Contents

- Basic concepts
- Linearization of the constrained problem
- Sequential Linear Programming
- Sequential Quadratic Programming
  - Search direction, step size
- Constrained steepest-descent method
- Potential constraint strategy
- Other numerical optimization method
  - Method of feasible directions
  - Gradient projection method
  - Generalized reduced gradient method

Classification

$$
\text{Find } \boldsymbol{x} = (x_1, \ldots, x_n) \text{ which minimizes } f(\boldsymbol{x}) = f(x_1, \ldots, x_n)
$$

$$
\text{subject to} \begin{cases} h_i(\boldsymbol{x}) = h_i(x_1, \ldots, x_n) = 0; & i = 1, \ldots, p \\ g_j(\boldsymbol{x}) = g_j(x_1, \ldots, x_n) \le 0; & i = 1, \ldots, m \\ x_{kl} \le x_k \le x_{ku}; & k = 1, \ldots, n \end{cases}
$$

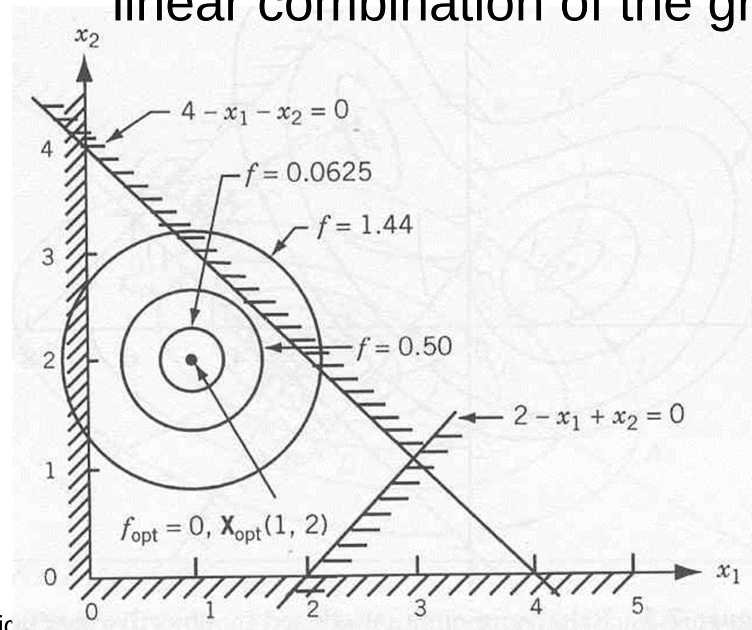
- Primal Method (direct method)
  - Search method that works on the original problem directly by searching through the feasible region for the optimal solution
- Transformation Method (indirect method): Ch.11.7
  - Convert a constrained optimization problem to a sequence of unconstrained optimization problems
  - Barrier and penalty function methods

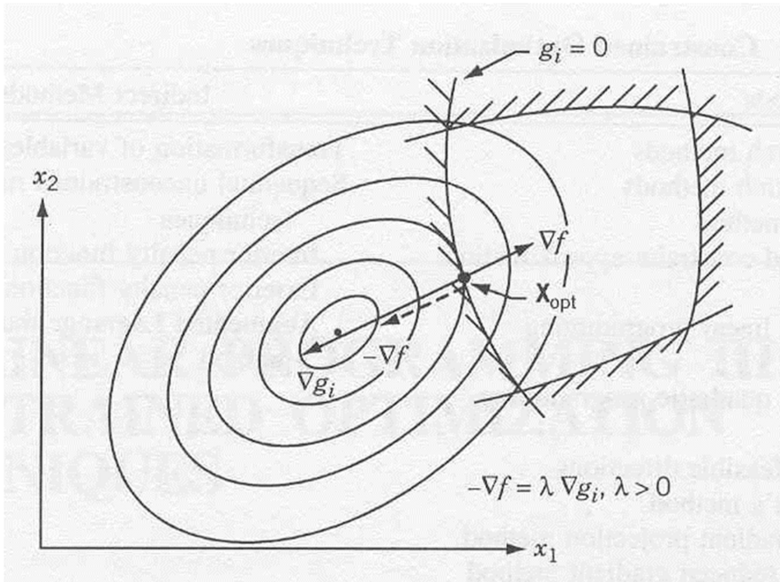
Constrained Optimization Methods

Direct (Primal) Methods	Indirect Methods
▪ Objective and constraint approximation methods – Sequential Linear Programming method – Sequential Quadratic Programming method ▪ Gradient Projection Method ▪ Methods of Feasible Directions ▪ Generalized Reduced Gradient Method	▪ Sequential unconstrained minimization technique – Interior penalty function method – Exterior penalty function method – Augmented Lagrange multiplier method

Characteristics of a Constrained Problem (1)

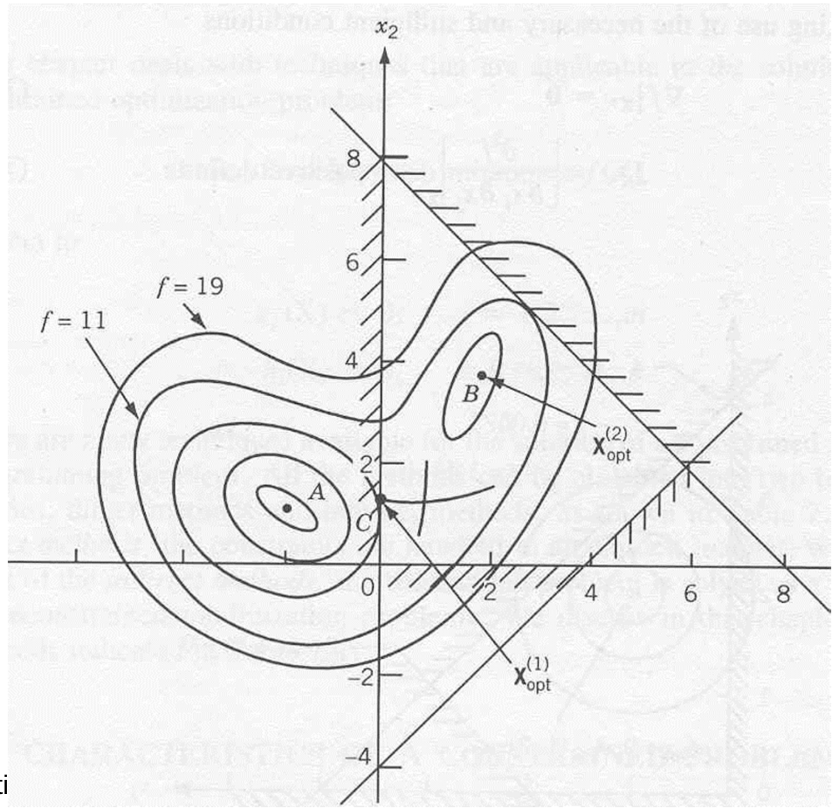
- The constraints may have no effect on the optimum point.
  - In most practical problems, it is difficult to identify whether the constraints have an influence on the minimum point.
- The optimum (unique) solution occurs on a constraint boundary.
  - The negative of the gradient must be expressible as a positive linear combination of the gradients of the active constraints.

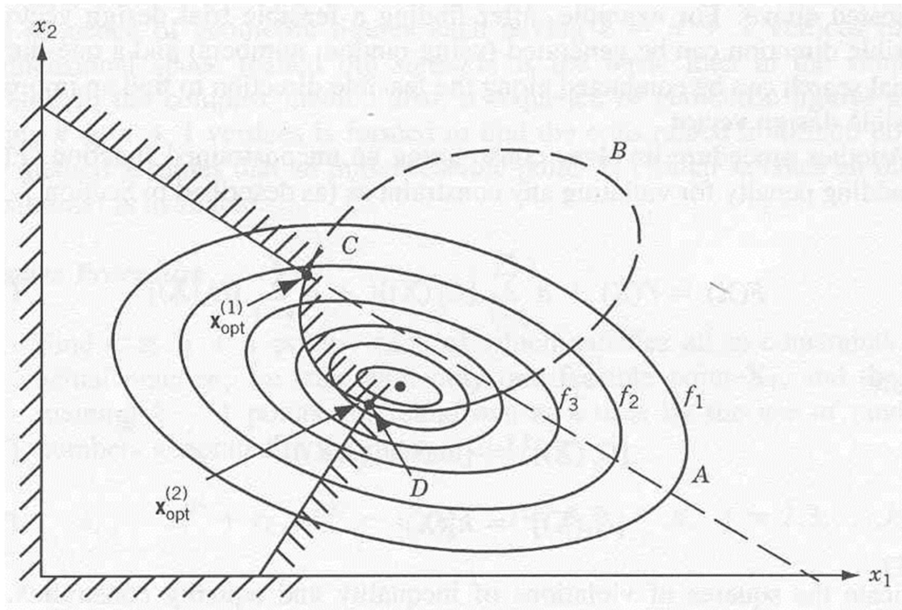




Characteristics of a Constrained Problem (2)

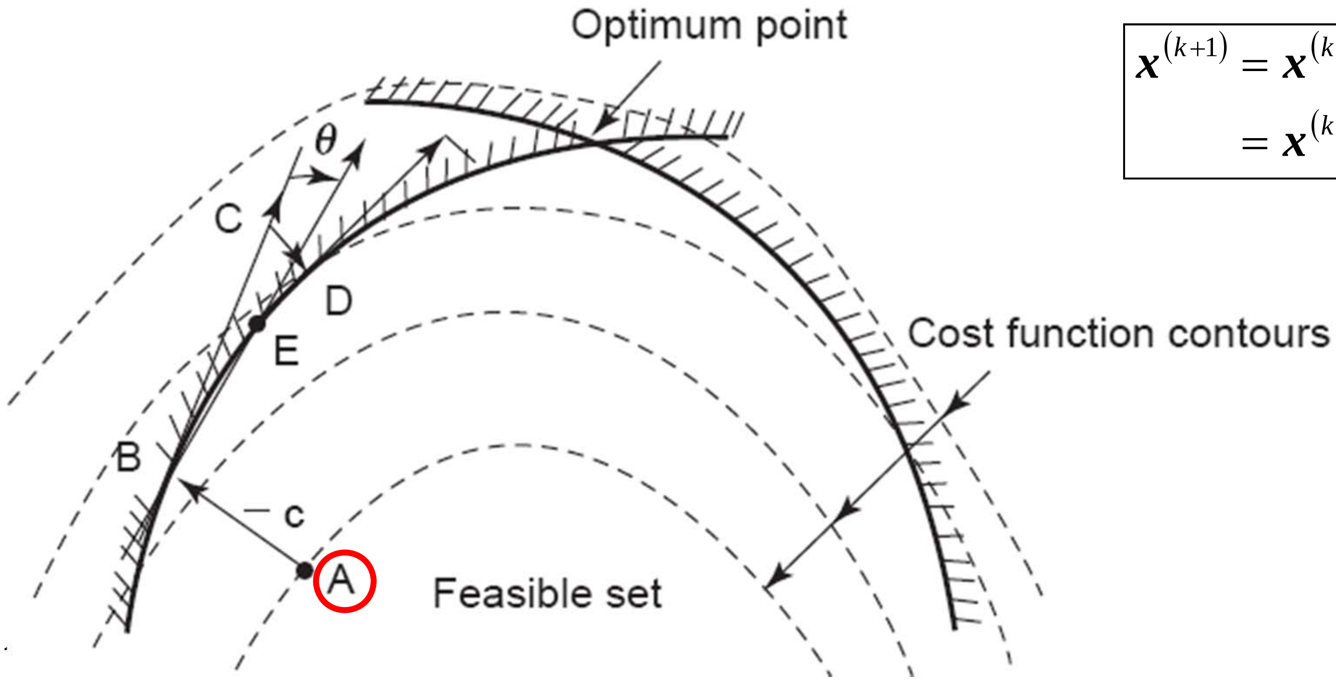
– If the objective function has two or more unconstrained local minima, the constrained problem may have multiple minima.
– Even if the objective function has a single unconstrained minimum, the constraints may introduce multiple local minima.





Basic Concepts (1)

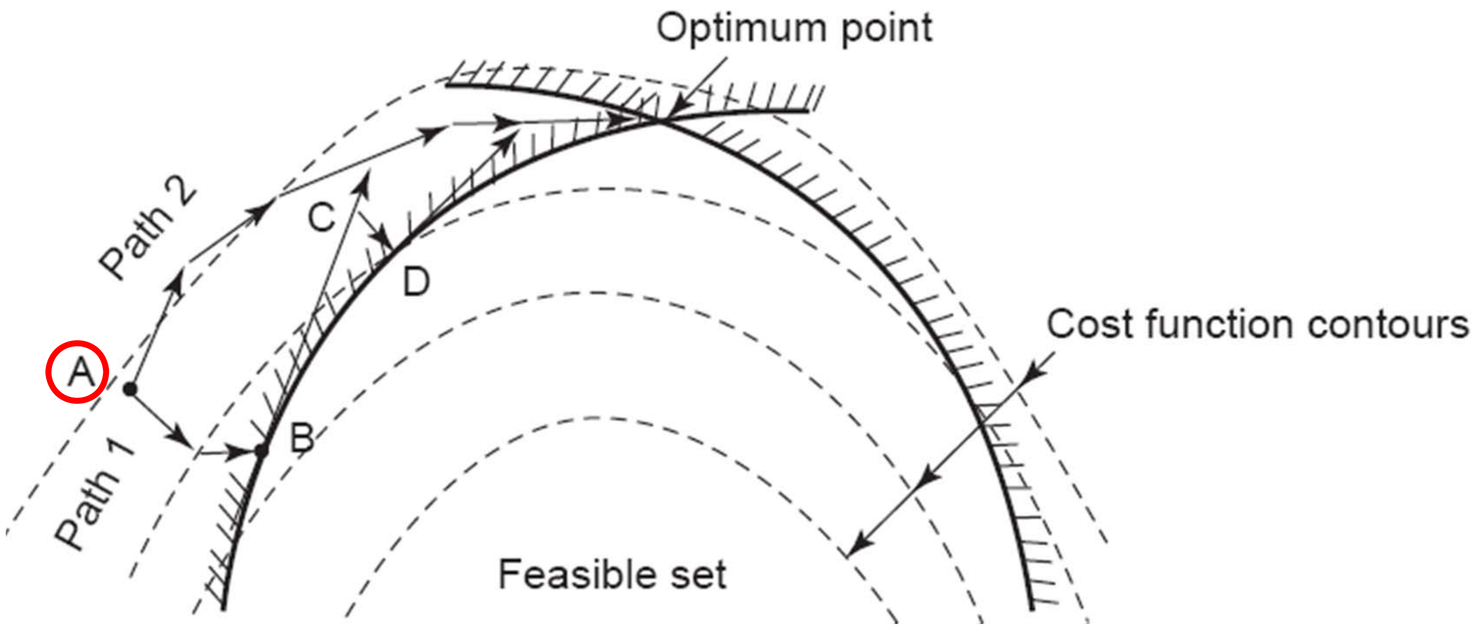
- From feasible starting point (inside the feasible region)
  - $\nabla f = 0$: Unconstrained stationary point$\rightarrow$check sufficient condition
  - $\nabla f \neq 0$: Moving along a descent direction
    - (Assume the optimum is on the boundary of the constraint set)
    - Travel along a tangent to the boundary $\rightarrow$correct to a feasible point
    - Deflect the tangential direction, toward the feasible region $\rightarrow$line search



$$\boxed{\begin{aligned} \boldsymbol{x}^{(k+1)} &= \boldsymbol{x}^{(k)} + \Delta\boldsymbol{x}^{(k)} \\ &= \boldsymbol{x}^{(k)} + \alpha_k d^{(k)} \end{aligned}}$$

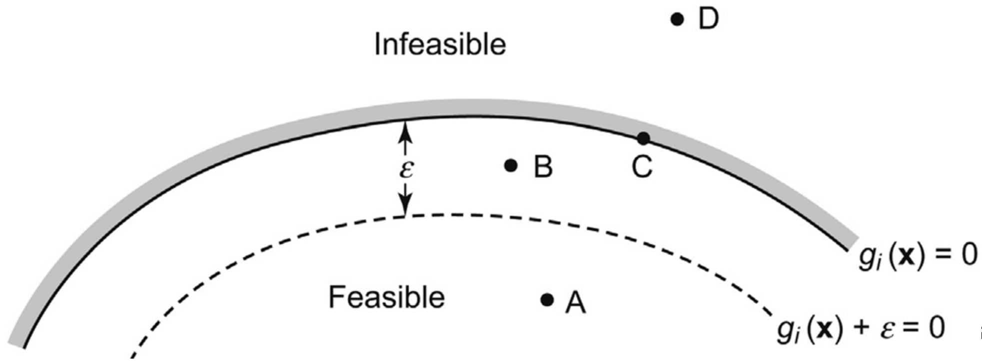
Basic Concepts (2)

- From infeasible starting point
  - Correct constraints to reach the constraint boundary $\rightarrow$same as previous steps
  - Iterate through the infeasible region to the optimum point



Basic Concepts (3)

- Numerical algorithm
  – Linearization of cost and constraint functions about the current design point
  – Definition of a search direction determination subproblem using the linearized functions
  – Solution of the subproblem that gives a search direction in the design space.
  – Calculation of a step size to minimize a descent function in the search direction
- Constraint status @ a design point
  – Active / Inactive / Violated / ε–Active



Basic Concepts (4)

- Constraint normalization
  - Same tolerance(ε) can be applied
  - Exception: divided by zero, undesirable situation(linear$\rightarrow$nonlinear)
- Descent (merit) function
  - A function used to monitor progress toward the minimum
  - Cost function + ?
- Convergent algorithm
  - Descent function, proper direction, closed and bound feasible set
  - Robust method
- Potential constraint strategy (Ch.13.1)
  - Numerical algorithms that use gradients of only a subset of the constraints

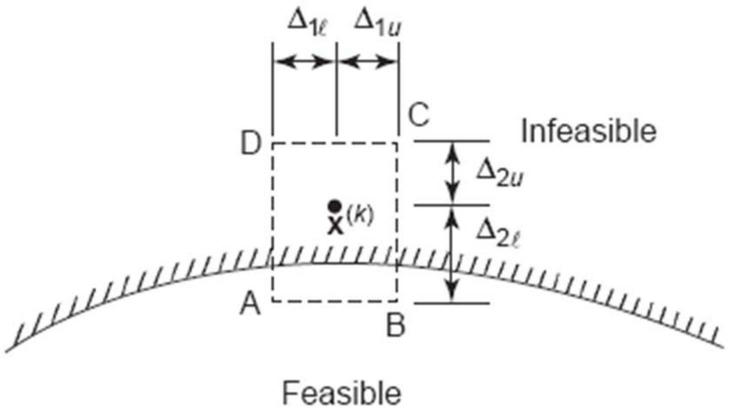
$$I_k = \left[\left\{ j \middle| j = 1 \text{ to } p \text{ for equalities} \right\} \text{and } \left\{ i \middle| g_i\left(\boldsymbol{x}^{(k)}\right) + \varepsilon \geq 0, i = 1 \text{ to } m \right\} \right]$$

Sequential Linear Programming

- Basic idea
  - Use linear approximation of the nonlinear functions and apply standard linear programming techniques
  - Repeated process successively as the optimization process
  - Major concern: How far from the point of interest are these approximations valid? move limits: depend on degree of nonlinearity)

$$-\Delta_{il}^{(k)} \le d_i \le \Delta_{iu}^{(k)}, \quad i = 1, \ldots, n$$

  - Some fraction of the current design variables (1~100%)
- Quite powerful and efficient for engineering design



Linearization

$$\min \quad f\left(\boldsymbol{x}^{(k)}+\Delta\boldsymbol{x}^{(k)}\right)\cong f\left(\boldsymbol{x}^{(k)}\right)+\nabla f^{T}\left(\boldsymbol{x}^{(k)}\right)\Delta\boldsymbol{x}^{(k)}$$

$$\text{subject to} \quad h_j\left(\boldsymbol{x}^{(k)}+\Delta\boldsymbol{x}^{(k)}\right)\cong h_j\left(\boldsymbol{x}^{(k)}\right)+\nabla h_j^{T}\left(\boldsymbol{x}^{(k)}\right)\Delta\boldsymbol{x}^{(k)}=0, \quad j=1,\ldots,p$$

$$g_j\left(\boldsymbol{x}^{(k)}+\Delta\boldsymbol{x}^{(k)}\right)\cong g_j\left(\boldsymbol{x}^{(k)}\right)+\nabla g_j^{T}\left(\boldsymbol{x}^{(k)}\right)\Delta\boldsymbol{x}^{(k)}\le 0, \quad j=1,\ldots,m$$

$$\boxed{\text{LP subproblem}}$$

$$\left.\begin{aligned}
&\min \quad \bar{f}=\sum_{i=1}^{n}\frac{\partial f\left(\boldsymbol{x}^{(k)}\right)}{\partial x_i}\Delta\boldsymbol{x}^{(k)}\\
&\text{s.t.} \quad \sum_{i=1}^{n}\frac{\partial h_j\left(\boldsymbol{x}^{(k)}\right)}{\partial x_i}\Delta\boldsymbol{x}^{(k)}=-h_j\left(\boldsymbol{x}^{(k)}\right)\\
&\qquad \sum_{i=1}^{n}\frac{\partial g_j\left(\boldsymbol{x}^{(k)}\right)}{\partial x_i}\Delta\boldsymbol{x}^{(k)}\le -g_j\left(\boldsymbol{x}^{(k)}\right)
\end{aligned}\right\}\rightarrow\left\{\begin{aligned}
&\min \quad \bar{f}=\sum_{i=1}^{n}c_i d_i\\
&\text{s.t.} \quad \sum_{i=1}^{n}n_{ij}d_i=e_j\\
&\qquad \sum_{i=1}^{n}a_{ij}d_i\le b_j
\end{aligned}\right\}\rightarrow\left\{\begin{aligned}
&\min \quad \bar{f}=\boldsymbol{c}^T\boldsymbol{d}\\
&\text{s.t.} \quad \underbrace{\boldsymbol{N}}_{(n\times p)}{}^{T}\boldsymbol{d}=\boldsymbol{e}\\
&\qquad \underbrace{\boldsymbol{A}}_{(n\times m)}{}^{T}\boldsymbol{d}\le\boldsymbol{b}
\end{aligned}\right.$$

SLP Algorithm

$$x^{(0)}, k = 0, \varepsilon_1, \varepsilon_2$$

$\begin{cases} \varepsilon_1\text{: constraint feasibility parameter} \\ \varepsilon_2\text{: stopping parameter} \end{cases}$

Evaluate $f\left(x^{(k)}\right), h_j\left(x^{(k)}\right), g_j\left(x^{(k)}\right)$

and $\dfrac{\partial f\left(x^{(k)}\right)}{\partial x_i}, \dfrac{\partial h_j\left(x^{(k)}\right)}{\partial x_i}, \dfrac{\partial g_j\left(x^{(k)}\right)}{\partial x_i}$

Define LP subproblem and Solve for $d^{(k)}$

(Select the proper move limit)

<Move limit>

- Bound the linearized subproblem
- Design change w/o line search

$$g_i \le \varepsilon_1, \left|h_i\right| \le \varepsilon_1, \left\|d^{(k)}\right\| \le \varepsilon_2 \ ?$$

Yes

$$x^* = x^{(k)}$$

No

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

Example 12.1+12.4 ← 4.31

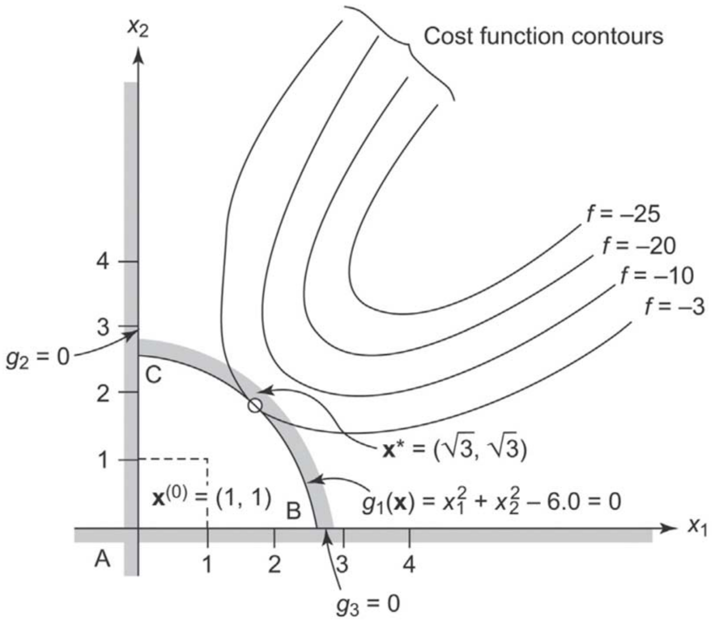
$$\text{Minimize } f(\boldsymbol{x}) = x_1^{\,2} + x_2^{\,2} - 3x_1x_2$$

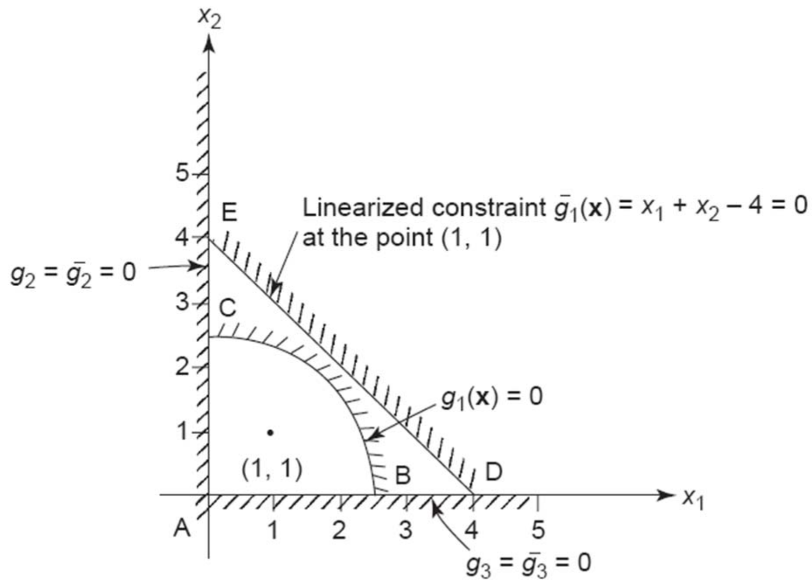
$$\text{subject to } g_1(\boldsymbol{x}) = \frac{1}{6}x_1^{\,2} + \frac{1}{6}x_2^{\,2} - 1 \le 0$$

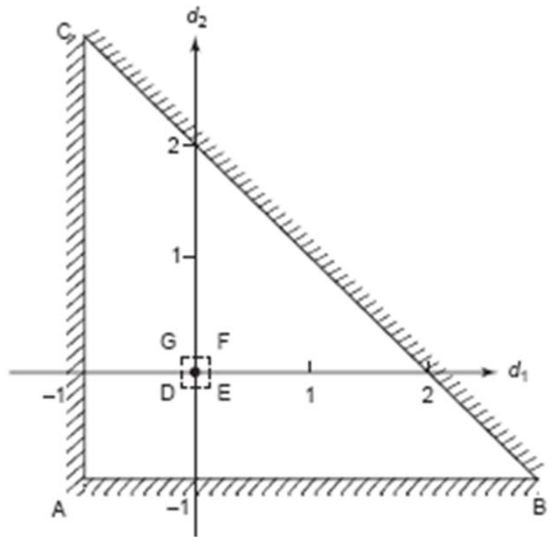
$$g_2(\boldsymbol{x}) = -x_1 \le 0$$

$$g_3(\boldsymbol{x}) = -x_2 \le 0$$

$$@\ \underbrace{\boldsymbol{x}^{(0)} = (1,1)}_{\text{feasible}} \text{ with 15\% move limit}$$







Example 12.3

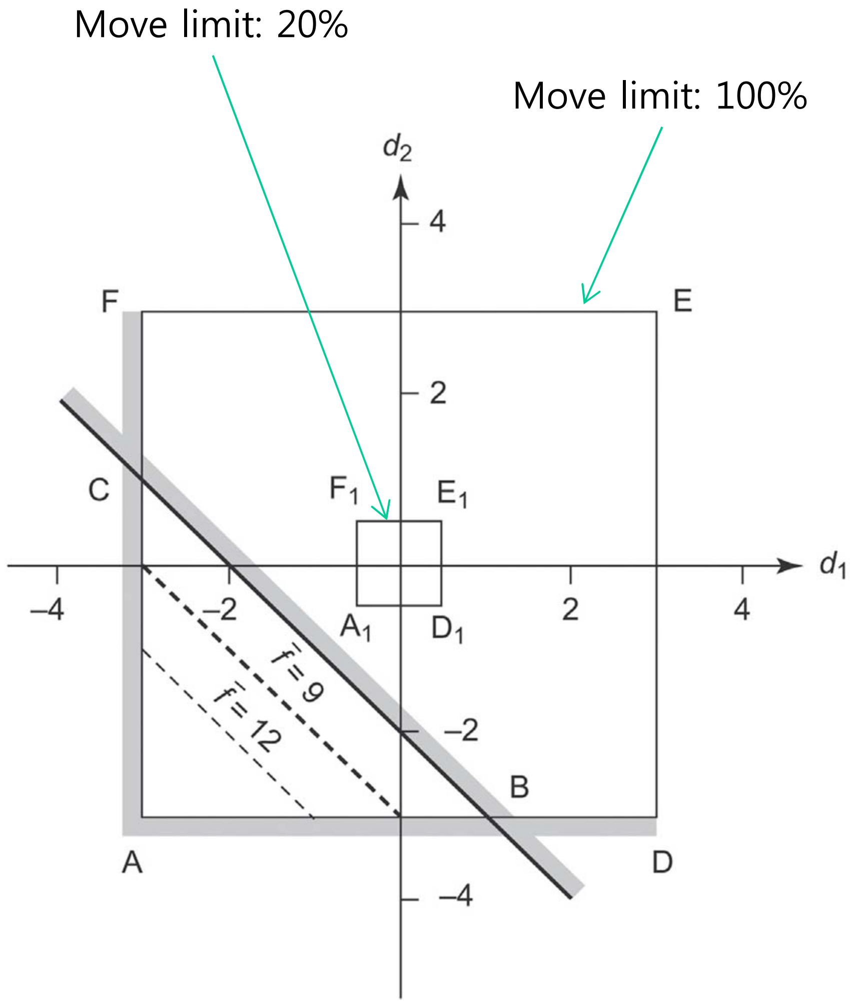
$$\text{Minimize } f(\boldsymbol{x}) = x_1^2 + x_2^2 - 3x_1x_2$$

$$\text{subject to } g_1(\boldsymbol{x}) = \frac{1}{6}x_1^2 + \frac{1}{6}x_2^2 - 1 \le 0$$

$$g_2(\boldsymbol{x}) = -x_1 \le 0$$

$$g_3(\boldsymbol{x}) = -x_2 \le 0$$

$@\ \underbrace{\boldsymbol{x}^{(0)} = (3,3)}_{\text{infeasible}}$ proper move limit?



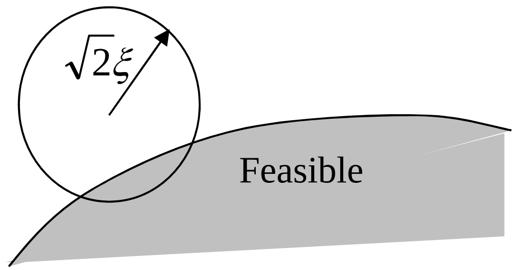
Observations

- Cannot be used as a black box approach for engineering design problems
  - Trial and error process for the selection of move limits
- May not convergent since no descent function is defined and line search is not performed along the search direction
- Rate of convergence and performance depend (to a large extent) on the selection of the move limits
  - Lack of robustness and uncertainty
- Simple conceptually as well as numerically
  - Improved designs in practice rather than the precise optimum

Quadratic Programming Subproblem

- Quadratic cost function + linear constraints
- SLP: linear move limits $\rightarrow$ quadratic step size constraint

$$-\Delta_{il}^{(k)} \le d_i \le \Delta_{iu}^{(k)} \rightarrow \|d\| \le \xi \rightarrow 0.5\sum_{i=1}^{n}\left(d_i\right)^2 \le \xi^2$$



$$\left.\begin{aligned} &\min \quad \bar{f} = \sum_{i=1}^{n} c_i d_i \\ &\text{s.t.} \quad \sum_{i=1}^{n} n_{ij} d_i = e_j, \quad j = 1, \ldots, p \\ &\qquad \sum_{i=1}^{n} a_{ij} d_i \le b_j, \quad j = 1, \ldots, m \\ &\qquad 0.5\sum_{i=1}^{n}\left(d_i\right)^2 \le \xi^2 \end{aligned}\right\} \rightarrow \left\{\begin{aligned} &\min \quad \bar{f} = \boldsymbol{c}^T \boldsymbol{d} \\ &\text{s.t.} \quad \underbrace{\boldsymbol{N}}_{(n\times p)}{}^T \boldsymbol{d} = \boldsymbol{e} \\ &\qquad \underbrace{\boldsymbol{A}}_{(n\times m)}{}^T \boldsymbol{d} \le \boldsymbol{b} \\ &\qquad 0.5\boldsymbol{d}^T \boldsymbol{d} \le \xi^2 \end{aligned}\right\} \leftrightarrow \left\{\begin{aligned} &\min \quad \bar{f} = \boldsymbol{c}^T \boldsymbol{d} + 0.5\boldsymbol{d}^T \boldsymbol{d} \\ &\text{s.t.} \quad \boldsymbol{N}^T \boldsymbol{d} = \boldsymbol{e} \\ &\qquad \boldsymbol{A}^T \boldsymbol{d} \le \boldsymbol{b} \end{aligned}\right.$$

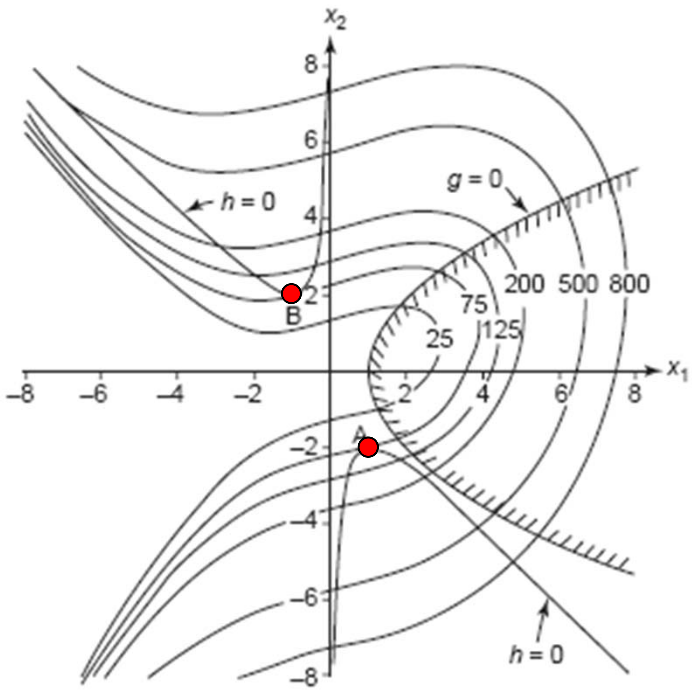
Strictly convex →
Minimum is global and unique

$$\left(d_1 + c_1\right)^2 + \left(d_2 + c_2\right)^2 = r^2 \rightarrow d_1^{\,2} + c_1^{\,2} + 2c_1 d_1 + d_2^{\,2} + c_2^{\,2} + 2c_2 d_2 = r^2$$

$$\frac{1}{2}\left(r^2 - c_1^{\,2} - c_2^{\,2}\right) = c_1 d_1 + c_2 d_2 + \frac{1}{2}\left(d_1^{\,2} + d_2^{\,2}\right)\text{: hypersphere with its center at } -\mathbf{c}$$

Example 12.5 (1)

$$\begin{aligned} &\textit{Minimize} \quad f(x) = 2x_1^{\,3} + 15x_2^{\,2} - 8x_1x_2 - 4x_1 \\ &\textit{subject to} \quad h(x) = x_1^{\,2} + x_1x_2 + 1.0 = 0 \\ &\qquad\qquad\quad g(x) = x_1 - \frac{1}{4}x_2^{\,2} - 1.0 \le 0 \end{aligned}$$



$$\nabla f = \left(6x_1^{\,2} - 8x_2 - 4,\ 30x_2 - 8x_1\right)$$
$$\nabla h = \left(2x_1 + x_2,\ x_1\right)$$
$$\nabla g = \left(1, -x_2/2\right)$$
$$\rightarrow \text{linearize}\ @(1,1)$$
$$f(1,1) = 5$$
$$h(1,1) = 3 \ne 0 \ \ \text{(violation)}$$
$$g(1,1) = -0.25 < 0 \ \ \text{(inactive)}$$
$$c = \nabla f = (-6, 22)$$
$$\nabla h = (3,1)$$
$$\nabla g = (1, -0.5)$$

point A: $\boldsymbol{x}^* = (1,-2),\ f(\boldsymbol{x}^*) = 74$

point B: $\boldsymbol{x}^* = (-1,2),\ f(\boldsymbol{x}^*) = 78$

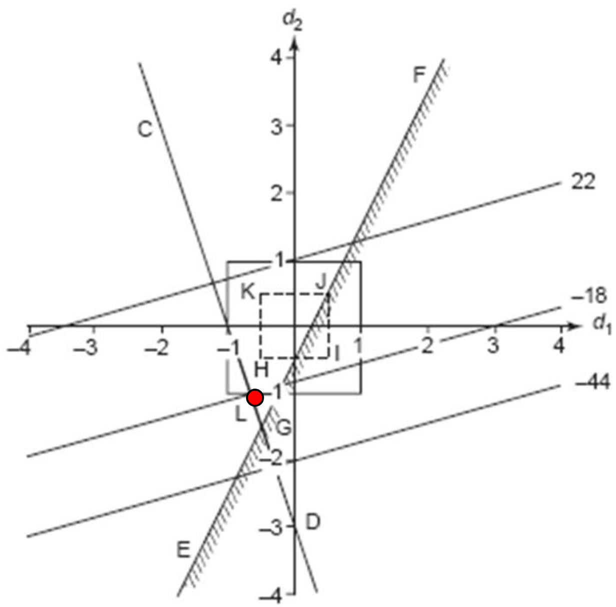
Example 12.5 (2)

< LP subproblem >

$$\begin{aligned} &\textit{Minimize} \quad \bar{f} = -6d_1 + 22d_2 \\ &\textit{subject to} \quad 3d_1 + d_2 = -3 \\ &\qquad\qquad\quad d_1 - 0.5d_2 \le 0.25 \\ &\qquad\qquad\quad -0.5 \le d_1, d_2 \le 0.5 \end{aligned}$$

[move limit : 50% (infeasible) $\rightarrow$ 100%]

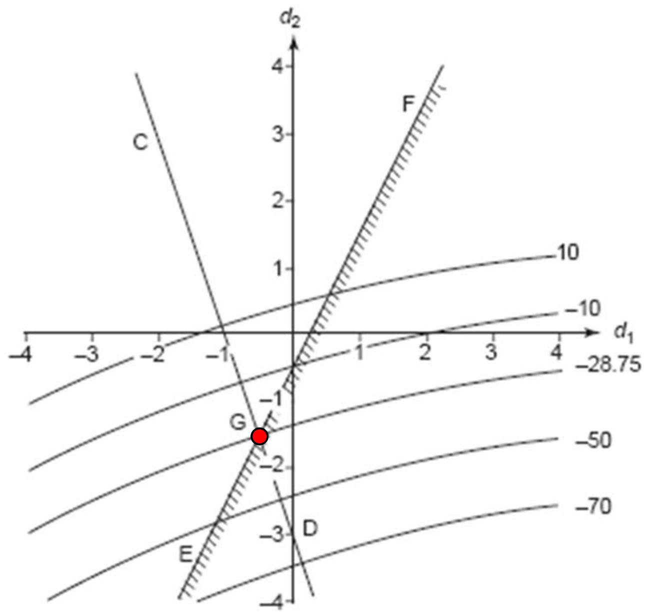
$$\rightarrow\ d_1 = -2/3,\ d_2 = -1.0,\ \bar{f} = -18$$



< QP subproblem >

$$\begin{aligned} &\textit{Minimize} \quad \bar{f} = -6d_1 + 22d_2 + 0.5\left(d_1^{\,2} + d_2^{\,2}\right) \\ &\textit{subject to} \quad 3d_1 + d_2 = -3 \\ &\qquad\qquad\quad d_1 - 0.5d_2 \le 0.25 \end{aligned}$$

$$\rightarrow\ d_1 = -0.5,\ d_2 = -1.5,\ \bar{f} = -28.75$$



Solution of QP problems

$$\begin{aligned}
&\textit{Minimize} \quad q(\mathbf{x}) = \mathbf{c}^T\mathbf{x} + 0.5\mathbf{x}^T\mathbf{H}\mathbf{x} \\
&\textit{subject to} \quad \mathbf{A}^T\mathbf{x} \le \mathbf{b} \rightarrow \mathbf{A}^T\mathbf{x} + \mathbf{s} = \mathbf{b} \\
&\qquad\qquad\quad \mathbf{N}^T\mathbf{x} = \mathbf{e} \\
&\qquad\qquad\quad \mathbf{x} \ge \mathbf{0} \rightarrow -\mathbf{x} \le \mathbf{0}
\end{aligned}$$

$$L = \mathbf{c}^T\mathbf{x} + 0.5\mathbf{x}^T\mathbf{H}\mathbf{x} + \mathbf{u}^T\left(\mathbf{A}^T\mathbf{x} + \mathbf{s} - \mathbf{b}\right) + \mathbf{v}^T\left(\mathbf{N}^T\mathbf{x} - \mathbf{e}\right) - \zeta^T\mathbf{x}$$

$$\partial L / \partial \mathbf{x} = \mathbf{c} + \mathbf{H}\mathbf{x} + \mathbf{A}\mathbf{u} + \mathbf{N}\mathbf{v} - \zeta = 0$$

$$\mathbf{A}^T\mathbf{x} + \mathbf{s} - \mathbf{b} = \mathbf{0}$$

$$\mathbf{N}^T\mathbf{x} - \mathbf{e} = \mathbf{0}$$

$$u_i s_i = 0; \quad i = 1, \ldots, m$$

$$\zeta_i x_i = 0; \quad i = 1, \ldots, n$$

$$s_i, u_i \ge 0; \quad i = 1, \ldots, m$$

$$\zeta_i \ge 0; \quad i = 1, \ldots, n$$

Solution for QP problem

$$\left.\begin{aligned} \frac{\partial L}{\partial \boldsymbol{x}} = \boldsymbol{c} + \boldsymbol{H}\boldsymbol{x} + \boldsymbol{A}\boldsymbol{u} + \boldsymbol{N}\boldsymbol{v} - \zeta = 0 \\ \boldsymbol{A}^T\boldsymbol{x} + \boldsymbol{s} - \boldsymbol{b} = \boldsymbol{0} \\ \boldsymbol{N}^T\boldsymbol{x} - \boldsymbol{e} = \boldsymbol{0} \end{aligned}\right\} \xrightarrow{\boldsymbol{v}=\boldsymbol{y}-\boldsymbol{z}} \boldsymbol{B}_{[(m+n+p)\times 2(m+n+p)]}\boldsymbol{X} = \boldsymbol{D}$$

$$(m+n+p)\underbrace{\begin{bmatrix} \boldsymbol{H}_{(n\times n)} & \boldsymbol{A}_{(n\times m)} & -\boldsymbol{I}_{(n\times n)} & \boldsymbol{0}_{(n\times m)} & \boldsymbol{N}_{(n\times p)} & -\boldsymbol{N}_{(n\times p)} \\ \boldsymbol{A}^T{}_{(m\times n)} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{N}^T{}_{(p\times n)} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}}_{2(m+n+p)} \begin{bmatrix} \boldsymbol{x}_{(n\times 1)} \\ \boldsymbol{u}_{(m\times 1)} \\ \zeta_{(n\times 1)} \\ \boldsymbol{s}_{(m\times 1)} \\ \boldsymbol{y}_{(p\times 1)} \\ \boldsymbol{z}_{(p\times 1)} \end{bmatrix} = \begin{bmatrix} -\boldsymbol{c} \\ \boldsymbol{b} \\ \boldsymbol{e} \end{bmatrix}$$

$$\left.\begin{aligned} & u_i s_i = 0; \quad i = 1,\dots,m \\ & \zeta_i x_i = 0; \quad i = 1,\dots,n \\ & s_i, u_i \geq 0; \quad i = 1,\dots,m \\ & \zeta_i \geq 0; \quad i = 1,\dots,n \end{aligned}\right\} \rightarrow \begin{cases} \underline{X_i X_{n+m+i} = 0;} \quad i = 1,\dots,n+m \\ X_i \geq 0; \quad i = 1,\dots,2(m+n+p) \end{cases}$$

Simplex method for solving QP problem

- Wolfe(1959)$\rightarrow$Hadley(1964)
  - Solve as a linear program using the Phase I simplex method
  - Fail to converge when H is positive semidefinite if c $\neq$ 0
- Lemke(1965)
  - Complementary pivot method
- Phase I:

$$\boldsymbol{BX} + \underbrace{\boldsymbol{Y}}_{\substack{\text{artificial}\\ \text{variable}}} = \underbrace{\boldsymbol{D}}_{\text{nonnegative}}$$

$$w = \sum_{i=1}^{m+n+p} Y_i = \sum_{i=1}^{m+n+p} D_i - \sum_{j=1}^{2(m+n+p)} \sum_{i=1}^{m+n+p} B_{ij} X_j = w_0 + \sum_{j=1}^{2(m+n+p)} C_j X_j$$

$$\left(\text{where} \quad w_0 = \sum_{i=1}^{m+n+p} D_i ,\; C_j = -\sum_{i=1}^{m+n+p} B_{ij} \right)$$