Ch.4 Optimality Conditions

Review Calculus & Algebra

PART - 2

ME511 – Principle of Optimum Design

Lecturer :

ANDIKA AJI WIJAYA, S.T, M.Sc

College of Engineering
Email : andika@ubt.edu.sa

جامعة الأعمال و التكنولوجيا
UNIVERSITY OF BUSINESS AND TECHNOLOGY

Introduction

- In previous lecture we have discussed the optimality condition to check if the solution we get is optimum solution for single variable optimization.
- What if our optimization problem contain multiple design variables?
- In this case, it is important to know how to calculate derivatives of functions of several variables to solve optimization problems as well as to perform matrix operation

Gradient of a Function

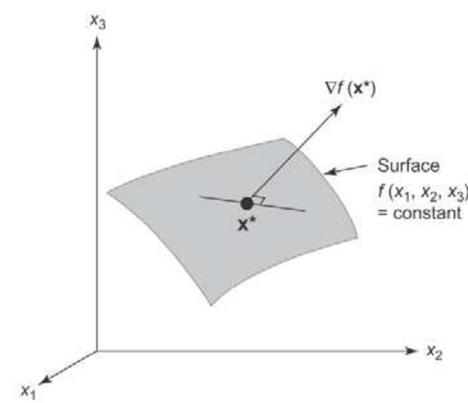
First Partial Derivatives

For a function $f(\mathbf{x})$ of n variables, the first partial derivatives are written as

$$\frac{\partial f(\mathbf{x})}{\partial x_i}; \quad i = 1 \text{ to } n$$

The n partial derivatives are usually arranged in a column vector known as the *gradient* of the function $f(\mathbf{x})$. The gradient is written as $\partial f/\partial \mathbf{x}$ or $\nabla f(\mathbf{x})$. Therefore,

Gradient → $$\mathbf{c} = \nabla f(\mathbf{x}) = \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} \\ \frac{\partial f(\mathbf{x})}{\partial x_2} \\ \vdots \\ \frac{\partial f(\mathbf{x})}{\partial x_n} \end{bmatrix}$$

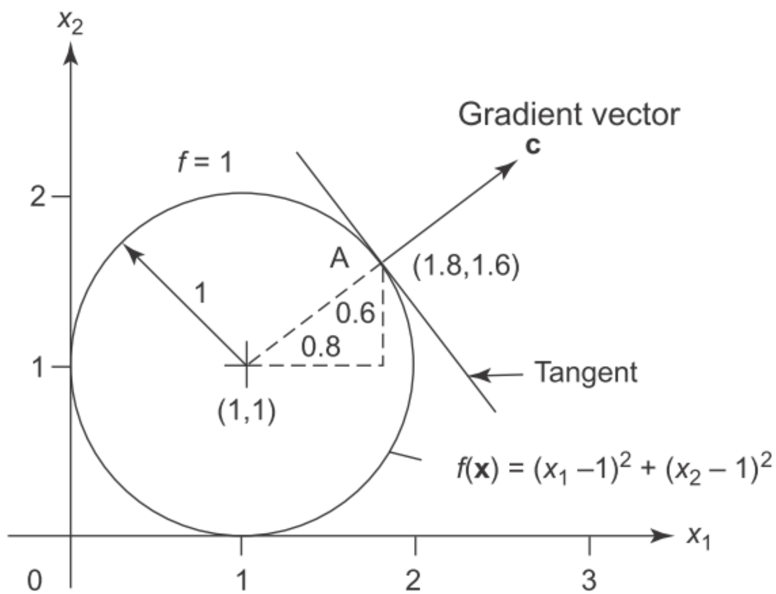


the gradient vector is normal to the tangent plane at the point x

Gradient of a Function - Example

Calculate the gradient of $f(\mathbf{x}) = (x_1 - 1)^2 + (x_2 - 1)^2$ at the point $\boldsymbol{x}^* = (1.8, 1.6)$

For $f(1.8, 1.6) = (1.8 - 1)^2 + (1.6 - 1)^2 = 1$ → a circle with center of (1,1) and radius of 1



$$\frac{\partial f}{\partial x_1}(1.8, 1.6) = 2(x_1 - 1) = 2(1.8 - 1) = 1.6$$

$$\frac{\partial f}{\partial x_2}(1.8, 1.6) = 2(x_2 - 1) = 2(1.6 - 1) = 1.2$$

$$c = \begin{bmatrix} 1.6 \\ 1.2 \end{bmatrix} = [1.6 \quad 1.2]^T$$

2^{nd} Derivatives of Functions –

Second Partial Derivatives

- If we differentiated again, we obtain second partial derivatives of $f(x)$

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j}; \quad i,\ j = 1 \text{ to } n$$

- If we arrange them in matrix form, it is known as "***Hessian matrix***", written as **H(*x*)**

$$\mathbf{H}(\mathbf{x}) = \nabla^2 f(\mathbf{x}) = \left[\frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j}\right]_{n \times n}$$

Hessian Matrix - Example

Derive with respect to x_2 →

For the following function, calculate the gradient vector and the Hessian matrix at the point (1, 2):

$$f(\mathbf{x}) = x_1^3 + x_2^3 + 2x_1^2 + 3x_2^2 - x_1x_2 + 2x_1 + 4x_2$$

Solution

The first partial derivatives of the function are given as

$$\mathbf{c} = \nabla f(\mathrm{x}^*) = \begin{bmatrix} \dfrac{\partial f(\mathrm{x}^*)}{\partial x_1} \\ \dfrac{\partial f(\mathrm{x}^*)}{\partial x_2} \end{bmatrix}$$

$$\frac{\partial f}{\partial x_1} = 3x_1^2 + 4x_1 - x_2 + 2$$

$$\frac{\partial f}{\partial x_2} = 3x_2^2 + 6x_2 - x_1 + 4$$

$$\mathbf{c} = \begin{bmatrix} 7 \\ 27 \end{bmatrix}$$

The second partial derivatives of the function are calculated by

$$\mathbf{H} \text{ or } \nabla^2 f = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} \end{bmatrix}$$

$$\frac{\partial^2 f}{\partial x_1^2} = 6x_1 + 4; \quad \frac{\partial^2 f}{\partial x_1 \partial x_2} = -1;$$

$$\frac{\partial^2 f}{\partial x_2 \partial x_1} = -1; \quad \frac{\partial^2 f}{\partial x_2^2} = 6x_2 + 6.$$

$$\mathbf{H}(\mathbf{x}) = \begin{bmatrix} 6x_1 + 4 & -1 \\ -1 & 6x_2 + 6 \end{bmatrix}$$

The Hessian matrix at the point (1, 2)

$$\mathbf{H}(1, 2) = \begin{bmatrix} 10 & -1 \\ -1 & 18 \end{bmatrix}$$

Eigenvalues

- In optimization, it is necessary to calculate the ***eigenvalues of Hessian matrix*** to check if it satisfy the sufficient condition.
- If the eigenvalues are **positive definite** then the solution satisfy the sufficient condition for ***minimum point***
- The eigenvalues can be found by solving the following:

$$\left|\mathbf{A} - \lambda \mathbf{I}\right| = 0$$

: where **A** is the Hessian Matrix

In order to do this we need to solve it using matrix operation

Taylor's Expansion

- The idea of ***Taylor's expansion*** is fundamental to the development of optimum design and numerical methods
- A function can be approximated by polynomials of any point in terms of its value and derivatives using Taylor's expansion.

$$f(x) = f(x^*) + \underbrace{\frac{df(x^*)}{dx}}_{\text{1st derivative}}(x - x^*) + \frac{1}{2}\underbrace{\frac{d^2f(x^*)}{dx^2}}_{\text{2nd derivative}}(x - x^*)^2 + R$$

where R is the remainder term that is smaller in magnitude than the previous terms. It sometimes can be neglected it the value is insignificant

Taylor's Expansion - Example

Using Taylor expansion, approximate $f(x) = \cos x$ around the point $x^* = 0$

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

Solution

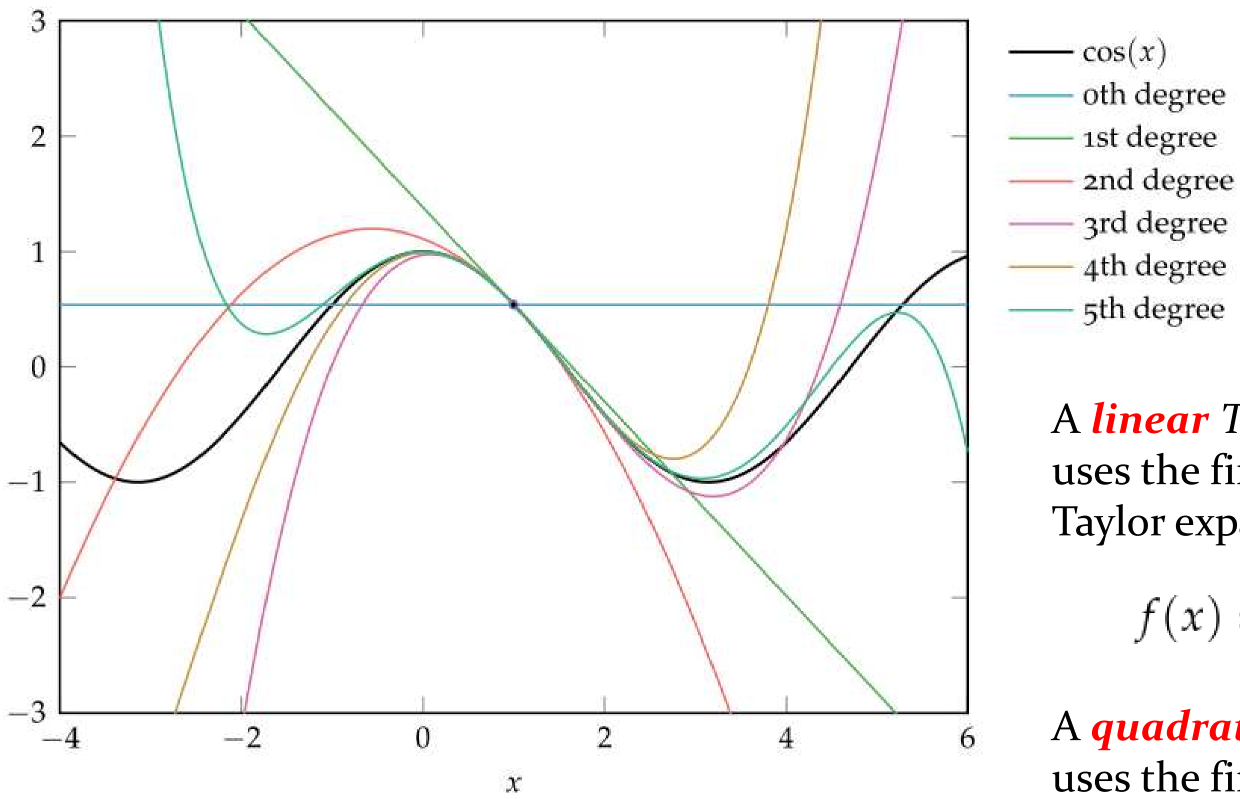
Derivatives of the function $f(x)$ are given as

$$\frac{df}{dx} = -\sin x, \quad \frac{d^2f}{dx^2} = -\cos x$$

Therefore, the second-order Taylor's expansion for $\cos x$ at the point $x^* = 0$ is given as

$$\cos x \approx \cos 0 - \sin\ 0(x - 0) + \frac{1}{2}(-\cos\ 0)(x - 0)^2 = 1 - \frac{1}{2}x^2$$

Taylor's Expansion - Example



$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

A ***linear*** *Taylor approximation* uses the first two terms of the Taylor expansion

$$f(x) \approx f(a) + f'(a)(x-a)$$

A ***quadratic*** Taylor approximation uses the first three terms:

$$f(x) \approx f(a) + f'(a)(x-a) + \frac{1}{2} f''(a)(x-a)^2$$

Vector & Matrix Algebra

- Consider this system of two simultaneous linear equations in three unknowns

$$x_1 + 2x_2 + 3x_3 = 6$$
$$-x_1 + 6x_2 - 2x_3 = 3$$

- We can represent above equations in the matrix form as

$$\begin{bmatrix} 1 & 2 & 3 \\ -1 & 6 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \end{bmatrix} \quad \Rightarrow \quad \mathbf{Ax} = \mathbf{B}$$

Matrix **A** Column Vector ***x*** Matrix **B**

Matrix Operations

- Let **A** and **B** are matrix of 2x2 below

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

1. Scaling with t → tA =

2. A + B =

3. Multiplication (inner product) → A.B =

4. Determinant of A → $|\mathbf{A}| = a_{11}a_{22} - a_{12}a_{21}$

5. Invers of A → $\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|}\begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$

UBT

جامعة الأعمال و التكنولوجيا

UNIVERSITY OF BUSINESS AND TECHNOLOGY

END OF THE SLIDES