Gradient-Based Non-Linear Optimization – Lecture 02

Dr. Serhat Hosder

Associate Professor of Aerospace Engineering

Mechanical and Aerospace Engineering

290B Toomey Hall

Missouri S&T

Rolla, MO 65409

Phone: 573-341-7239

E-mail: hosders@mst.edu

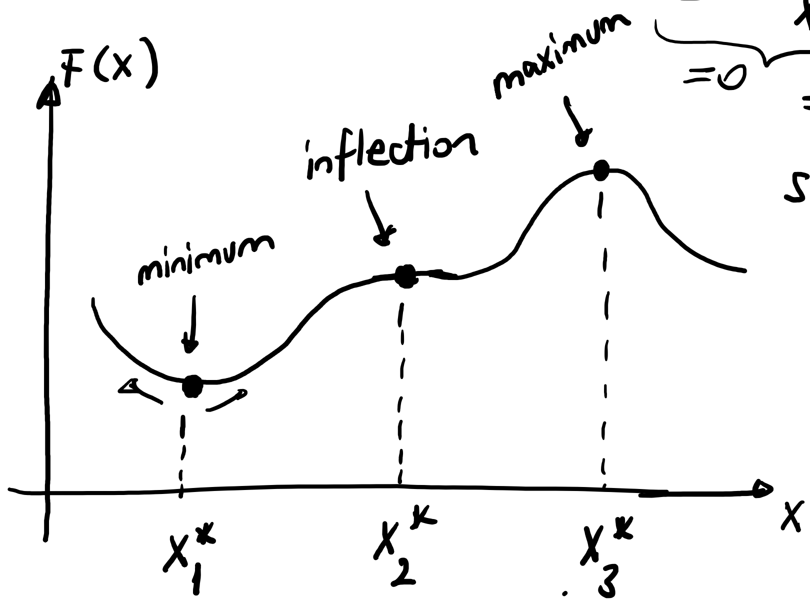
Optimality Conditions:

(1) One-Dimensional Case:

Consider an objective function $F(x)$ of a single variable "x":

→ Taylor series expansion at $x = x^*$ where x^* is the optimum

$$F(x^* + \Delta x) - F(x^*) = \underbrace{\left.\frac{dF}{dx}\right|_{x^*} \cdot \Delta x}_{=0} + \frac{1}{2!}\left.\frac{d^2F}{dx^2}\right|_{x^*} \Delta x^2 + \frac{1}{3!}\left.\frac{d^3F}{dx^3}\right|_{x^*} \cdot \Delta x^3 + \dots$$



$\underbrace{x_1^*, x_2^*, x_3^*}$ = stationary points for $F(x)$

For all three stationary points ⇒ $\boxed{\left.\frac{dF}{dx}\right|_{x^*} = 0}$ Necessary condition for a stationary point

For a minimum → $\left.\frac{d^2F}{dx^2}\right|_{x^*} > 0$

For a maximum → $\left.\frac{d^2F}{dx^2}\right|_{x^*} < 0$

For an inflection point $\left.\frac{d^2F}{dx^2}\right|_{x^*} = 0$

② Multi-Dimensional Case (Optimality Conditions)

→ Consider a function $F(x_1, x_2)$ of two variables $\vec{x} = \begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix}$

→ Write Taylor Series Expansion at $\vec{x}^*$ where $\vec{x}^* = \begin{Bmatrix} x_1^* \\ x_2^* \end{Bmatrix}$

$$F(x_1^* + \Delta x_1, x_2^* + \Delta x_2) = F(x_1^*, x_2^*) + \frac{\partial F}{\partial x_1}\Big|_{\vec{x}^*} \Delta x_1 + \frac{\partial F}{\partial x_2}\Big|_{\vec{x}^*} \Delta x_2$$

$$+ \frac{1}{2!}\left[\frac{\partial^2 F}{\partial x_1^2}\Big|_{\vec{x}^*} \Delta x_1^2 + 2 \frac{\partial^2 F}{\partial x_1 \partial x_2}\Big|_{\vec{x}^*} \Delta x_1 \Delta x_2 + \frac{\partial^2 F}{\partial x_2^2}\Big|_{\vec{x}^*} \Delta x_2^2 \right] + \cdots$$

In Matrix Form:

$$F(\vec{x}^* + \vec{\Delta x}) = F(\vec{x}^*) + (\vec{\Delta x})^T \cdot \nabla F(\vec{x}^*) + \frac{1}{2} (\vec{\Delta x})^T \cdot [H(\vec{x}^*)] \cdot \vec{\Delta x} + \cdots$$

$$\vec{\Delta x} = \begin{Bmatrix} \Delta x_1 \\ \Delta x_2 \end{Bmatrix}, \quad \underbrace{\nabla F(\vec{x}^*)}_{\text{Gradient vector}} = \begin{Bmatrix} \frac{\partial F}{\partial x_1} \\ \frac{\partial F}{\partial x_2} \end{Bmatrix}_{\vec{x}^*} ; \quad \underbrace{H(\vec{x}^*)}_{\text{Hessian matrix}} = \begin{bmatrix} \frac{\partial^2 F}{\partial x_1^2} & \frac{\partial^2 F}{\partial x_1 \partial x_2} \checkmark \\ \frac{\partial^2 F}{\partial x_2 \partial x_1} \checkmark & \frac{\partial^2 F}{\partial x_2^2} \end{bmatrix}_{\vec{x}^*}$$

MISSOURI S&T

$$F(\vec{x}^* + \vec{\Delta x}) - F(\vec{x}^k) = \underbrace{(\vec{\Delta x})^T \cdot \nabla F(\vec{x}^k)}_{=0} + \frac{1}{2}(\vec{\Delta x})^T \cdot [H(\vec{x}^*)] \cdot \vec{\Delta x} + \cdots$$

→ For $\vec{x}^k = \begin{Bmatrix} x_1^k \\ x_2^k \end{Bmatrix}$ to be a stationary point, $\nabla F(\vec{x}^k) = 0$ (Necessary Condition)

For <u>small enough and any $\vec{\Delta x}$</u>

→ <u>To have $\vec{x}^k$ as a minimum</u> : $(\vec{\Delta x})^T \cdot [H(\vec{x}^*)] \cdot \vec{\Delta x} > 0$

* Then $[H(\vec{x}^*)]$ should be a positive definite matrix
* A positive definite matrix is a matrix with all its eigenvalues being positive.

For n-dimensional case

If $[H(\vec{x}^k)]$ has all its "n" eigenvalues positive, then

$[H(\vec{x}^k)] \rightarrow$ positive definite matrix

$$H = \begin{bmatrix} \frac{\partial^2 F}{\partial x_1^2} & \frac{\partial^2 F}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 F}{\partial x_1 \partial x_n} \\ \vdots & & \ddots & \\ \frac{\partial^2 F}{\partial x_n \partial x_1} & \frac{\partial^2 F}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 F}{\partial x_n^2} \end{bmatrix}$$

$\rightarrow$ If $(\vec{\Delta x})^T \cdot [H(\vec{x^*})] \cdot \vec{\Delta x} \geq 0 \Rightarrow \vec{x^*}$ can be a minimum, a saddle point or an inflection point.

$\hookrightarrow$ Positive "semi-definite" matrix

$\rightarrow$ If $(\vec{\Delta x})^T \cdot [H(\vec{x^*})] \vec{\Delta x} < 0 \Rightarrow \vec{x^*}$ is a maximum

$\hookrightarrow$ Negative definite matrix (all eigenvalues of $[H(\vec{x^*})]$ negative)

$\rightarrow$ If $(\vec{\Delta x})^T \cdot [H(\vec{x^*})] \vec{\Delta x} \leq 0 \Rightarrow \vec{x^*}$ can be a maximum, a saddle point or an inflection point.

$\rightarrow$ If $[H(\vec{x^*})]$ has mixed eigenvalues (some positive, some negative)

$\hookrightarrow x^* \rightarrow$ saddle point.

Example for Saddle point of a 2-D function