# Chapter 16. Calculus Methods of Optimization

Min Soo KIM

Department of Mechanical and Aerospace Engineering
Seoul National University

# Chapter 16. Calculus Methods of Optimization

## 16.1 Continued exploration of calculus methods

- A major portion of this chapter deals with principles of calculus method
- Substantiation for the Largrange multiplier equations will be provided
- Physical interpretation of λ, and test for maxima-minima are also provided

# Chapter 16. Calculus Methods of Optimization

## 16.1 Continued exploration of calculus methods

- Lagrange multipliers (from Chap.8)

$$y = y(x_1, x_2, \cdots, x_n)$$

$$\phi_1(x_1, x_2, \cdots, x_n) = 0$$

$$\vdots$$

$$\phi_m(x_1, x_2, \cdots, x_n) = 0$$

⇒

$$\left.\begin{aligned} &\nabla y - \sum_{i=1}^{m} \lambda_i \nabla \phi_i = 0 \\ &\phi_1(x_1, x_2, \cdots, x_n) = 0 \\ &\vdots \\ &\phi_m(x_1, x_2, \cdots, x_n) = 0 \end{aligned}\right\}$$

n+m scalar equations

## 16.2 The nature of the gradient vector ($\nabla y$)

1. **normal to the surface of constant y**

For arbitrary vector : $dx_1\hat{i}_1 + dx_2\hat{i}_2 + dx_3\hat{i}_3$

To be tangent to the surface : $$dy = \frac{\partial y}{\partial x_1}dx_1 + \frac{\partial y}{\partial x_2}dx_2 + \frac{\partial y}{\partial x_3}dx_3 = 0$$

$$\rightarrow \quad dx_1 = -\frac{(\partial y/\partial x_2)dx_2 + (\partial y/\partial x_3)dx_3}{\partial y/\partial x_1}$$

Tangent vector : $$T = \left[-\frac{(\partial y/\partial x_2)dx_2 + (\partial y/\partial x_3)dx_3}{\partial y/\partial x_1}\right]\hat{i}_1 + dx_2\hat{i}_2 + dx_3\hat{i}_3$$

Gradient vector : $$\nabla y = \frac{\partial y}{\partial x_1}\hat{i}_1 + \frac{\partial y}{\partial x_2}\hat{i}_2 + \frac{\partial y}{\partial x_3}\hat{i}_3$$

$T \cdot \nabla y = 0$ → gradient vector is normal to all tangent vectors

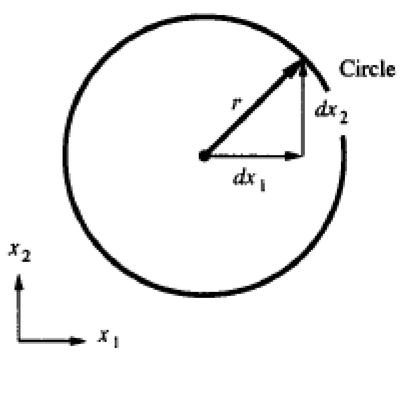
→ gradient vector normal to the surface

## 16.2 The nature of the gradient vector ($\nabla y$)

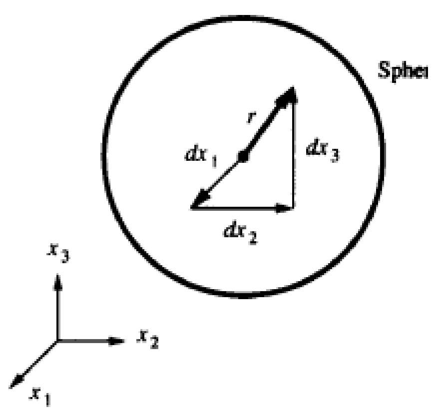
2. **indicate direction of maximum rate of change of y with respect to x**

to find maximum dy : $$dy = \frac{\partial y}{\partial x_1} dx_1 + \frac{\partial y}{\partial x_2} dx_2 + \cdots + \frac{\partial y}{\partial x_n} dx_n$$

constraints : $$(dx_1)^2 + (dx_2)^2 + \cdots + (dx_n)^2 = r^2$$



(a)



(b)

The constraint indicates a circle of radious for 2 dimensions, and a sphere for 3 dimensions

## 16.2 The nature of the gradient vector ($\nabla y$)

using Lagrange multipliers : $\dfrac{\partial y}{\partial x_i} - 2\lambda dx_i = 0 \;\rightarrow\; dx_i = \dfrac{1}{2\lambda}\dfrac{\partial y}{\partial x_i}$

In vector form :

$$dx_1 \hat{i}_1 + dx_2 \hat{i}_2 + \cdots + dx_n \hat{i}_n \rightarrow \frac{1}{2\lambda}\left[\frac{\partial y}{\partial x_1}\hat{i}_1 + \frac{\partial y}{\partial x_2}\hat{i}_2 + \cdots + \frac{\partial y}{\partial x_n}\hat{i}_n\right] = \frac{1}{2\lambda}\nabla y$$

→ $\nabla y$ indicates the direction of maximum change for a given distance in the space

## 16.2 The nature of the gradient vector ($\nabla y$)

3. **points in the direction of increasing y**

small move in the $x_1$-$x_2$ space : $dy = \dfrac{\partial y}{\partial x_1} dx_1 + \dfrac{\partial y}{\partial x_2} dx_2$

If the move is made in the direction of $\nabla y$ :

$$\frac{dx_i}{\partial y / \partial x_i} = c \;\rightarrow\; dx_1 = c(\partial y / \partial x_1),\;\; dx_2 = c(\partial y / \partial x_2)$$

$$dy = c\left[\left(\frac{\partial y}{\partial x_1}\right)^2 + \left(\frac{\partial y}{\partial x_2}\right)^2\right] \geq 0$$

→ dy is equal or greater than zero

→ y always increase in the direction of $\nabla y$

## 16.5 Two variables and one constraint (prove Lagrange multiplier method)

Optimize $y(x_1, x_2)$ subject to $\phi(x_1, x_2) = 0$

Taylor expansion : $\Delta y \approx \left(\frac{\partial y}{\partial x_1}\right)\Delta x_1 + \left(\frac{\partial y}{\partial x_2}\right)\Delta x_2$

$$d\phi = \left(\frac{\partial \phi}{\partial x_1}\right)\Delta x_1 + \left(\frac{\partial \phi}{\partial x_2}\right)\Delta x_2 = 0 \rightarrow \Delta x_1 = -\frac{\partial \phi / \partial x_2}{\partial \phi / \partial x_1}\Delta x_2$$

substituting the result for $\Delta x_1$

$$\Delta y = \left[-\frac{\partial y}{\partial x_1}\frac{\partial \phi / \partial x_2}{\partial \phi / \partial x_1} + \frac{\partial y}{\partial x_2}\right]\Delta x_2$$

## 16.5 Two variables and one constraint (prove Lagrange multiplier method)

In order for no improvements of Δy,

$$\Delta y = \left[ -\frac{\partial y}{\partial x_1} \frac{\partial \phi / \partial x_2}{\partial \phi / \partial x_1} + \frac{\partial y}{\partial x_2} \right] \Delta x_2 = 0 \quad \Rightarrow \quad -\frac{\partial y}{\partial x_1} \frac{\partial \phi / \partial x_2}{\partial \phi / \partial x_1} + \frac{\partial y}{\partial x_2} = 0$$

If λ is defined as $\dfrac{\frac{\partial y}{\partial x_1}}{\frac{\partial \phi}{\partial x_1}}$

$$\Rightarrow \quad \frac{\partial y}{\partial x_2} - \lambda \frac{\partial \phi}{\partial x_2} = 0 \qquad \frac{\partial y}{\partial x_1} - \lambda \frac{\partial \phi}{\partial x_1} = 0$$

## 16.6 Three variables and one constraint

As a same manner with 2 variables, $y(x_1, x_2, x_3)$ is need to be optimized subject to the constraint, $\phi(x_1, x_2, x_3)$

The first degree terms in the Taylor seriers are

$$\left(\frac{\partial y}{\partial x_1}\right)\Delta x_1 + \left(\frac{\partial y}{\partial x_2}\right)\Delta x_2 + (\frac{\partial y}{\partial x_3})\Delta x_3$$

$$\left(\frac{\partial \phi}{\partial x_1}\right)\Delta x_1 + \left(\frac{\partial \phi}{\partial x_2}\right)\Delta x_2 + (\frac{\partial \phi}{\partial x_3})\Delta x_3$$

## 16.6 Three variables and one constraint

Any two of the three variables can be moved independently, but the motion of the third variables must abide by the constraint. Arbitraily choosing $x_2$ as a dependant one,

$$\Delta y = \frac{\partial y}{\partial x_1}\Delta x_1 - \frac{\partial y}{\partial x_2}\left[\frac{\partial \phi/\partial x_1}{\partial \phi/\partial x_2}\Delta x_1 + \frac{\partial \phi/\partial x_3}{\partial \phi/\partial x_2}\Delta x_3\right] + \frac{\partial y}{\partial x_3}\Delta x_3 = 0$$

In order for no improvement to be possible

$$\frac{\partial y}{\partial x_1} - \frac{\partial y}{\partial x_2}\frac{\frac{\partial \phi}{\partial x_1}}{\frac{\partial \phi}{\partial x_2}} = 0 \qquad \frac{\partial y}{\partial x_3} - \frac{\partial y}{\partial x_2}\frac{\frac{\partial \phi}{\partial x_3}}{\frac{\partial \phi}{\partial x_2}} = 0$$

## 16.6 Three variables and one constraint

Define $\lambda$ as

$$\frac{\frac{\partial y}{\partial x_1}}{\frac{\partial \phi}{\partial x_1}}$$

Then

$$\frac{\partial y}{\partial x_1} - \lambda \frac{\partial \phi}{\partial x_1} = 0 \qquad \frac{\partial y}{\partial x_2} - \lambda \frac{\partial \phi}{\partial x_2} = 0 \qquad \frac{\partial y}{\partial x_3} - \lambda \frac{\partial \phi}{\partial x_3} = 0$$

## 16.6 Three variables and one constraint

Define $\lambda$ as

$$\frac{\frac{\partial y}{\partial x_1}}{\frac{\partial \phi}{\partial x_1}}$$

Then

$$\frac{\partial y}{\partial x_1} - \lambda \frac{\partial \phi}{\partial x_1} = 0 \qquad \frac{\partial y}{\partial x_2} - \lambda \frac{\partial \phi}{\partial x_2} = 0 \qquad \frac{\partial y}{\partial x_3} - \lambda \frac{\partial \phi}{\partial x_3} = 0$$

## 16.8 Alternate expression of constrained optimization problem

optimize $y(x_1, x_2)$

subject to $\phi(x_1, x_2) = b$

unconstrained function $L(x_1, x_2) = y(x_1, x_2) - \lambda[\phi(x_1, x_2) - b]$

optimum occurs where $\nabla L = 0$

$$\left.\begin{aligned} &\frac{\partial L}{\partial x_1} = \frac{\partial y}{\partial x_1} - \lambda \frac{\partial \phi}{\partial x_1} = 0 \\ &\frac{\partial L}{\partial x_2} = \frac{\partial y}{\partial x_2} - \lambda \frac{\partial \phi}{\partial x_2} = 0 \\ &\phi(x_1, x_2) - b = 0 \end{aligned}\right\} \quad \text{find optimum point}$$

## 16.9 Interpretation λ of as the sensitivity coefficient

sensitivity coefficient (SC) :

$$y(x_1, x_2) \;\rightarrow\; SC = \frac{\partial y^*}{\partial b} = \frac{\partial y^*}{\partial x_1}\frac{\partial x_1}{\partial b} + \frac{\partial y^*}{\partial x_2}\frac{\partial x_2}{\partial b} \quad \cdots(1)$$

$$\phi(x_1, x_2) = b \;\rightarrow\; \frac{\partial \phi}{\partial b} = \frac{\partial \phi}{\partial x_1}\frac{\partial x_1}{\partial b} + \frac{\partial \phi}{\partial x_2}\frac{\partial x_2}{\partial b} - 1 = 0 \quad \cdots(2)$$

$$\lambda = \frac{(\partial y^* / \partial x_1^*)}{(\partial \phi / \partial x_1^*)} = \frac{(\partial y^* / \partial x_2^*)}{(\partial \phi / \partial x_2^*)}$$

$$(2) \times \lambda \;\rightarrow\; \frac{\partial y}{\partial x_1}\frac{\partial x_1}{\partial b} + \frac{\partial y}{\partial x_2}\frac{\partial x_2}{\partial b} - \lambda = 0$$

$$(1) \;\rightarrow\; SC = \lambda$$