*,x_2^*)/\partial x_1}{\partial h(x_1^*,x_2^*)/\partial x_2}$$

$$\frac{\partial f(x_1^*,x_2^*)}{\partial x_1} - \frac{\partial f(x_1^*,x_2^*)}{\partial x_2}\left[\frac{\partial h(x_1^*,x_2^*)/\partial x_1}{\partial h(x_1^*,x_2^*)/\partial x_2}\right] = 0$$

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

$$\begin{cases} \dfrac{\partial f(x_1^*,x_2^*)}{\partial x_1} + v\dfrac{\partial h(x_1^*,x_2^*)}{\partial x_1} = 0 \\ \dfrac{\partial f(x_1^*,x_2^*)}{\partial x_2} + v\dfrac{\partial h(x_1^*,x_2^*)}{\partial x_2} = 0 \end{cases}$$

$$h(x_1,x_2) = 0$$

라그란지승수의 기하학적 의미

$$L(x_1, x_2, v) = f(x_1, x_2) + vh(x_1, x_2)$$

$$\left.\begin{aligned} \frac{\partial L(x_1^*, x_2^*)}{\partial x_1} = 0 \\ \frac{\partial L(x_1^*, x_2^*)}{\partial x_2} = 0 \end{aligned}\right\} \rightarrow \nabla L(x_1^*, x_2^*) = \nabla f(\mathbf{x}^*) + v\nabla h(\mathbf{x}^*) = 0$$

$$\boxed{\nabla f(\mathbf{x}^*) = -v\nabla h(\mathbf{x}^*)}$$

- 후보최적점에서 목적함수 및 제약함수들의 경사도벡터들은 동일 작용선상에 있고 서로 비례함
- 라그란지승수는 비례상수 (제약을 가하기 위해 필요한 힘으로 해석 가능)
- 등호제약조건에 대한 라그란지승수의 부호는 제약이 없음

라그란지승수정리

– 등호제약조건으로 $h_j(\mathbf{x})=0;\ j=1,\ldots,p$ 가 있는 $f(x)$ 의 최소화문제를 고려해 보자. x* 를 이 문제의 국부적 최소인 정칙점이라고 하면 다음을 만족하는 라그란지승수 $v_j^*,\ j=1,\ldots,p$ 가 존재한다.

$$\frac{df(x^*)}{dx_i}+\sum_{j=1}^{p}v_j^*\frac{dh_j(x^*)}{dx_i}=0 \rightarrow \frac{df(x^*)}{dx_i}=-\sum_{j=1}^{p}v_j^*\frac{dh_j(x^*)}{dx_i};\qquad i=1,\ldots,n$$

$$h_j(x^*)=0;\qquad j=1,\ldots,p$$

– 후보최적점에서 목적함수의 경사도벡터는 제약함수의 경사도벡터들의 선형결합

$$L(x,v)=f(x)+\sum_{j=1}^{p}v_jh_j(x)=f(x)+v^Th(x)$$

$$\rightarrow\begin{cases}\nabla L(x^*,v^*)=0 \quad \text{or} \quad \dfrac{\partial L(x^*,v^*)}{\partial x_i}=0; \qquad i=1,\ldots,n\\[2ex] \dfrac{\partial L(x^*,v^*)}{\partial v_j}=h_j(x^*)=0; \qquad j=1,\ldots,p\end{cases}$$

Example 4.25 Cylindrical Tank Design

- Section 2.8 → Example 4.21

$$\left.\begin{aligned} &\underset{R,l}{\text{minimize}} \;\; \bar{f} = R^2 + RH \\ &\text{subject to} \;\; h = \pi R^2 H - V = 0 \end{aligned}\right\} \rightarrow R^* = \left(\frac{V}{2\pi}\right)^{1/3}, V^* = \left(\frac{4V}{\pi}\right)^{1/3}, \bar{f}^* = 3\left(\frac{V}{2\pi}\right)^{1/3}$$

Example 4.26~28 Role of Inequalities

– status of the inequality constraint (active or inactive) must be determined as a part of the solution for the optimization problem

$$\begin{cases} \underset{\mathbf{x}}{\text{Minimize}} \; f(x_1, x_2) = (x_1 - 1.5)^2 + (x_2 - 1.5)^2 \\ \underset{\mathbf{x}}{\text{Minimize}} \; f(x_1, x_2) = (x_1 - 0.5)^2 + (x_2 - 0.5)^2 \end{cases}$$

$$\begin{aligned} \text{subject to } \; & g_1(\mathbf{x}) = x_1 + x_2 - 2 \le 0 \\ & g_2(\mathbf{x}) = -x_1 \le 0 \\ & g_3(\mathbf{x}) = -x_2 \le 0 \end{aligned}$$

$$\underset{\mathbf{x}}{\text{Minimize}} \; f(x_1, x_2) = (x_1 - 2)^2 + (x_2 - 2)^2$$

$$\begin{aligned} \text{subject to } \; & g_1(\mathbf{x}) = x_1 + x_2 - 2 \le 0 \\ & g_2(\mathbf{x}) = -x_1 + x_2 + 3 \le 0 \\ & g_3(\mathbf{x}) = -x_1 \le 0 \\ & g_4(\mathbf{x}) = -x_2 \le 0 \end{aligned}$$

필요조건: 부등호제약조건

- 완화변수(slack variable)를 더하여 등호제약조건으로 변환

$$g_i(\boldsymbol{x}) \le 0 \rightarrow g_i(\boldsymbol{x}) + s_i^2 = 0 \quad i = 1, \ldots, m$$

$$\text{slack variable} : s_i^2 \text{ (why? avoid additional constraints } s_i \ge 0)$$

$$\begin{cases} s_i^2 = 0 : \text{equality} \rightarrow \text{active (tight) constraint} \\ s_i^2 > 0 : \text{inequality} \rightarrow \text{inactive constraint} \end{cases}$$

- "≤ type" 제약조건의 라그란지승수에 대한 추가적인 필요조건
  - u_j^* 는 j번째 부등호제약조건에 대한 라그란지승수: $u_j^* \ge 0 \ \ (j = 1, \ldots, m)$

	minimize	maximize
$g_i(\boldsymbol{x}) \le 0$	$u_i^* \ge 0$	$u_i^* \le 0$
$g_i(\boldsymbol{x}) \ge 0$	$u_i^* \le 0$	$u_i^* \ge 0$

Example 4.29

$$\text{Minimize } f(x_1, x_2) = (x_1 - 1.5)^2 + (x_2 - 1.5)^2$$

$$\text{subject to } g(\mathbf{x}) = x_1 + x_2 - 2 \le 0$$

Karush-Kuhn-Tucker necessary conditions (1)

- *x** 가 제약집합내의 정칙점이고, 다음의 제약조건 하에서 함수 *f(x)*의 국부적 최소점이라 하자.

$$h_i(\boldsymbol{x}) = 0;\quad i = 1,\ldots,p \quad \text{and} \quad g_j(\boldsymbol{x}) \le 0;\quad j = 1,\ldots,m$$

- 이 문제의 라그란지함수를 다음과 같이 정의한다.

$$L(\boldsymbol{x},\boldsymbol{v},\boldsymbol{u},\boldsymbol{s}) = f(\boldsymbol{x}) + \sum_{i=1}^{p} v_i h_i(\boldsymbol{x}) + \sum_{i=1}^{m} u_i\left[g_i(\boldsymbol{x}) + s_i^{\,2}\right] = f(\boldsymbol{x}) + \boldsymbol{v}^T\boldsymbol{h}(\boldsymbol{x}) + \boldsymbol{u}^T\left[\boldsymbol{g}(\boldsymbol{x}) + \mathrm{s}^2\right]$$

- 그러면 다음 조건을 만족하는 라그란지승수 v* 와 u* 가 존재한다.

$$\begin{cases} \dfrac{\partial L}{\partial x_j} \equiv \dfrac{\partial f}{\partial x_j} + \sum_{i=1}^{p} v_i^{\,*}\dfrac{\partial h_i}{\partial x_j} + \sum_{i=1}^{m} u_i^{\,*}\dfrac{\partial g_i}{\partial x_j} = 0; & j = 1,\ldots,n \\ h_i(\mathbf{x}^*) = 0; \quad i = 1,\ldots,p \\ g_i(\mathbf{x}^*) + s_i^{\,2} = 0; \quad i = 1,\ldots,m \\ u_i^{\,*} s_i = 0; \quad i = 1,\ldots,m \ \text{(switching conditions)} \\ u_i^{\,*} \ge 0; \quad i = 1,\ldots,m \end{cases}$$

Karush-Kuhn-Tucker necessary conditions (2)

- 1차 필요조건
  - 어떤 주어진 점에 대한 최적성을 점검 / 후보최적점을 결정
- 기하학적 의미
  - 목적함수의 음의 경사도벡터방향이 제약함수의 경사도벡터들의 선형결합이며, 라그란지승수가 선형결합의 상수로서 사용

$$-\frac{\partial f}{\partial x_j} = \sum_{i=1}^{p} v_i^* \frac{\partial h_i}{\partial x_j} + \sum_{i=1}^{m} u_i^* \frac{\partial g_i}{\partial x_j}; \quad j = 1, \ldots, n$$

- 미지수의 개수: x, u, s, v (n+2m+p) = # of eqns
- 전환조건(switching condition) 또는 보충완화조건 (complementary slackness condition)

$$\begin{cases} g_i\left(\mathbf{x}^*\right) < 0 \text{ (inactive, } s_i^2 > 0) \rightarrow u_i^* = 0 \\ g_i\left(\mathbf{x}^*\right) = 0 \text{ (active, } s_i^2 = 0) \rightarrow u_i^* \geq 0 \end{cases}$$

Karush-Kuhn-Tucker necessary conditions (3)

- K-T conditions are *not applicable* at the points that are not *regular*.
- Any point that *does not satisfy* K-T conditions *cannot be a local minimum* unless it is an irregular points.
- The points satisfying K-T conditions can be *constrained or unconstrained*.
- If there are equality constraints and no inequalities are active, then the points satisfying K-T conditions are *only stationary*.
- If some inequality constraints are active and their multipliers are positive, then the points satisfying K-T conditions cannot be local maxima for the cost function.
- The value of the *Lagrange multiplier* for each constraint depends on the functional form for the constraint.
  - Optimum solution ? / Lagrange multiplier ?

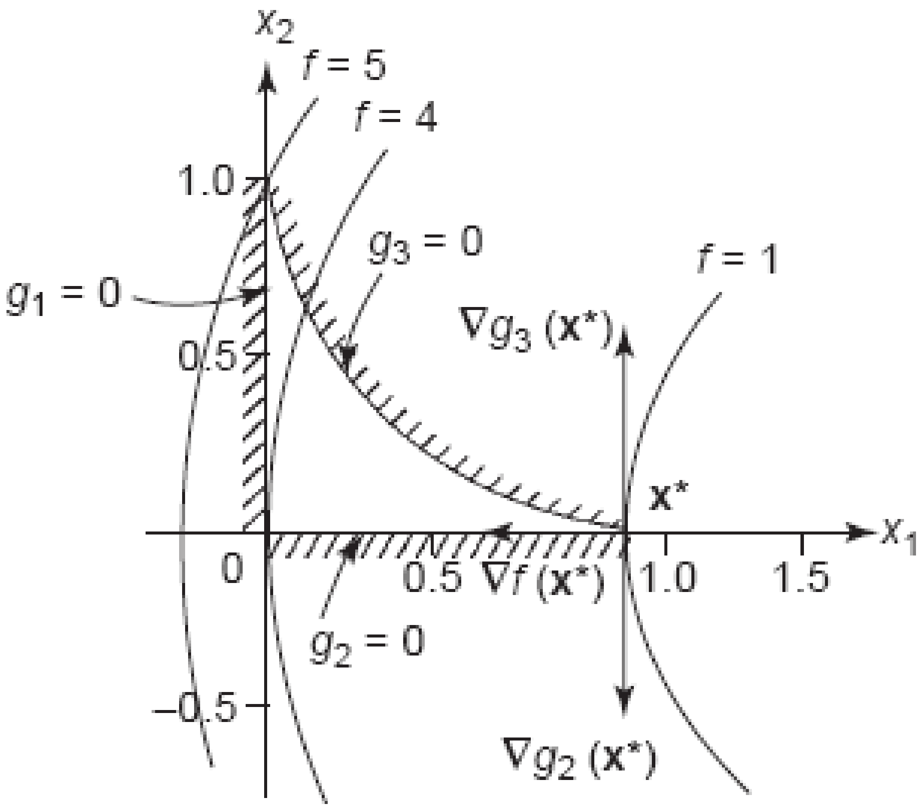
$$\text{(i)}\ \frac{x}{y} - 10 \le 0 \left(y > 0\right)\ \ \text{(ii)}\ x - 10y \le 0\ \ \text{(iii)}\ \frac{0.1x}{y} - 1 \le 0$$

Example 5.3

– Necessary conditions are applicable only if the assumption for regularity of x* is satisfied.
  - Gradients of all the active constraints @ x* is linearly independent

$$\begin{aligned}
&\text{Minimize} \quad f(x_1, x_2) = x_1^{\,2} + x_2^{\,2} - 4x_1 + 4 \\
&\text{subject to} \quad g_1 = -x_1 \le 0 \\
&\qquad\qquad\quad\; g_2 = -x_2 \le 0 \\
&\qquad\qquad\quad\; g_3 = x_2 - (1 - x_1)^3 \le 0
\end{aligned}$$

Check if the minimum point (1,0) satisfies K-T conditions.



Kuhn-Tucker 필요조건의 변형

- Without slack variables

$$\begin{cases} g_i(\mathbf{x}^*) + s_i^2 = 0 \rightarrow s_i^2 = -g_i(\mathbf{x}^*) \geq 0 \rightarrow g_i(\mathbf{x}^*) \leq 0 \\ u_i^* s_i = 0 \rightarrow u_i^* s_i^2 = 0 \rightarrow u_i^* g_i(\mathbf{x}^*) = 0 \end{cases}$$

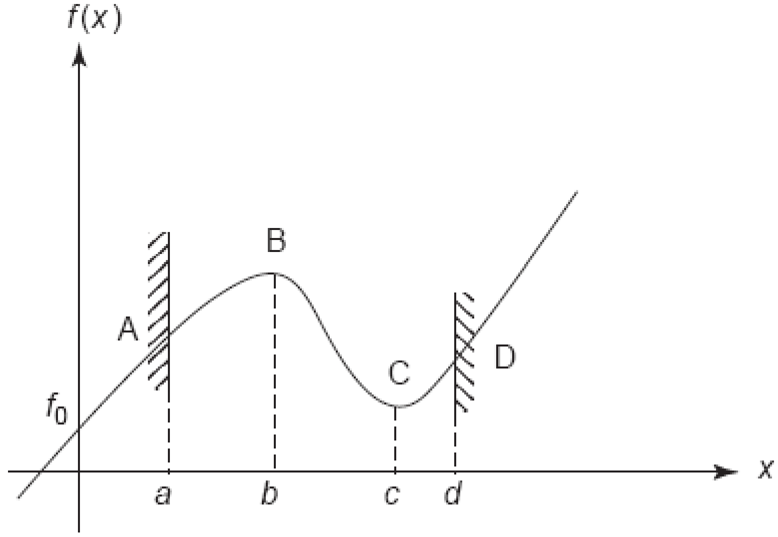
$$L(\mathbf{x}, \mathbf{v}, \mathbf{u}) = f(\mathbf{x}) + \sum_{i=1}^{p} v_i h_i(\mathbf{x}) + \sum_{i=1}^{m} u_i g_i(\mathbf{x}) = f(\mathbf{x}) + \mathbf{v}^T \mathbf{h}(\mathbf{x}) + \mathbf{u}^T \mathbf{g}(\mathbf{x})$$

$$\begin{cases} \dfrac{\partial L}{\partial x_j} \equiv \dfrac{\partial f}{\partial x_j} + \sum_{i=1}^{p} v_i \dfrac{\partial h_i}{\partial x_j} + \sum_{i=1}^{m} u_i \dfrac{\partial g_i}{\partial x_j} = 0; \quad j = 1, \ldots, n \\ h_i(\mathbf{x}^*) = 0; \quad i = 1, \ldots, p \\ g_i(\mathbf{x}^*) \leq 0; \quad i = 1, \ldots, m \\ u_i^* g_i(\mathbf{x}^*) = 0; \quad i = 1, \ldots, m \ \text{(switching conditions)} \\ u_i^* \geq 0; \quad i = 1, \ldots, m \end{cases}$$

Example 4.30

$$\text{Minimize} \quad f(\boldsymbol{x}) = \frac{1}{3}x^3 - \frac{1}{2}(b+c)x^2 + bcx + f_0$$

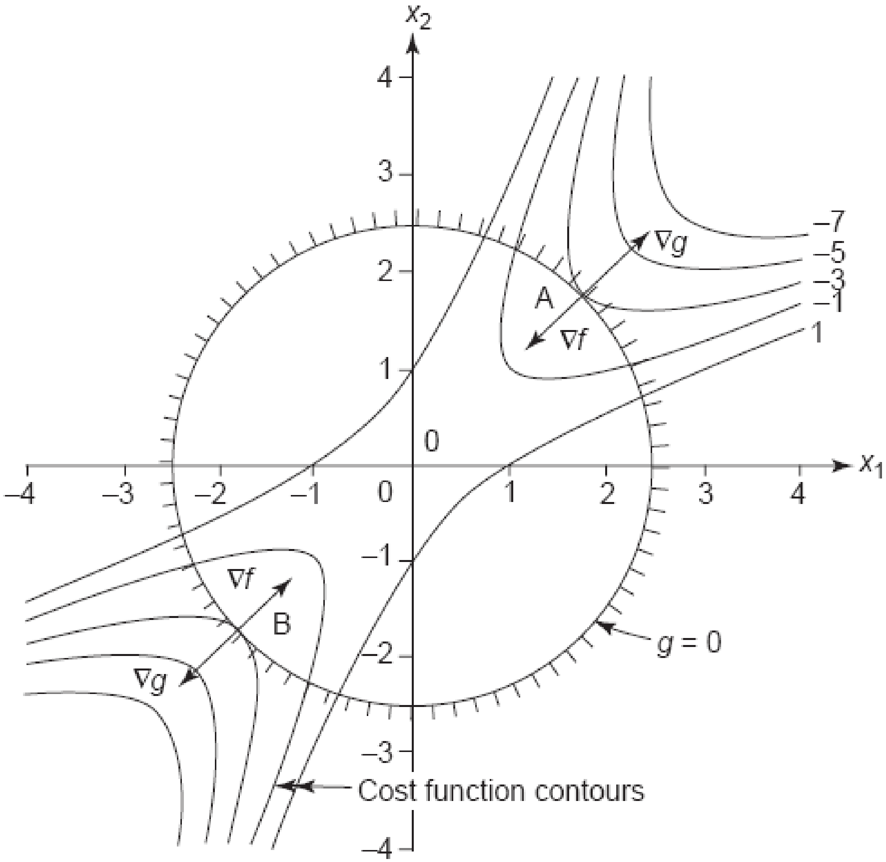
$$\text{subject to} \quad a \le x \le d \quad (0 < a < b < c < d \text{ and } f_0 \text{ are constants})$$



Example 4.31

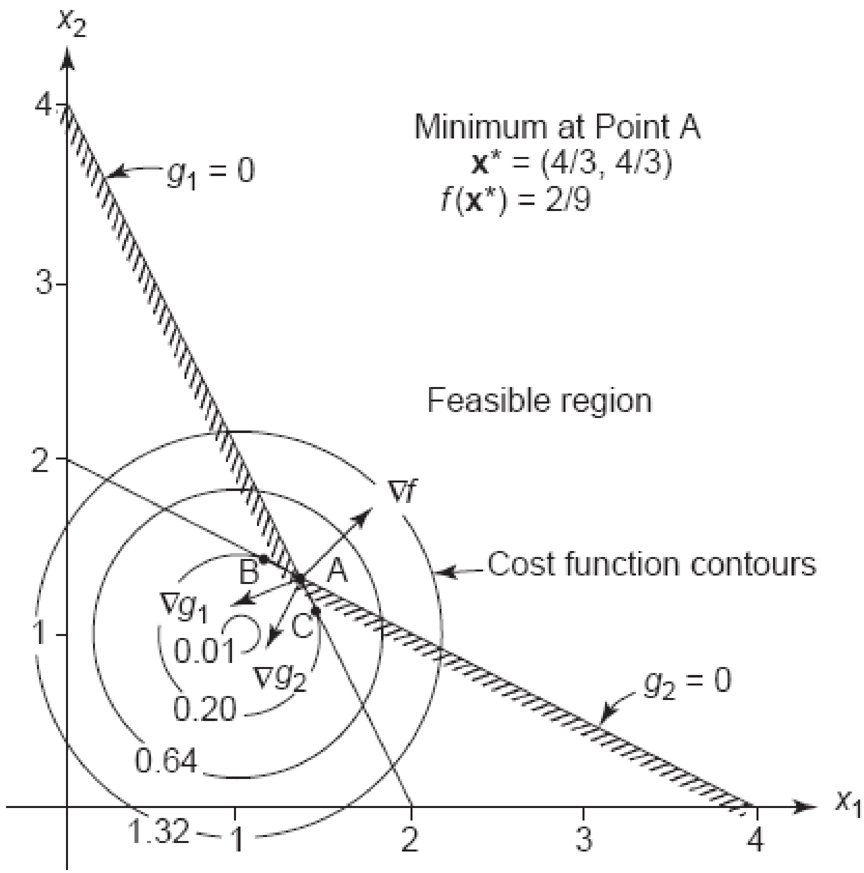
$$\text{Minimize } f(\mathbf{x}) = x_1^{\ 2} + x_2^{\ 2} - 3x_1x_2$$

$$\text{subject to } g = x_1^{\ 2} + x_2^{\ 2} - 6 \le 0$$



Example 4.32

$$\text{Minimize } \; f(\mathbf{x}) = x_1^{\,2} + x_2^{\,2} - 2x_1 - 2x_2 + 2$$
$$\text{subject to } \; g_1 = -2x_1 - x_2 + 4 \le 0$$
$$g_2 = -x_1 - 2x_2 + 4 \le 0$$



Summary of the KKT Solution Approach

- *to check* whether a given point is a candidate minimum
  - the point must be feasible
  - the gradient of the Lagrangian with respect to the design variables must be zero
  - the Lagrange multipliers for the inequality constraints must be nonnegative
- *to find* candidate minimum points
  - Several cases defined by the switching conditions must be considered and solved
    - Check all inequality constraints for feasibility
    - Calculate all of the Lagrange multipliers for each solution point
    - Ensure that the Lagrange multipliers for all of the inequality constraints are nonnegative

Limitation of the KKT Solution Approach

- addition of an inequality to the problem formulation doubles the number of KKT solution cases

No.	Cases	Active constraints
1	$u_1 = 0, u_2 = 0, u_3 = 0, u_4 = 0$	No inequality active
2	$s_1 = 0, u_2 = 0, u_3 = 0, u_4 = 0$	*One* active inequality: $g_1 = 0$
3	$u_1 = 0, s_2 = 0, u_3 = 0, u_4 = 0$	*One* active inequality: $g_2 = 0$
4	$u_1 = 0, u_2 = 0, s_3 = 0, u_4 = 0$	*One* active inequality: $g_3 = 0$
5	$u_1 = 0, u_2 = 0, u_3 = 0, s_4 = 0$	*One* active inequality: $g_4 = 0$
6	$s_1 = 0, s_2 = 0, u_3 = 0, u_4 = 0$	*Two* active inequalities: $g_1 = 0, g_2 = 0$
7	$u_1 = 0, s_2 = 0, s_3 = 0, u_4 = 0$	*Two* active inequalities: $g_2 = 0, g_3 = 0$
8	$u_1 = 0, u_2 = 0, s_3 = 0, s_4 = 0$	*Two* active inequalities: $g_3 = 0, g_4 = 0$
9	$s_1 = 0, u_2 = 0, u_3 = 0, s_4 = 0$	*Two* active inequalities: $g_1 = 0, g_4 = 0$
10	$s_1 = 0, u_2 = 0, s_3 = 0, u_4 = 0$	*Two* active inequalities: $g_1 = 0, g_3 = 0$
11	$u_1 = 0, s_2 = 0, u_3 = 0, s_4 = 0$	*Two* active inequalities: $g_2 = 0, g_4 = 0$
12	$s_1 = 0, s_2 = 0, s_3 = 0, u_4 = 0$	*Three* active inequalities: $g_1 = 0, g_2 = 0, g_3 = 0$
13	$u_1 = 0, s_2 = 0, s_3 = 0, s_4 = 0$	*Three* active inequalities: $g_2 = 0, g_3 = 0, g_4 = 0$
14	$s_1 = 0, u_2 = 0, s_3 = 0, s_4 = 0$	*Three* active inequalities: $g_1 = 0, g_3 = 0, g_4 = 0$
15	$s_1 = 0, s_2 = 0, u_3 = 0, s_4 = 0$	*Three* active inequalities: $g_1 = 0, g_2 = 0, g_4 = 0$
16	$s_1 = 0, s_2 = 0, s_3 = 0, s_4 = 0$	All *four* inequalities active

라그란지승수의 물리적 의미

- 후최적성해석(Post-optimality analysis) 또는 민감도해석(Sensitivity analysis)
  - 최적설계문제의 매개변수를 변화시켰을 때 최적해의 변화에 대한 고찰
- 제약한계값을 0으로부터 변화시켰을 때 목적함수의 최적해에는 어떤 변화?
  - 라그란지승수 (v*, u*)가 이러한 민감도문제에 대한 해답을 제공
- 왜 "≤ type" 제약조건에 대한 라그란지승수가 음수가 아니어야 하는가?
- 제약조건을 완화(relaxation)함에 따라 얻어지는 이점이나 속박(tightening)에 따른 불리한 점
- 목적함수와 제약함수를 축적화(scaling)했을때 라그란지승수의 변화?

제약한계변화의 영향

- b_i 와 e_j 는 0근처에서 매우 작은 변화량

$$\boldsymbol{x}^* = \boldsymbol{x}^*(\boldsymbol{b},\boldsymbol{e}), f = f(\boldsymbol{b},\boldsymbol{e})$$

$$h_i(\boldsymbol{x}) = b_i; \;\; i = 1,\ldots,p \;\; \text{and} \;\; g_j(\boldsymbol{x}) \le e_j; \;\; j = 1,\ldots,m$$

- 제약함수의 민감도 정리
  - v_i*, u_j*: satisfying both necessary and sufficient conditions

$$\begin{cases} \dfrac{\partial L}{\partial b_i} \equiv \dfrac{\partial f}{\partial b_i} + \displaystyle\sum_{i=1}^{p} v_i \dfrac{\partial h_i}{\partial b_i} + \sum_{j=1}^{m} u_j \dfrac{\partial g_j}{\partial b_i} = 0 \rightarrow \dfrac{\partial f(\boldsymbol{x}^*(0,0))}{\partial b_i} = -v_i^*; \quad i = 1,\ldots,p \\ \dfrac{\partial L}{\partial e_j} \equiv \dfrac{\partial f}{\partial e_j} + \displaystyle\sum_{i=1}^{p} v_i \dfrac{\partial h_i}{\partial e_j} + \sum_{j=1}^{m} u_j \dfrac{\partial g_j}{\partial e_j} = 0 \rightarrow \dfrac{\partial f(\boldsymbol{x}^*(0,0))}{\partial e_j} = -u_j^*; \quad j = 1,\ldots,m \end{cases}$$

$$f(b_i, e_j) = f(0,0) + \frac{\partial f(0,0)}{\partial b_i} b_i + \frac{\partial f(0,0)}{\partial e_j} e_j = f(0,0) - v_i^* b_i - u_j^* e_j$$

$$\Delta f = f(b_i, e_j) - f(0,0) = -v_i^* b_i - u_j^* e_j$$

$$\Delta f = -\sum_i v_i^* b_i - \sum_j u_j^* e_j$$

Example 4.33 ($\leftarrow$4.31)

- Nonnegativity of Lagrange multipliers

relax an inequality constraint $\left(g_j \le 0\right): e_j > 0$

$\rightarrow$ feasible set for the design problem expands

$\rightarrow$ expect the optimum cost function to reduce further or at the most remain unchanged

if $u_j^* < 0$, then $\Delta f = -u_j^* e_j > 0$ (contradiction!)

$\therefore$ Lagrange multiplier corresponding to a "$\le$ *type*" constraint must be nonnegative.

$$\begin{aligned} &\textit{Minimize} \quad f\left(x_1, x_2\right) = x_1^2 + x_2^2 - 3x_1x_2 \\ &\textit{subject to} \quad g\left(x_1, x_2\right) = x_1^2 + x_2^2 - 6 \le 0 \\ &\Rightarrow x_1^* = x_2^* = \sqrt{3}, \quad u^* = \frac{1}{2}, \quad f\left(\mathbf{x}^*\right) = -3 \end{aligned}$$

$$\begin{cases} e = 1 \text{ (i.e., radius of circle } = \sqrt{6} \rightarrow \sqrt{7}\text{) ?} \\ f\left(0,1\right) = f\left(0,0\right) - u^* e = -3 - \left(0.5\right)\left(1\right) = -3.5 \\ e = -1 \text{ (smaller feasible region) } f\left(0,-1\right) = -2.5 \end{cases}$$

Effect of scaling a cost function

- No change on the optimum point
- Change in the optimum value for the cost function

$$\bar{f}(\mathbf{x}) = Kf(\mathbf{x}) \quad \text{where} \quad K > 0$$

$$\begin{cases} \bar{v}_i^* = Kv_i^*; & i = 1, \ldots, p \\ \bar{u}_j^* = Ku_j^*; & j = 1, \ldots, m \end{cases}$$

$$L = K\left(x_1^2 + x_2^2 - 3x_1x_2\right) + \bar{u}\left(x_1^2 + x_2^2 - 6 + \bar{s}^2\right)$$

$$\left.\begin{aligned} &\frac{\partial L}{\partial x_1} = 2Kx_1 - 3Kx_2 + 2\bar{u}x_1 = 0 \\ &\frac{\partial L}{\partial x_2} = 2Kx_2 - 3Kx_1 + 2\bar{u}x_2 = 0 \\ &x_1^2 + x_2^2 - 6 + \bar{s}^2 = 0 \\ &\bar{u}\bar{s} = 0 \\ &\bar{u} \ge 0 \end{aligned}\right\} \rightarrow \begin{cases} x_1^* = x_2^* = \sqrt{3}, \quad \bar{u}^* = \dfrac{K}{2}, \quad \bar{f}(\mathbf{x}^*) = -3K \\ x_1^* = x_2^* = -\sqrt{3}, \quad \bar{u}^* = \dfrac{K}{2}, \quad \bar{f}(\mathbf{x}^*) = -3K \\ \bar{u}^* = Ku^* \end{cases}$$

Effect of scaling a constraint

- No change on the constraint boundary (no effect on the optimum solution)
- Change in the Lagrange multiplier

$$\begin{cases} \bar{v}_i^* = v_i^* / P_i ; & i = 1, \ldots, p \\ \bar{u}_j^* = u_j^* / M_j ; & j = 1, \ldots, m \quad \text{where} \quad M_j > 0 \end{cases}$$

$$L = \left(x_1^2 + x_2^2 - 3x_1x_2\right) + \bar{u}\left[M\left(x_1^2 + x_2^2 - 6\right) + \bar{s}^2\right]$$

$$\left.\begin{aligned} &\frac{\partial L}{\partial x_1} = 2x_1 - 3x_2 + 2\bar{u}Mx_1 = 0 \\ &\frac{\partial L}{\partial x_2} = 2x_2 - 3x_1 + 2\bar{u}Mx_2 = 0 \\ &M\left(x_1^2 + x_2^2 - 6\right) + \bar{s}^2 = 0 \\ &\bar{u}\bar{s} = 0 \\ &\bar{u} \ge 0 \end{aligned}\right\} \rightarrow \begin{cases} x_1^* = x_2^* = \sqrt{3}, \quad \bar{u}^* = \dfrac{1}{2M}, \quad \bar{f}\left(\mathbf{x}^*\right) = -3 \\ x_1^* = x_2^* = -\sqrt{3}, \quad \bar{u}^* = \dfrac{1}{2M}, \quad \bar{f}\left(\mathbf{x}^*\right) = -3 \\ \bar{u}^* = \dfrac{u^*}{M} \end{cases}$$

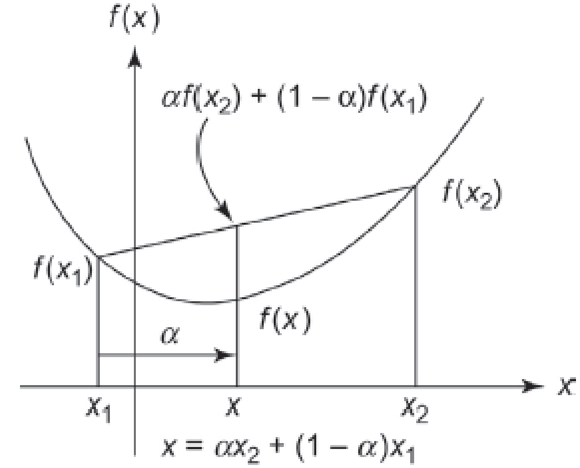
Global Optimality

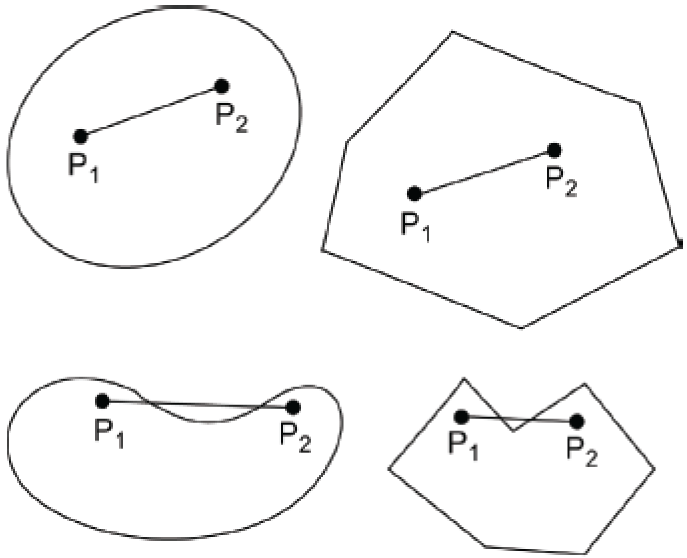
- Question of *global optimum*
  - Weierstrass theorem (→ exhaustive search)
    - Cost function is continuous on a closed and bounded feasible region
  - Showing the optimization problem is convex
- Convex set *S*
  - If P_1 and P_2 are any points in *S*, then the entire line segment P_1 - P_2 is also in *S* $\left[\mathbf{x} = \alpha\mathbf{x}^{(2)} + (1-\alpha)\mathbf{x}^{(1)}; \quad 0 \le \alpha \le 1\right]$
- Convex functions

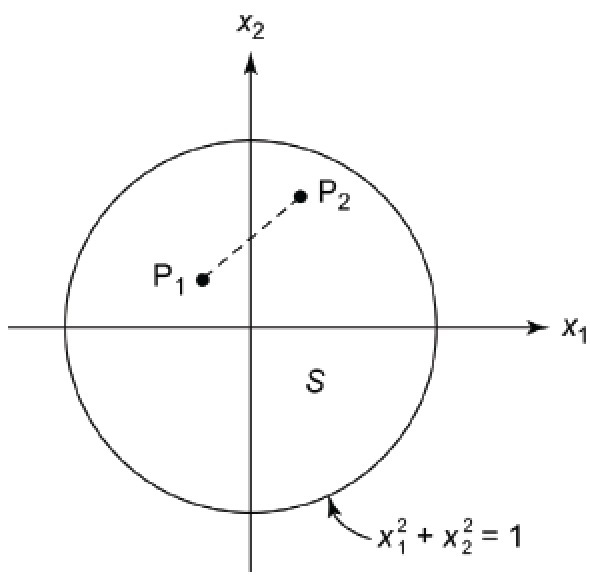
$$f\left(\alpha\mathbf{x}^{(2)} + (1-\alpha)\mathbf{x}^{(1)}\right) \le \alpha f\left(\mathbf{x}^{(2)}\right) + (1-\alpha) f\left(\mathbf{x}^{(1)}\right)$$

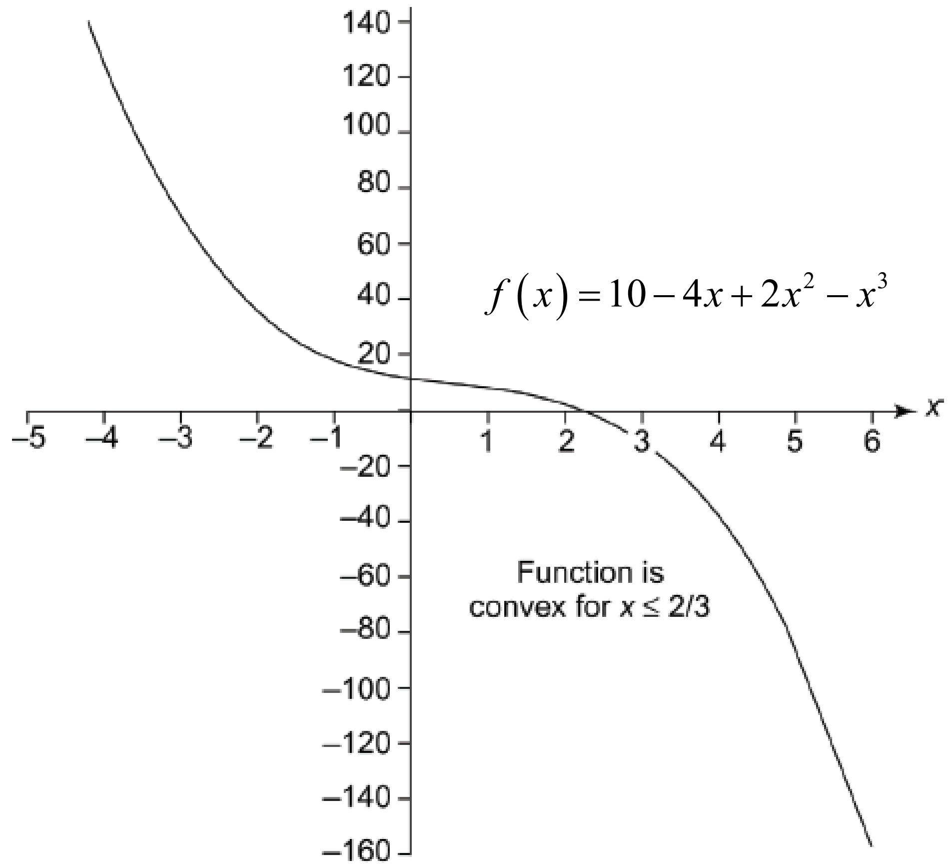
  - Check : *iff* Hessian matrix of a function is positive semidefinite or positive definite at all points in the set S
    - Hessian matrix is positive definite $\rightarrow$ $\leftarrow$(x) strictly convex

Convexity


f(x)
αf(x2) + (1 − α)f(x1)
f(x2)
f(x1)
f(x)
α
x
x1
x
x2
x = αx2 + (1 − α)x1


P1
P2
P1
P2
P1
P2
P1
P2


x2
P2
P1
x1
S
x1² + x2² = 1


140
120
100
80
60
40
20
–20
–40
–60
–80
–100
–120
–140
–160
–5
–4
–3
–2
–1
1
2
3
4
5
6
x
f(x) = 10 − 4x + 2x² − x³
Function is
convex for x ≤ 2/3

Convex Programming Problem

- Constraint set *S* is convex and cost function is also convex over *S*
  - Nonlinear equality constraint set $\rightarrow$ always nonconvex
  - Linear equality/inequality constraint set $\rightarrow$ always convex
- KKT necessary conditions are also sufficient
  - (theorem 4.9)

$$\left\langle S = \left\{ \boldsymbol{x} \middle| h_j(\boldsymbol{x}) = 0;\ j = 1, \ldots, p;\quad g_i(\boldsymbol{x}) \le 0; i = 1, \ldots, m \right\} \text{is a convex set} \right\rangle$$

$$\rightarrow (\times) \leftarrow (\mathrm{o}) \left\langle \text{function } g_i \text{ are convex and } h_j \text{ are linear} \right\rangle$$

  - (theorem 4.10) Any local minimum is also a global minimum
    - Proof ?
    - Convexity check failure $\rightarrow$(x) no global minimum point

Check for Convexity (1)

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

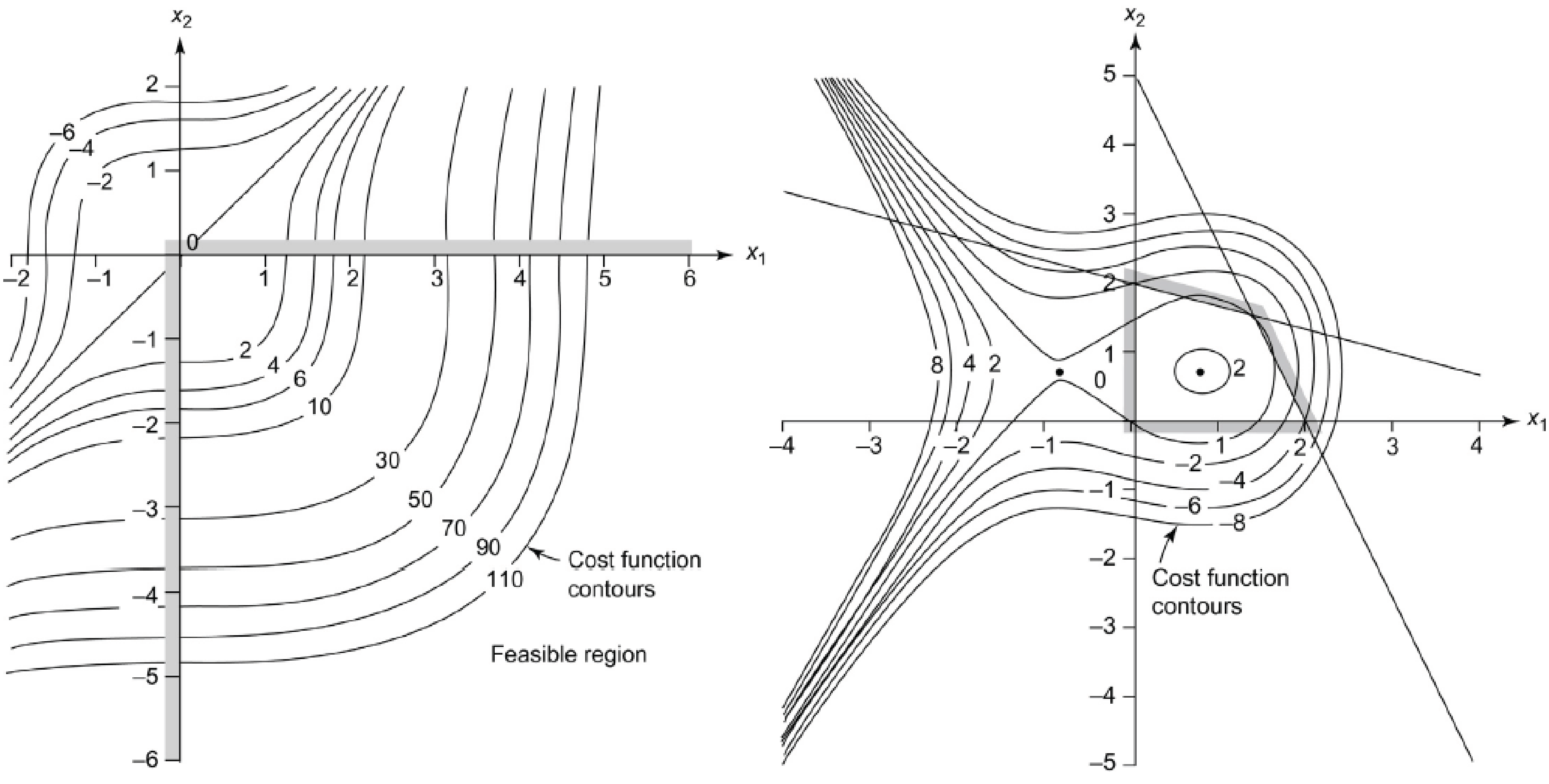
[2] $f(x) = 10 - 4x + 2x^2 - x^3$

[3] $\begin{cases} \min f(\boldsymbol{x}) = x_1^{\ 3} - x_2^{\ 3} \\ \text{s.t.} \quad x_1 \geq 0, x_2 \leq 0 \end{cases}$

[4] $\begin{cases} \min f(\boldsymbol{x}) = 2x_1 + 3x_2 - x_1^{\ 3} - 2x_2^{\ 2} \\ \text{s.t.} \quad x_1 + 3x_2 \leq 6 \\ \qquad 5x_1 + 2x_2 \leq 10 \\ \qquad x_1, x_2 \geq 0 \end{cases}$

[5] $\begin{cases} \min f(\boldsymbol{x}) = 9x_1^{\ 2} - 18x_1x_2 + 13x_2^{\ 2} - 4 \\ \text{s.t.} \quad x_1^{\ 2} + x_2^{\ 2} + 2x_1 \geq 16 \end{cases}$

Check for Convexity (2)



Transformation of a constraint

- Form of function: convex $\leftrightarrow$ nonconvex
- Convexity of the feasible region: no change

$$g_1 = \frac{a}{x_1 x_2} - b \le 0 \qquad g_2 = a - b x_1 x_2 \le 0$$

$$\nabla^2 g_1 = \frac{2a}{x_1^2 x_2^2}\begin{bmatrix} x_2/x_1 & 0.5 \\ 0.5 & x_1/x_2 \end{bmatrix} \qquad \nabla^2 g_2 = \begin{bmatrix} 0 & -b \\ -b & 0 \end{bmatrix}$$

$$(\textit{positive definite}) \qquad (\textit{indefinite})$$

- Sufficient Conditions for Convex Programming Problems
  - If *f* (**x**) is a convex cost function defined on a convex feasible set, then the first-order KKT conditions are necessary as well as sufficient for a global minimum

Second-order conditions (1)

- Convex problems
  - First-order K-T conditions are necessary as well as sufficient for a global minimum

$$\begin{aligned}&\textit{Minimize} \;\; f(x_1, x_2) = (x_1 - 1.5)^2 + (x_2 - 1.5)^2 \\ &\textit{subject to} \;\; g(\mathbf{x}) = x_1 + x_2 - 2 \le 0\end{aligned}$$

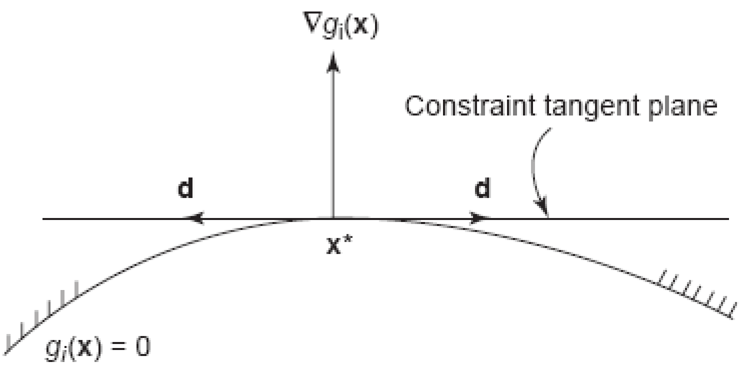
- General problems
  - Let **x*** satisfy the first-order KKT necessary conditions
  - Consider active constraints @**x*** to determine feasible changes ***d***

$$\nabla h_i^T \boldsymbol{d} = 0 \quad \text{and} \quad \nabla g_i^T \boldsymbol{d} = 0$$

  - If the number of active inequality constraints is equal to the number of independent design variables and all other K-T conditions are satisfied, then the candidate point is a local minimum

Second-order conditions (2)

- Necessary condition



for nonzero feasible directions ($\boldsymbol{d} \neq 0$) satisfying

$$\nabla h_i^T \boldsymbol{d} = 0; \quad i = 1, \ldots, p$$

$$\nabla g_i^T \boldsymbol{d} = 0; \quad \text{for all active constraints}$$

$$Q = \boldsymbol{d}^T \nabla^2 \boldsymbol{L}(\boldsymbol{x}^*)\boldsymbol{d} \geq 0 \ \text{ if } \boldsymbol{x}^* \text{ is a local minimum point}$$

- Sufficient condition

for nonzero feasible directions ($\mathbf{d} \neq 0$) satisfying

$$\nabla h_i^T \mathbf{d} = 0; \quad i = 1, \ldots, p$$

$$\nabla g_i^T \mathbf{d} = 0; \quad i = 1, \ldots, m \text{ for active inequalities with } u_i^* > 0$$

$$\nabla g_i^T \mathbf{d} \leq 0; \quad \text{for constraints with } u_i^* = 0$$

if $Q = \mathbf{d}^T \nabla^2 \mathbf{L}\left(\mathbf{x}^*\right)\mathbf{d} > 0$, then $\mathbf{x}^*$ is an isolated local minimum point

Check for Sufficient Conditions

[Example 4.30]

$$\begin{cases} \text{Minimize } f(\mathbf{x}) = \frac{1}{3}x^3 - \frac{1}{2}(b+c)x^2 + bcx + f_0 \\ \text{subject to } a \le x \le d \quad (0 < a < b < c < d \text{ and } f_0 \text{ are constants}) \end{cases}$$

$\rightarrow x = a$

[Example 4.31]

$$\begin{cases} \text{Minimize } f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2 \\ \text{subject to } g = x_1^2 + x_2^2 - 6 \le 0 \end{cases}$$

$\rightarrow (1)\,\mathbf{x}^* = (0,0), u^* = 0 \;\; (2)\,\mathbf{x}^* = (\sqrt{3},\sqrt{3}), u^* = 0.5 \;\; (3)\,\mathbf{x}^* = (-\sqrt{3},-\sqrt{3}), u^* = 0.5$

[Example 4.32]

$$\begin{cases} \text{Minimize } f(\mathbf{x}) = x_1^2 + x_2^2 - 2x_1 - 2x_2 + 2 \\ \text{subject to } g_1 = -2x_1 - x_2 + 4 \le 0 \\ \qquad\qquad\quad g_2 = -x_1 - 2x_2 + 4 \le 0 \end{cases}$$

$\rightarrow \mathbf{x}^* = (4/3, 4/3), \mathbf{u}^* = (2/9, 2/9)$

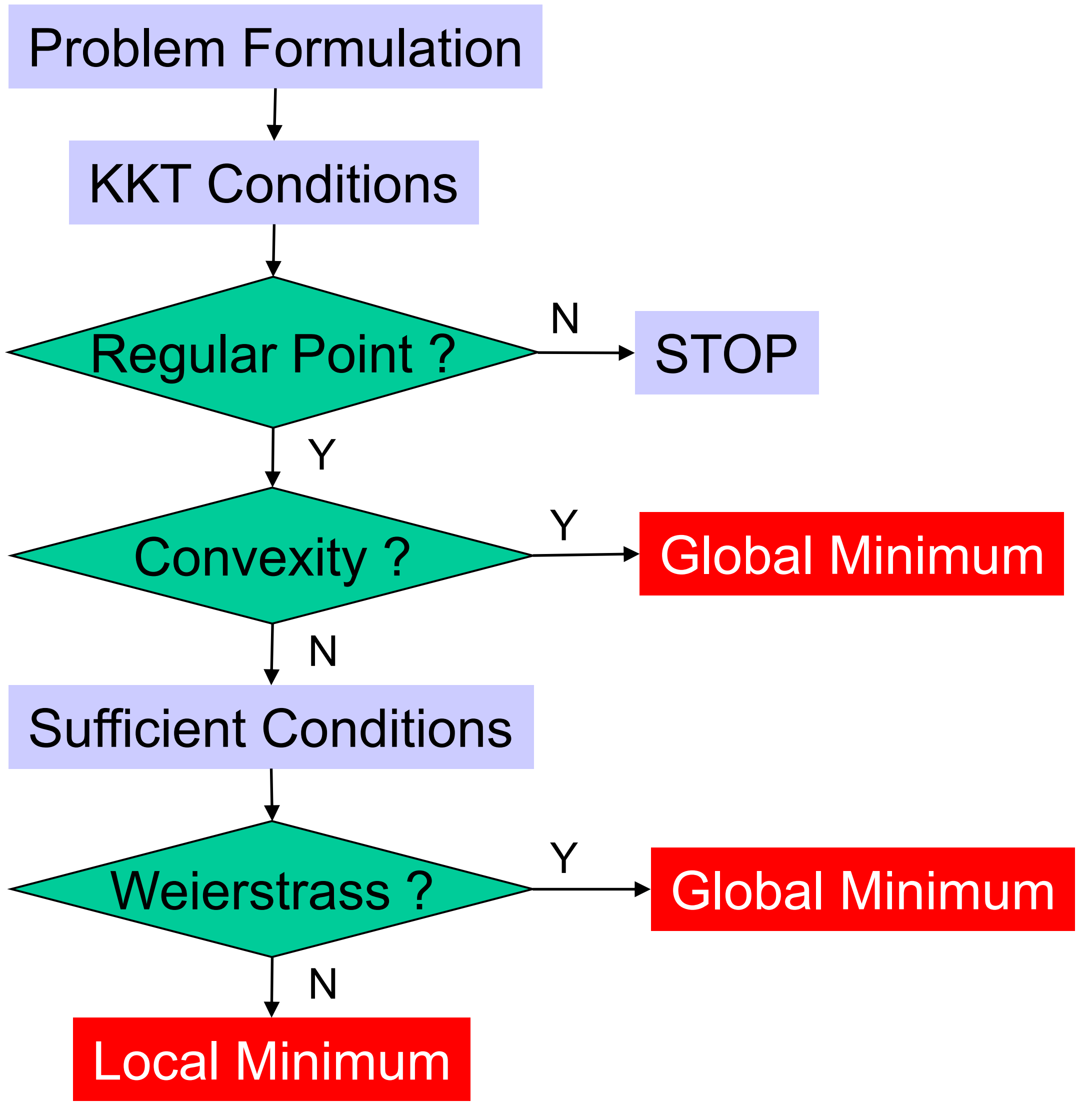
Summary: Optimality Conditions

	Unconstrained	Constrained
Necessary	$\nabla \boldsymbol{f} = 0$	$\nabla \boldsymbol{L} = 0$
Sufficient	$\nabla^2 \boldsymbol{f}$: positive definite	(1) convex problem: global (2) $\nabla^2 \mathbf{L}$: positive definite $\rightarrow$ strong (3) $\left.\begin{matrix} \nabla \mathbf{h}^T \mathbf{d} = 0 \\ \nabla \mathbf{g}^T \mathbf{d} = 0 (\text{active}) \end{matrix}\right\} \rightarrow \mathbf{d}^T \nabla^2 \mathbf{L} \mathbf{d} > 0$ $\rightarrow$ isolated local

Procedures (1)

- Problem formulation: DVs, objective, constraints
- Convexity check: global optimum ?
- K-T conditions: solutions
- Sufficiency check
- Sensitivity analysis: changes in the constraint limits

Procedures (2)



Design of a Wall Bracket

$$\underset{A_1, A_2}{\text{Minimize}} \quad f(A_1, A_2) = l_1 A_1 + l_2 A_2$$

subject to

$$g_1 = \frac{2.0E+06}{A_1} - 16000 \leq 0$$

$$g_2 = \frac{1.6E+06}{A_2} - 16000 \leq 0$$

$$g_3 = -A_1 \leq 0$$

$$g_4 = -A_2 \leq 0$$

Design of a Rectangular Beam

$$\underset{b,d}{\text{Minimize}}\ f = bd$$

subject to

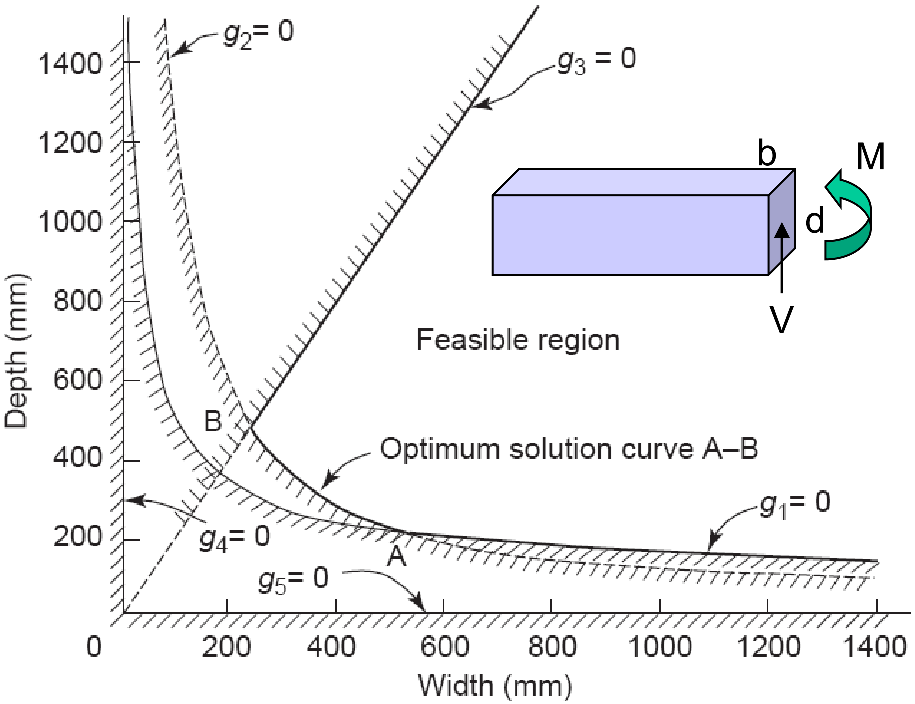
$$\left.\begin{aligned} \sigma &= \frac{6M}{bd^2} \le (\sigma_a)_{bending} \\ \tau &= \frac{3V}{2bd} \le (\tau_a)_{shear} \\ d &\le 2b \\ b, d &\ge 0 \end{aligned}\right\} \rightarrow \left\{\begin{aligned} g_1 &= \frac{2.40E+08}{bd^2} - 10 \le 0 \\ g_2 &= \frac{2.25E+05}{bd} - 2 \le 0 \\ g_3 &= d - 2b \le 0 \\ g_4 &= -b \le 0 \\ g_5 &= -d \le 0 \end{aligned}\right.$$

$$M = 40kN \cdot m$$

$$V = 150kN$$

$$(\sigma_a)_{bending} = 10MPa$$

$$(\tau_a)_{shear} = 2MPa$$



$b = 237mm,\ d = 474mm$ @ point B

$b = 527.3mm,\ d = 213.3mm$ @ point A