Design Theories of Ship and Offshore Plant

Part II. Optimum Design

Ch. 3 Penalty Function Method

Fall 2014

Myung-Il Roh

Department of Naval Architecture and Ocean Engineering
Seoul National University

Contents

Ch. 3 Penalty Function Method

3.1 Interior Penalty Function Method

Lagrange Multiplier

Constrained Optimal Design Problem

Minimize $f(\mathbf{x})$

Subject to $\mathbf{h}(\mathbf{x}) = \mathbf{0}$ Equality constraint

$\mathbf{g}(\mathbf{x}) \le \mathbf{0}$ Inequality constraint

Transforming this problem to unconstrained optimal design problem by using the Lagrangian function

$$L(\mathbf{x}, \mathbf{v}, \mathbf{u}, \mathbf{s}) = f(\mathbf{x}) + \mathbf{v}^T \mathbf{h}(\mathbf{x}) + \mathbf{u}^T (\mathbf{g}(\mathbf{x}) + \mathbf{s}^2)$$

By using the necessary condition for the candidate local optimal solution(∇L=0), are calculated.

1) If the constraints are satisfied at the current design point,

In case of the equality constraints: $\mathbf{h}(\mathbf{x}) = \mathbf{0}$

In case of the inequality constraints: $\mathbf{u} = \mathbf{0}$ (The constraints are inactive, i.e, the design point is in feasible region.)

$\mathbf{s} = \mathbf{0} \Rightarrow \mathbf{g}(\mathbf{x}) = \mathbf{0}$ (The constraints are active, i.e, the design point is on the constraints)

Therefore, $L(\mathbf{x}, \mathbf{v}, \mathbf{u}, \mathbf{s}) = f(\mathbf{x}) + \mathbf{v}^T \mathbf{h}(\mathbf{x}) + \mathbf{u}^T (\mathbf{g}(\mathbf{x}) + \mathbf{s}^2) \Rightarrow f(\mathbf{x})$ ➡ If all the constrains are satisfied, the Lagrange function is same with the original objective function.

2) If the constraints are violated at the current design point,

In case of the equality constraints: $\mathbf{v}^T \mathbf{h}(\mathbf{x}) \ne \mathbf{0}$

In case of the inequality constraints: $\mathbf{u}^T (\mathbf{g}(\mathbf{x}) + \mathbf{s}^2) > \mathbf{0}$

Therefore, $L(\mathbf{x}, \mathbf{v}, \mathbf{u}, \mathbf{s}) = f(\mathbf{x}) + \mathbf{v}^T \mathbf{h}(\mathbf{x}) + \mathbf{u}^T (\mathbf{g}(\mathbf{x}) + \mathbf{s}^2)$

➡ This term means augmenting a penalty to the original objective function when the constraints are violated.

5

SUMT: Sequential Unconstrained Minimization Technique (Interior Penalty Function Method) (1/2)

Constrained Optimal Design Problem

Minimize $f(\mathbf{x})$

Subject to $\mathbf{h}(\mathbf{x}) = \mathbf{0}$ Equality constraint

$\mathbf{g}(\mathbf{x}) \le \mathbf{0}$ Inequality constraint

Fiacco and McCormick suggested a method which transforms the constrained optimization problem into the unconstrained optimization problem by using the modified objective function in 1968.
The modified objective function is a function augmenting a penalty to the original objective function.

SUMT: Sequential Unconstrained Minimization Technique

$$\Phi(\mathbf{x}, r_k) = f(\mathbf{x}) - r_k \sum_{j=1}^{m} \frac{1}{g_j(\mathbf{x})}$$

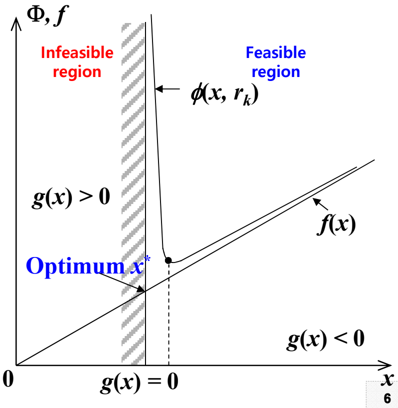
where r_k is given and positive value and getting smaller each iteration.

If the design point approaches to the boundary of the inequality constraints in the feasible region,

$g_j(\mathbf{x}) \le 0$, the absolute value of this is decreased.

$-r_k \frac{1}{g_j(\mathbf{x})} > 0$, the absolute value of this is increased.

Since the modified objective function is increased as the design point approaches to the boundary of the inequality constraint, this method prevents the design point violating the constraints.



6

SUMT: Sequential Unconstrained Minimization Technique (Interior Penalty Function Method) (2/2)

☑ If the design point approaches to the boundary of the constraints in the feasible region, the objective function is augmented by a penalty.

☑ The starting design point has to be in the feasible region.

$$\Phi(\mathbf{x}, r_k) = f(\mathbf{x}) - r_k \sum_{j=1}^{m} \frac{1}{g_j(\mathbf{x})}$$ (r_k is decreased, when k is increased.)

[Example] Function of a single variable

$f(x) = \alpha x, \quad g(x) = \beta - x \le 0 \quad ,(x \ge \beta \Rightarrow \beta - x \le 0)$

Transform the constrained optimization problem into the unconstrained optimization problem.

$$\Phi(x, r_k) = f(x) - r_k \frac{1}{g(x)} = \alpha x - r_k \frac{1}{\beta - x}$$

- k is the number of iteration.
- In each iteration, the optimal design point can be obtained by using the Gradient method, Hooke & Jeeves or Nelder & Mead method.

$k = 1$, Starting design point: x^*_0

$\Phi(x, r_1) = \alpha x - r_1 \dfrac{1}{\beta - x}$ ⇨ Optimal design point: x^*_1

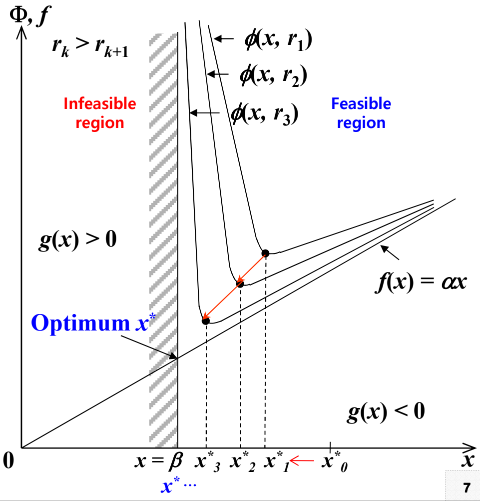
$k = 2$, Starting design point : x^*_1

$\Phi(x, r_2) = \alpha x - r_2 \dfrac{1}{\beta - x}$ ⇨ Optimal design point : x^*_2

$k = 3$, Starting design point : x^*_2

$\Phi(x, r_3) = \alpha x - r_3 \dfrac{1}{\beta - x}$ ⇨ Optimal design point : x^*_3

By iterating the above process, we find the optimal design point(x^*).



3.2 Exterior Penalty Function Method

sydlab SEOUL NAT'L UNIV.

Exterior Penalty Function Method (1/2)

☑ There will be a penalty for only violating the constraints.

$$\Phi(\mathbf{x}, r_k) = f(\mathbf{x}) + r_k \sum_{j=1}^{m} \left[\max\left\{ g_j(\mathbf{x}), 0 \right\} \right]^2 \quad (r_k \text{ is increased, when } k \text{ is increased.})$$

[Example] Function of a single variable

$$f(x) = \alpha x, \quad g(x) = \beta - x \le 0, (x \ge \beta \Rightarrow \beta - x \le 0)$$

Transform the constrained optimization problem into the unconstrained optimization problem.

$$\Phi(x, r_k) = f(x) + r_k \max\left\{ g(x), 0 \right\}^2 = \alpha x + r_k \max\left\{ g(x), 0 \right\}^2$$

- k is the number of iteration.
- In each iteration, the optimal design point can be obtained by using the Gradient method, Hooke & Jeeves or Nelder & Mead method.

$k = 1$, Starting design point: x^*_0

$\Phi(x, r_1) = \alpha x + r_1 \left[\max\left\{ g(x), 0 \right\} \right]^2$ ⇨ Optimal design point: x^*_1

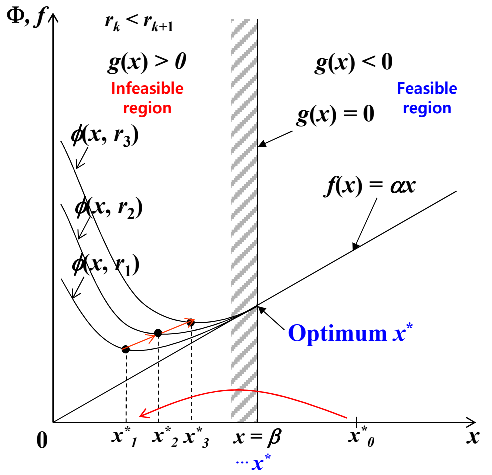
$k = 2$, Starting design point: x^*_1

$\Phi(x, r_2) = \alpha x + r_2 \left[\max\left\{ g(x), 0 \right\} \right]^2$ ⇨ Optimal design point: x^*_2

$k = 3$, Starting design point: x^*_2

$\Phi(x, r_3) = \alpha x + r_3 \left[\max\left\{ g(x), 0 \right\} \right]^2$ ⇨ Optimal design point: x^*_3

By iterating the above process, we find the optimal design point(x^*).



Exterior Penalty Function Method (2/2)

☑ There will be a penalty for only violating the constraints.

$$\Phi(\mathbf{x}, r_k) = f(\mathbf{x}) + r_k \sum_{j=1}^{m} \max\left\{ g_j(\mathbf{x}), 0 \right\} \quad (r_k \text{ is increased, when } k \text{ is increased.})$$

[Example] Function of a single variable

$$f(x) = \alpha x, \quad g(x) = \beta - x \le 0 \ , (x \ge \beta \Rightarrow \beta - x \le 0)$$

Transform the constrained optimization problem into the unconstrained optimization problem.

$$\Phi(x, r_k) = f(x) + r_k \max\left\{ g(x), 0 \right\} = \alpha x + r_k \max\left\{ g(x), 0 \right\}$$

- k is the number of iteration.
- In each iteration, the optimal design point can be obtained by using the Gradient method, Hooke & Jeeves or Nelder & Mead method.

$k = 1$, Starting design point: x^*_0

$\Phi(x, r_1) = \alpha x + r_1 \max\left\{ g(x), 0 \right\}$ ⇨ Optimal design point: we can not find it.

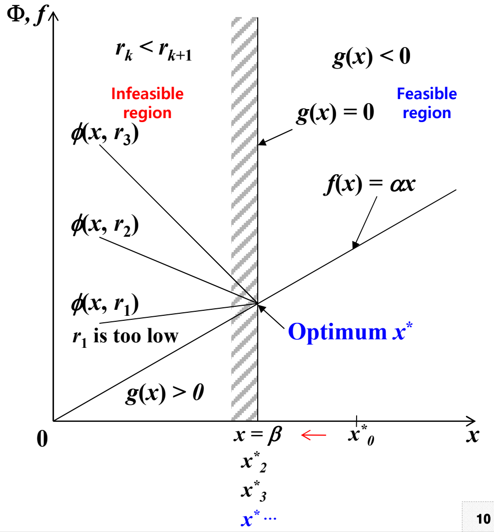
$k = 2$, Starting design point: x^*_1

$\Phi(x, r_2) = \alpha x + r_2 \max\left\{ g(x), 0 \right\}$ ⇨ Optimal design point: x^*_2

$k = 3$, Starting design point: x^*_2

$\Phi(x, r_3) = \alpha x + r_3 \max\left\{ g(x), 0 \right\}$ ⇨ Optimal design point: x^*_3

If r_k is determined properly, the optimal design point(x^*) is not changed.



Relationship between External Penalty Function and Feasible Region (1/2)

☑ Since there will be a penalty for only violating the constraints, **if the minimum design point is in the feasible region**, the result of the optimization method by using the exterior penalty function is the same with that only using the objective function.

[Example] Function of a single variable

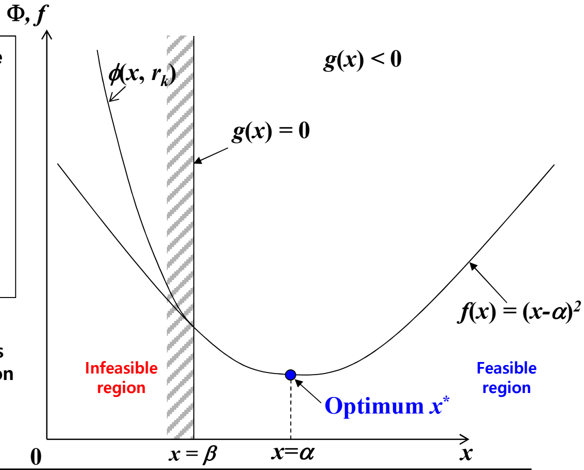
$f(x)=(x-\alpha)^2,\ g(x)=\beta-x\le 0$

$\Phi(\mathbf{x},r_k)=f(\mathbf{x})+r_k\underbrace{\sum_{j=1}^{m}\left[\cancelto{0}{\max\{g_j(\mathbf{x}),0\}}\right]^2}_{\text{Penalty term}}$

$\Phi(\mathbf{x},r_k)=f(\mathbf{x})$

$where,\ \ g(\mathbf{x})\le 0,\ \ \max\{g_j(\mathbf{x}),0\}=0$

If the minimum design point(x^*) is in the feasible region, the penalty term is equal to zero. So, the objective function augmented by the penalty is the same with the original objective function.



Relationship between External Penalty Function and Feasible Region (2/2)

☑ Since there will be a penalty for only violating the constraints, **if the minimum design point is not in the feasible region (infeasible region)**, the result of the optimization method by using the exterior penalty function is different with that only using the objective function.

[Example] Function of a single variable

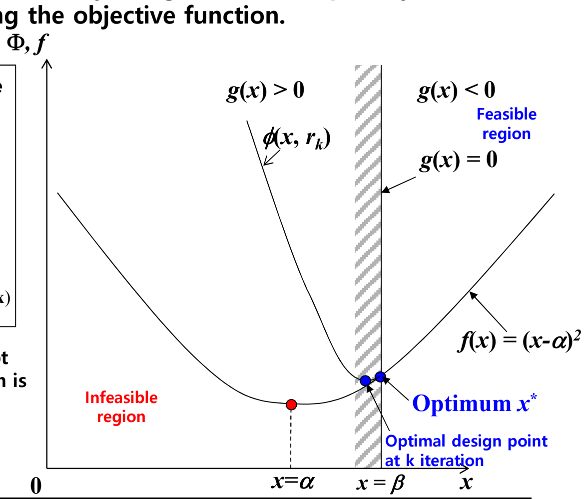
$f(x)=(x-\alpha)^2,\ g(x)=\beta-x\le 0$

$\Phi(\mathbf{x},r_k)=f(\mathbf{x})+r_k\underbrace{\sum_{j=1}^{m}\left[\max\{g_j(\mathbf{x}),0\}\right]^2}_{\text{Penalty term}}$

$\Phi(\mathbf{x},r_k)=f(\mathbf{x})+r_k\sum_{j=1}^{m}g_j(\mathbf{x})^2$

$where,\ \ g(\mathbf{x})>0,\ \ \max\{g_j(\mathbf{x}),0\}=g_j(\mathbf{x})$

If the minimum design point(x^*) is not in the feasible region, the penalty term is larger than zero. So, the objective function augmented by the penalty is different with the original objective function.



3.3 Augmented Lagrange Multiplier Method

Augmented Lagrange Multiplier Method

- ☑ This method combines the Lagrange multiplier and the penalty function methods.
- ☑ There is no need for the penalty parameter "r" to go to infinity for the exterior penalty function method or zero for the interior penalty function method.
- ☑ Starting point does not have to be in feasible region for the interior penalty function method.
- ☑ It has been proven that they possess a faster rate of convergence than interior and exterior penalty function method.

Augmented Lagrange Multiplier Method in Equality Constrained Problem (1/4)

$$\begin{array}{ll} \textit{Minimize} & f(\mathbf{x}) \\ \textit{Subject to} & h_j(\mathbf{x}) = \mathbf{0}, \quad j = 1, 2, \ldots, m \end{array}$$

Lagrangian function of this problem is as follows.

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \sum_{j=1}^{m} \lambda_j h_j(\mathbf{x})$$

Augmented Lagrangian function of this problem is follows.

$$\Phi(\mathbf{x}, \boldsymbol{\lambda}, r_k) = f(\mathbf{x}) + \sum_{j=1}^{m} \lambda_j h_j(\mathbf{x}) + \boxed{r_k \sum_{j=1}^{m} h_j^2(\mathbf{x})}$$

Augmented term to Lagrangian function

r_k : arbitrary constant

Augmented Lagrange Multiplier Method in Equality Constrained Problem (2/4)

$$\begin{array}{ll} \textit{Minimize} & f(\mathbf{x}) \\ \textit{Subject to} & h_j(\mathbf{x}) = \mathbf{0}, \quad j = 1, 2, \ldots, m \end{array}$$

Lagrangian function

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \sum_{j=1}^{m} \lambda_j h_j(\mathbf{x})$$

Augmented Lagrangian function

$$\Phi(\mathbf{x}, \boldsymbol{\lambda}, r_k) = f(\mathbf{x}) + \sum_{j=1}^{m} \lambda_j h_j(\mathbf{x}) + \boxed{r_k \sum_{j=1}^{m} h_j^2(\mathbf{x})}$$

Augmented term to Lagrangian function

r_k : arbitrary constant

Necessary conditions for the minimum of Lagrangian function

$$\frac{\partial L}{\partial x_i} = \frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \boxed{\lambda_j^*} \frac{\partial h_j}{\partial x_i} = 0$$

Necessary conditions for the minimum of Augmented Lagrangian function

$$\frac{\partial \Phi}{\partial x_i} = \frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \boxed{(\lambda_j + 2r_i h_j)} \frac{\partial h_j}{\partial x_i} = 0$$

Find iterative relation

$$\lambda_j^* = \lambda_j + 2r_k h_j \quad j = 1, 2, \ldots, m$$

$$\Downarrow$$

$$\lambda_j^{(k+1)} = \lambda_j^{(k)} + 2r_k h_j(\mathbf{x}^{(k)}) \quad j = 1, 2, \ldots, m$$

Augmented Lagrange Multiplier Method in Equality Constrained Problem (3/4)

Minimize	$f(\mathbf{x})$
Subject to	$h_j(\mathbf{x}) = \mathbf{0}, \quad j = 1,2,\dots,m$

Augmented Lagrangian function

$$\Phi(\mathbf{x},\boldsymbol{\lambda},r_k) = f(\mathbf{x}) + \sum_{j=1}^{m}\lambda_j h_j(\mathbf{x}) + \boxed{r_k \sum_{j=1}^{m} h_j^2(\mathbf{x})}$$

Augmented term to Lagrangian function

r_k **: arbitrary constant**

Iterative relation

$$\lambda_j^{(k+1)} = \lambda_j^{(k)} + 2r_k h_j(\mathbf{x}^{(k)}) \quad j = 1,2,\dots,m$$

1. In the first iteration(k=1), the values of $\lambda_j^{(1)}$ are chosen as zero, the value of r_k is set equal to an arbitrary constant.

2. Find the $\mathbf{x}^{(k)*}$ that minimize Φ by using any unconstrained optimization method and set $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)*}$.

Augmented Lagrange Multiplier Method in Equality Constrained Problem (4/4)

Minimize	$f(\mathbf{x})$
Subject to	$h_j(\mathbf{x}) = \mathbf{0}, \quad j = 1,2,\dots,m$

Augmented Lagrangian function

$$\Phi(\mathbf{x},\boldsymbol{\lambda},r_k) = f(\mathbf{x}) + \sum_{j=1}^{m}\lambda_j h_j(\mathbf{x}) + \boxed{r_k \sum_{j=1}^{m} h_j^2(\mathbf{x})}$$

Augmented term to Lagrangian function

r_k **: arbitrary constant**

Iterative relation

$$\lambda_j^{(k+1)} = \lambda_j^{(k)} + 2r_k h_j(\mathbf{x}^{(k)}) \quad j = 1,2,\dots,m$$

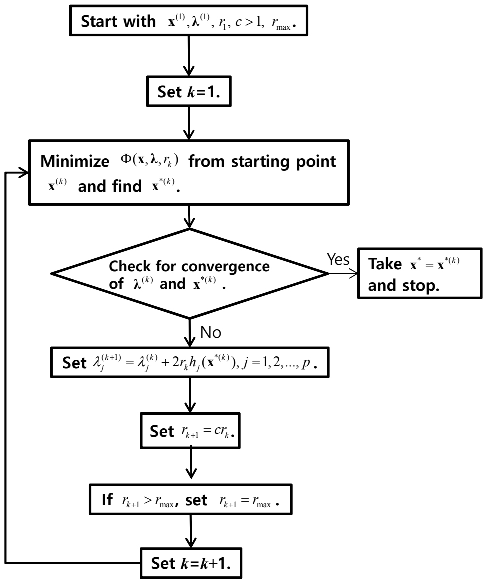
3. The values of $\lambda_j^{(k)}$ and r_k are then updated by using the iterative relation to start the next iteration.

$$r_{k+1} = cr_k,\ c > 1$$

$$\lambda_j^{(k+1)} = \lambda_j^{(k)} + 2r_k h_j(\mathbf{x}^{(k)}) \quad j = 1,2,\dots,m$$

4. If $\left|\lambda_j^{(k+1)} - \lambda_j^{(k)}\right| < \varepsilon$, stop the iteration and take $\mathbf{x}^* = \mathbf{x}^{(k)*}$.

Algorithm of Augmented Lagrange Multiplier Method



Augmented Lagrangian function

$$\Phi(\mathbf{x},\boldsymbol{\lambda},r_k)=f(\mathbf{x})+\sum_{j=1}^{m}\lambda_j h_j(\mathbf{x})+r_k\sum_{j=1}^{m}h_j^2(\mathbf{x})$$

Iterative relation

$$\lambda_j^{(k+1)}=\lambda_j^{(k)}+2r_k h_j(\mathbf{x}^{(k)})\quad j=1,2,...,m$$

$$r_{k+1}=cr_k,\ c>1$$

Augmented Lagrange Multiplier Method in Inequality Constrained Problem

Minimize $f(\mathbf{x})$
Subject to $g_j(\mathbf{x})\le\mathbf{0},\quad j=1,2,...,m$

Augmented Lagrangian function in the inequality constrained problem

$$\Phi(\mathbf{x},\mathbf{u},\mathbf{s},r_k)=f(\mathbf{x})+\sum_{j=1}^{m}u_j[g_j(\mathbf{x})+s_j^2]+r_k\boxed{\sum_{j=1}^{m}[g_j(\mathbf{x})+s_j^2]^2}$$

Augmented term to Lagrangian function

r_k **: arbitrary constant**
s_j **: slack variable**

This function is equivalent to*

$$\Phi(\mathbf{x},\mathbf{u},r_k)=f(\mathbf{x})+\sum_{j=1}^{m}u_j\alpha_j+r_k\sum_{j=1}^{m}\alpha_j^{\,2},\quad \alpha_j=\max\left\{g_j(\mathbf{x}),-\frac{u_j}{2r_k}\right\}$$

Iterative relation

$$u_j^{(k+1)}=u_j^{(k)}+2r_k\alpha_j^{(k)}$$

* Rockafellar, R.T., "The Multiplier Method of Hestenes and Powell Applied to Convex Programming", Journal of Optimization Theory and Applications, 1973

3.4 Descent Function Method

Descent Function Method

Constrained Optimal Design Problem

Minimize $f(\mathbf{x})$

Subject to $\mathbf{h}(\mathbf{x}) = \mathbf{0}$ Equality constraint

$\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$ Inequality constraint

* Descent Function
- Modified objective function by augmenting a penalty to the original objective function
- It has the same meaning with penalty function.

Pshenichny and Danilin suggested a method which transforms the constrained optimization problem into the unconstrained optimization by using the descent function* in 1978.

$V(\mathbf{x}) = \max\{0; |\mathbf{h}|; \mathbf{g}\}$: Maximum penalty by the constraints

$$\Phi(\mathbf{x}) = f(\mathbf{x}) + R \cdot V(\mathbf{x}) \qquad R = \max\left\{R_0,\ r\left(=\sum_{i=1}^{p}|v_i| + \sum_{i=1}^{m} u_i\right)\right\}$$

: Penalty parameters which is the summation of the all Lagrange multipliers (Positive value)

R_0: The positive value defined by user

1) If all constraints are satisfied at the current design point,

$V(\mathbf{x}) = 0 \Rightarrow R \cdot V(\mathbf{x}) = 0$

$\Rightarrow \Phi(\mathbf{x}) = f(\mathbf{x}) + R \cdot V(\mathbf{x}) \Rightarrow f(\mathbf{x})$

➡ If all constraints are satisfied at the current design point, the descent function is the same with the original objective function.

2) If one of more constraints are violated at the current design point,

$R \cdot V(\mathbf{x}) > 0$

$\Rightarrow \Phi(\mathbf{x}) = f(\mathbf{x}) + R \cdot V(\mathbf{x}) > f(\mathbf{x})$

➡ If one of more constraints are violated at the current design point, the value of the positive penalty is augmented to the original objective function.

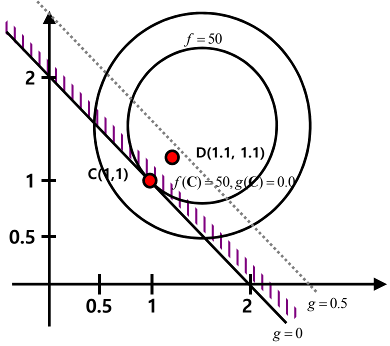
[Reference] The Meaning of the Constant 'R' in the Decent Function (1/2)

Original Problem

Minimize $f(\mathbf{x}) = 100(x_1 - 1.5)^2 + 100(x_2 - 1.5)^2$
Subject to $g(\mathbf{x}) = x_1 + x_2 - 2 \le 0$

$$\Phi(\mathbf{x}) = f(\mathbf{x}) + R \cdot V(\mathbf{x})$$
$$V(\mathbf{x}) = \max\{0; |\mathbf{h}|; \mathbf{g}\}$$
$$R = \max\left\{R_0,\ r\left(= \sum_{i=1}^{p} |v_i| + \sum_{i=1}^{m} u_i\right)\right\}$$

If 'R' is assumed as a constant '10',



Since the constraint is satisfied at the point C(1,1), the value of the decent function is as follows:

$$\Phi(\mathbf{C}) = f(\mathbf{C}) + R \cdot V(\mathbf{C}) = 50 + R \cdot \max\{0, g(\mathbf{C})\}$$
$$= 50 + 10 \cdot \max\{0, 0\} = 50$$

Since the constraint is violated at the point D(1.1, 1.1), the value of the decent function is as follows:

$$\Phi(\mathbf{D}) = f(\mathbf{D}) + R \cdot V(\mathbf{D}) = 32 + R \cdot \max\{0, g(\mathbf{D})\}$$
$$= 32 + 10 \cdot \max\{0, 0.2\} = 32 + 2 = 34$$

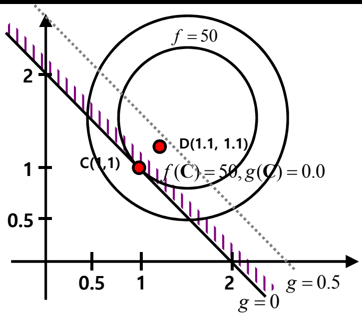
Although the constraint is violated, the value of the decent function is decreased, because the change ("decrease") in the original objective function f is larger than the change ("increase") in the constraint g. Therefore, if the decrease in the original objective function f is larger than the increase in the constraint g, the value of the penalty parameter 'R' has to be increased.

[Reference] The Meaning of the Constant 'R' in the Decent Function (2/2)

Original Problem

Minimize $f(\mathbf{x}) = 100(x_1 - 1.5)^2 + 100(x_2 - 1.5)^2$
Subject to $g(\mathbf{x}) = x_1 + x_2 - 2 \le 0$

$$\Phi(\mathbf{x}) = f(\mathbf{x}) + R \cdot V(\mathbf{x})$$
$$V(\mathbf{x}) = \max\{0; |\mathbf{h}|; \mathbf{g}\}$$
$$R = \max\left\{R_0,\ r\left(= \sum_{i=1}^{p} |v_i| + \sum_{i=1}^{m} u_i\right)\right\}$$



At point C, the value of $-\nabla f(\mathbf{x}^*) = u^* \nabla g(\mathbf{x}^*)$ is as follows.

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix}_{\mathbf{x}^*(1,1)} = \begin{bmatrix} 200(x_1 - 1.5) \\ 200(x_2 - 1.5) \end{bmatrix}_{\mathbf{x}^*(1,1)} = \begin{bmatrix} -100 \\ -100 \end{bmatrix}$$

$$\nabla g(\mathbf{x}) = \begin{bmatrix} \frac{\partial g}{\partial x_1} \\ \frac{\partial g}{\partial x_2} \end{bmatrix}_{\mathbf{x}^*(1,1)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}_{\mathbf{x}^*(1,1)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$u^* = 100$$

If we use the value of the Lagrange Multiplier, 100, as the value of 'R', the value of the decent function at the point D increases by 52.

$$\Phi(\mathbf{D}) = f(\mathbf{D}) + R \cdot V(\mathbf{D}) = 32 + R \cdot \max\{0, g(\mathbf{D})\}$$
$$= 32 + 100 \cdot \max\{0, 0.2\} = 32 + 20 = 52$$

If the change in the objective function($\nabla f(\mathbf{x})$) is larger than the change in the constraint($\nabla g(\mathbf{x})$) respectively, the value of the Lagrange multiplier is increased. Therefore, we use the value of the Lagrange multiplier as the value of 'R'.