Gradient-Based Non-Linear Optimization – Lecture 01

Dr. Serhat Hosder

Associate Professor of Aerospace Engineering

Mechanical and Aerospace Engineering

290B Toomey Hall

Missouri S&T

Rolla, MO 65409

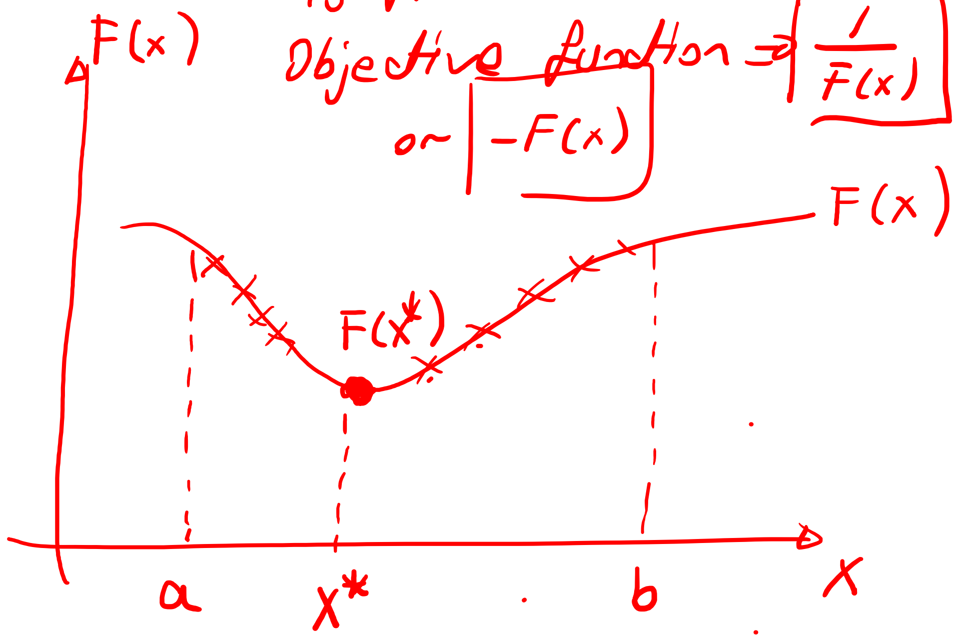
Phone: 573-341-7239

E-mail: hosders@mst.edu

General Problem Statement

1-D optimization

To find the maximum Objective function $\Rightarrow \boxed{\dfrac{1}{F(x)}}$ or $\boxed{-F(x)}$ $\Rightarrow$ minimize $\dfrac{1}{F(x)}$



$F(x) \equiv$ Objective Function

↳ The function to be minimized

$x \equiv$ design variable

$$\underbrace{a \leq x \leq b}_{\text{Design space}}$$

Design space (the region where the minimum is searched)

$F(x^*) =$ minimum of $F(x)$ in $x \in [a, b]$

$x^* =$ minimizer (optimum point)

$\boxed{\text{Find the minimum of } F(x)}$ $\Rightarrow$ For minimum

Classification of Optimization methods

① Optimization (Based on the Objective Function and Constraint Functions)

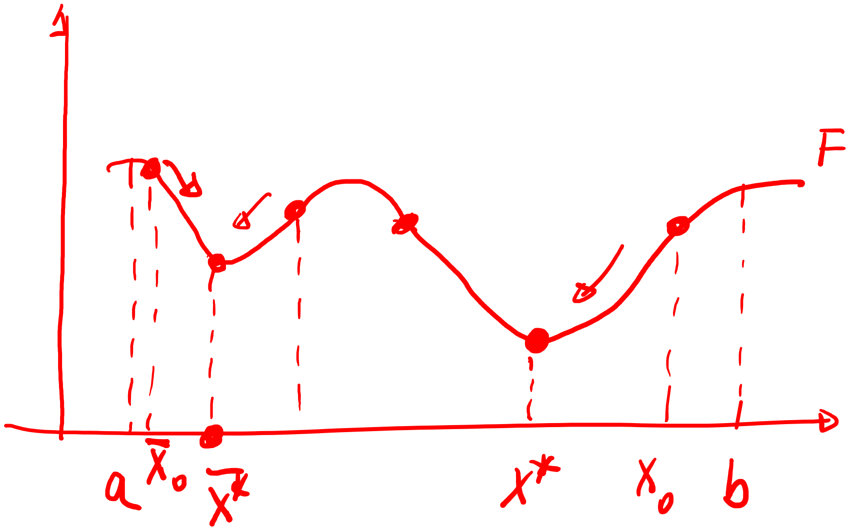
- Linear
- Non-Linear ✓

② Optimization (Based on the existence of any Constraints)

- Constraint
- Unconstraint ✓

③ Optimization (Depending on the search approach for the minimum in the design space)

- Global
- Local ✓ (Gradient-based methods)



$F(x)$ = objective Function

x = design variable

x^* = global optimum

$\bar{x}^*$ = local optimum

MISSOURI S&T

General Problem Statement for Non-Linear Constraint Optimization:

Minimize $F(\vec{x})$ where $F(\vec{x}) \equiv$ objective Function

$$\vec{x} = \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Bmatrix} \Rightarrow \text{Design Variable Vector}$$

(n number of design variables)

Subject to

$g_j(\vec{x}) \leq 0 \quad j=1,\ldots,m \Rightarrow$ Inequality Constraints

$h_k(\vec{x}) = 0 \quad k=1,\ldots,\ell \Rightarrow$ Equality constraints

$x_i^L \leq x_i \leq x_i^U \quad i=1,\ldots,n \Rightarrow$ Side constraints

Example 1-2 Constrained function minimization

Figure 1-2a depicts a tubular column of height h which is required to support a concentrated load P as shown. We wish to find the mean diameter D and the wall thickness t to minimize the weight of the column. The column weight is given by

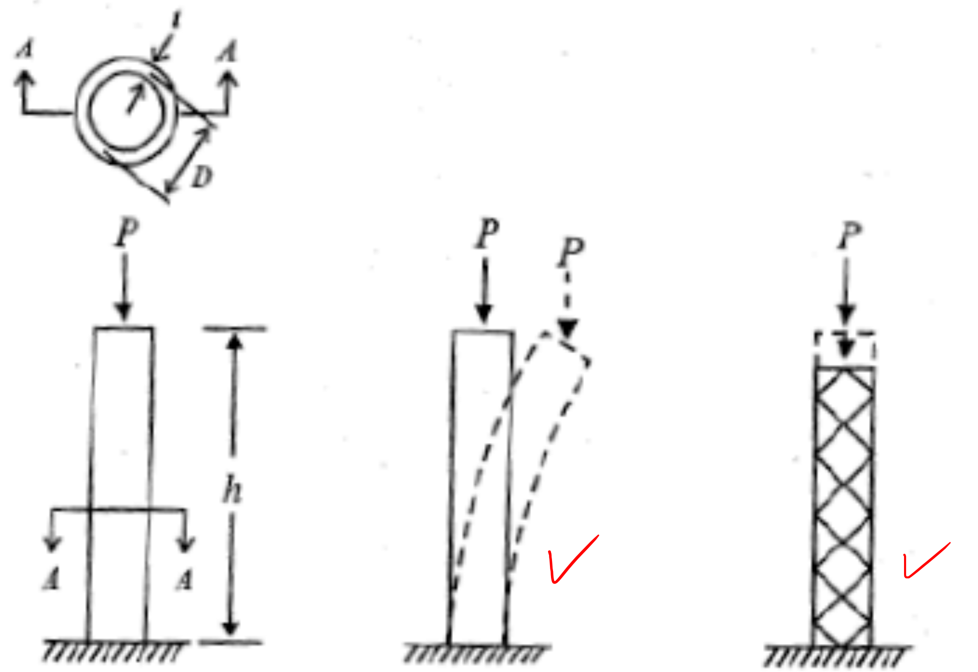
$$\boxed{W = \rho A h = \rho \pi D t h} \tag{1-4}$$

where A is the cross-sectional area and ρ is the material's unit weight.

We will consider the axial load only, and for simplicity will ignore any eccentricity, lateral loads, or column imperfections. The stress in the column is given by

$$\sigma = \frac{P}{A} = \frac{P}{\pi D t} \tag{1-5}$$

D, t } Design variables



(a) Column (b) Euler buckling (c) Local buckling

Figure 1-2 Column design for least weight.

where stress is taken as positive in compression. In order to prevent material failure, this stress must not exceed the allowable stress $\bar{\sigma}$. In addition to preventing material failure, the stress must not exceed that at which Euler buckling or local shell buckling will occur, as shown in Figs. 1-2b and c. The stress at which Euler buckling occurs is given by

$$\sigma_b = \frac{\pi^2 EI}{4Ah^2} = \frac{\pi^2 E(D^2 + t^2)}{8h^2} \quad \checkmark \tag{1-6}$$

where E = Young's modulus
I = moment of inertia

The stress at which shell buckling occurs is given by

$$\sigma_s = \frac{2Et}{D\sqrt{3(1-\nu^2)}} \quad \checkmark \tag{1-7}$$

where ν = Poisson's ratio

The column must now be designed so that the magnitude of the stress is less than the minimum of $\bar{\sigma}$, σ_b, and σ_s. These requirements can be written algebraically as

Inequility Constraints

$$\sigma \le \bar{\sigma} \qquad \sigma - \bar{\sigma} \le 0 \tag{1-8}$$

$$\sigma \le \sigma_b \qquad \sigma - \bar{\sigma}_b \le 0 \tag{1-9}$$

$$\sigma \le \sigma_s \qquad \sigma - \sigma_s \le 0 \tag{1-10}$$

$$\frac{\sigma}{\bar{\sigma}} - 1 \le 0$$

$$\frac{\sigma}{\bar{\sigma}_b} - 1 \le 0$$

$$\frac{\sigma}{\bar{\sigma}_s} - 1 \le 0$$

$$t - D \le 0$$

In addition to the stress limitations, the design must satisfy the geometric conditions that the mean diameter be greater than the wall thickness and that both the diameter and thickness be positive

$$D \ge t \rightarrow t - D \le 0 \tag{1-11}$$

$$D \ge 10^{-6} \tag{1-12}$$

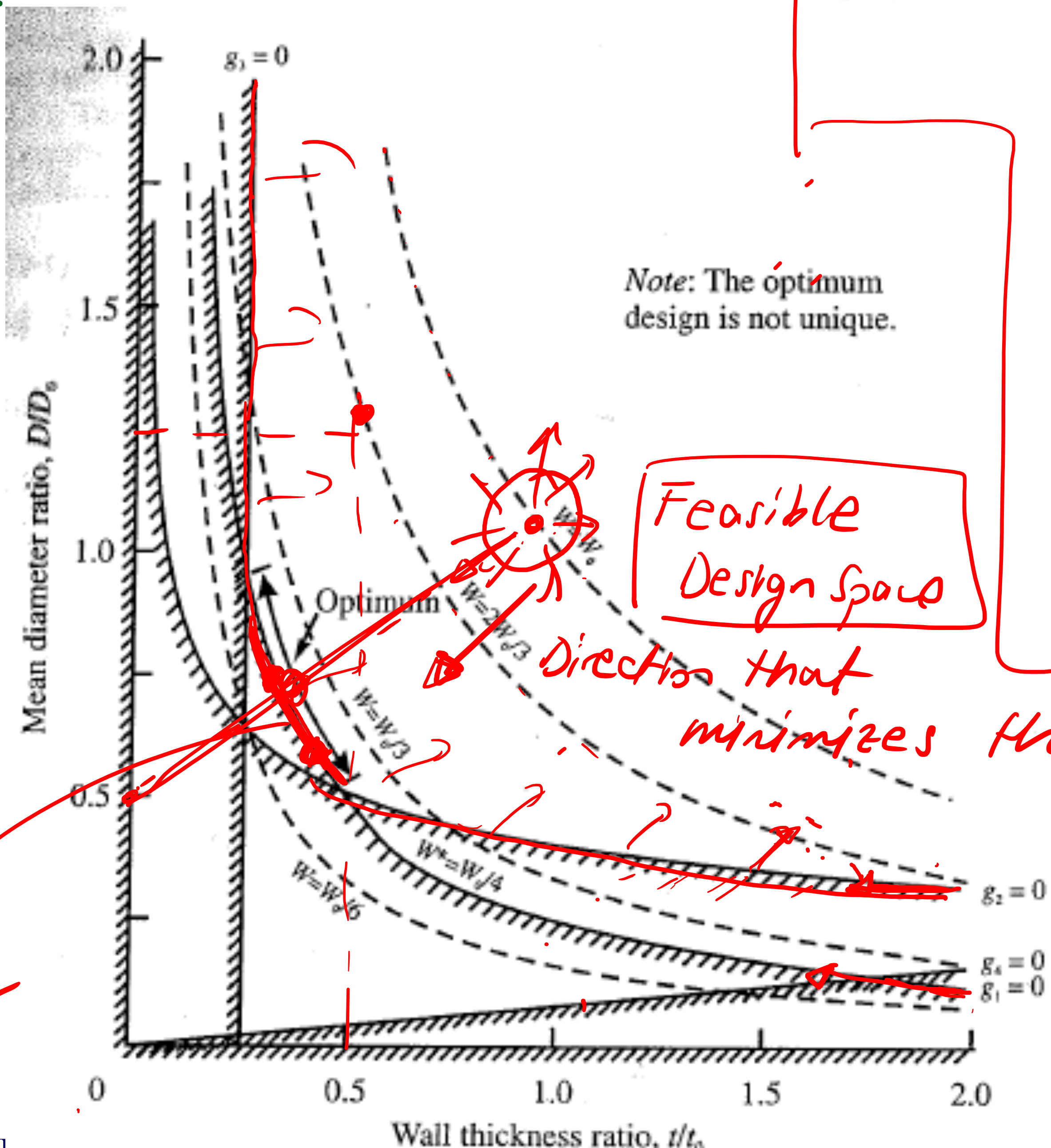
$$t \ge 10^{-6} \tag{1-13}$$

side constraints

Bounds of 10^{-6} are imposed on D and t to ensure that σ in Eq. (1-5) and σ_s in Eq. (1-7) will be finite.

The design problem can now be stated compactly as

MISSOURI S&T



Direction that minimizes the objective function

Pareto Front

Minimize: $W(D,t) = \rho \pi D t h$

$g_1(D,t) = \frac{\sigma}{\bar{\sigma}} - 1 \leq 0$

$g_2(D,t) = \frac{\sigma}{\sigma_b} - 1 \leq 0$

$g_3(D,t) = \frac{\sigma}{\sigma_s} - 1 \leq 0$

$g_4(D,t) = t - D \leq 0$

$D \geq 10^{-6}$

$t \geq 10^{-6}$

Iterative Optimization Procedure

Specify / Start with $\vec{X}_0 = \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Bmatrix}_0 \Rightarrow$ Initial design variable vector

The most common form of the iterative procedure:

$$\vec{X}_k = \vec{X}_{k-1} + \alpha_k^* \vec{S}_k$$

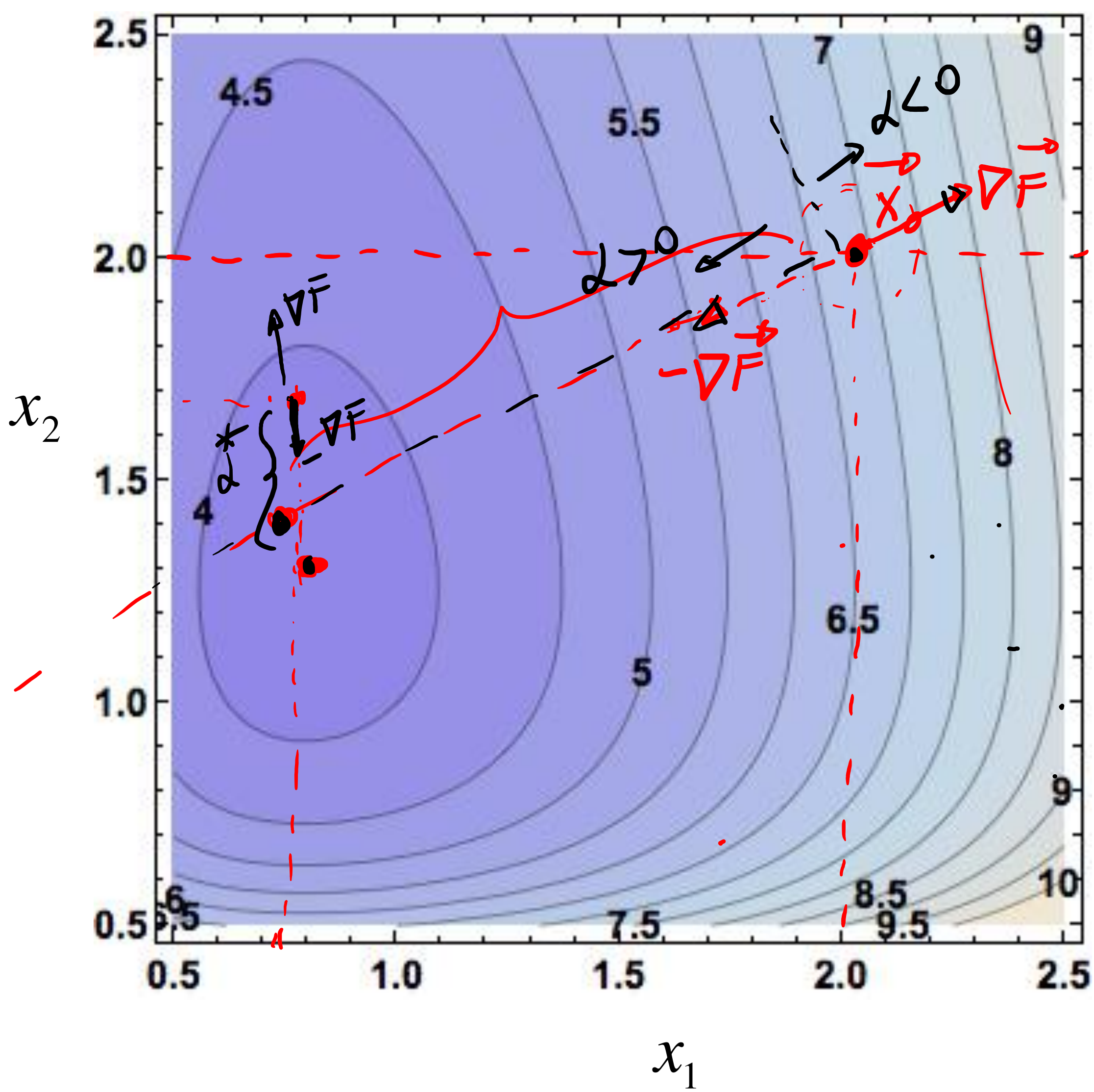
k = iteration number

$\vec{X}_k$ = design variable vector obtained at the end of iteration "k".

$\vec{S}_k$ = Search direction

α_k^* = optimum value of a scaler multiplier determining the amount of change in $\vec{X}$ vector along the $\vec{S}$ direction.

An Example



Unconstraint non-linear optimization with two design variables x_1 & x_2.

objective fnction

$$F(x_1, x_2) = x_1^2 + \frac{1}{x_1} + x_2 + \frac{1}{x_2^2}$$

$\vec{x} = \begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix} \Rightarrow$ Design Variable Vector

$\vec{x}_0 = \begin{Bmatrix} 2.0 \\ 2.0 \end{Bmatrix} \Rightarrow$ initial design variable vector

$$\boxed{F(x_1, x_2) = x_1^2 + \frac{1}{x_1} + x_2 + \frac{1}{x_2^2}}$$

$$\vec{x}_0 = \begin{Bmatrix} 2.0 \\ 2.0 \end{Bmatrix}$$

$$\nabla F(x_1, x_2)\Big|_{\vec{x}_0} = \frac{\partial F}{\partial x_1}\Big|_{\vec{x}_0} \cdot \vec{e}_1 + \frac{\partial F}{\partial x_2}\Big|_{\vec{x}_0} \cdot \vec{e}_2$$

$\vec{e}_1$ = unit normal vector in x_1 direction

$\vec{e}_2$ = unit normal vector in x_2 direction

$$\left.\begin{aligned} \frac{\partial F}{\partial x_1} &= 2x_1 - \frac{1}{x_1^2} \;; \quad \frac{\partial F}{\partial x_1}\Big|_{\vec{x}_0} = 3.75 \\ \frac{\partial F}{\partial x_2} &= 1 - \frac{2}{x_2^3} \;; \quad \frac{\partial F}{\partial x_2}\Big|_{\vec{x}_0} = 0.75 \end{aligned}\right\} \quad \vec{\nabla} F(\vec{x}_0) = 3.75\,\vec{e}_1 + 0.75\,\vec{e}_2$$

$$\tilde{s} = -\nabla F(\vec{x}_0) = -3.75\,\vec{e}_1 - 0.75\,\vec{e}_2$$

$$\boxed{\vec{s} = -1.0\,\vec{e}_1 - 0.2\,\vec{e}_2}$$

Search direction at $\vec{x}_0$

{ We scaled each component of $\tilde{s}$ with $\max|\tilde{s}_i|$ $i = 1, 2, \ldots n$ }

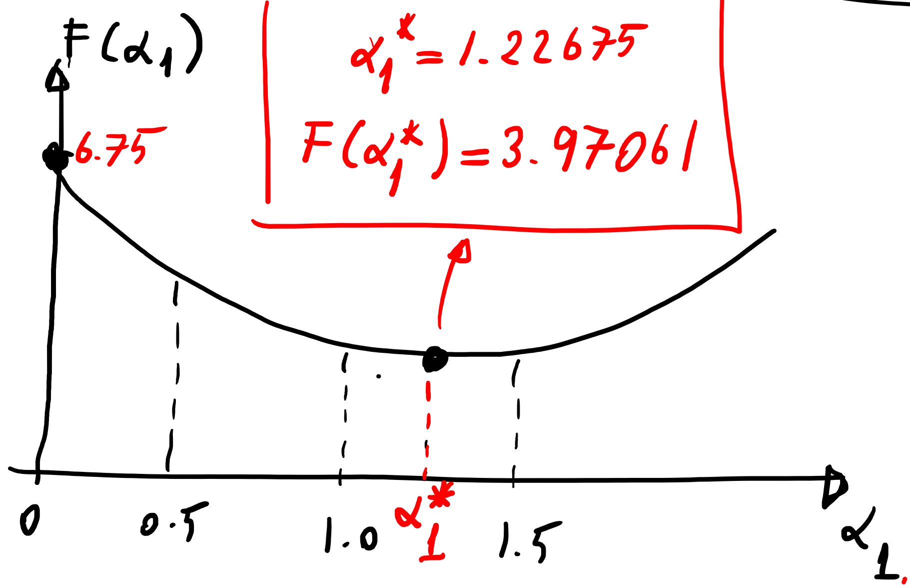
Another approach for scaling: $\vec{s} = \dfrac{-\nabla F}{\|\nabla F\|}$ $\|\nabla F\| \rightarrow$ length of $|\nabla F|$

$$\vec{X}_k = \vec{X}_{k-1} + \alpha_k^* \vec{S}_k \quad , \quad \text{For the } 1^{st} \text{ iteration } (k=1) \quad \boxed{\vec{X}_1 = \vec{X}_0 + \alpha_1 \vec{S}_1}$$

$$\vec{X}_0 = \begin{Bmatrix} 2.0 \\ 2.0 \end{Bmatrix} \quad ; \quad \vec{S}_1 = \begin{Bmatrix} -1.0 \\ -0.2 \end{Bmatrix} \qquad \begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix}_1 = \begin{Bmatrix} 2.0 \\ 2.0 \end{Bmatrix} + \alpha_1 \cdot \begin{Bmatrix} -1.0 \\ -0.2 \end{Bmatrix}$$

$$(x_1)_1 = 2.0 - \alpha_1 \quad ; \quad (x_2)_1 = 2.0 - 0.2\alpha_1$$

$$F(\vec{X}_1) = F\big(x_1(\alpha_1), x_2(\alpha_1)\big) = \underbrace{(2-\alpha_1)^2 + \frac{1}{(2-\alpha_1)} + (2-0.2\alpha_1) + \frac{1}{(2-0.2\alpha_1)^2}}_{F(\alpha_1)}$$



$$\boxed{\begin{aligned} \alpha_1^* &= 1.22675 \\ F(\alpha_1^*) &= 3.97061 \end{aligned}}$$

α_1^* = optimum value of α_1

$\boxed{\text{For } \alpha_1 = 0 \rightarrow F(0) = 6.75}$

For $\alpha_1 = 0.5 \rightarrow F(0.5) = 5.0937$

For $\alpha_1 = 1.0 \rightarrow F(1.0) = 4.1086$

For $\alpha_1 = 1.5 \rightarrow F(1.5) = 4.2960$

$$\boxed{\alpha_1^* = 1.22675}$$

$$\alpha_1^* = 1.22675$$

$$F(\alpha_1^*) = 3.97061$$

$$\vec{X}_1 = \vec{X}_0 + \alpha_1^* \vec{S}_1$$

$$\begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix}_1 = \begin{Bmatrix} 2.0 \\ 2.0 \end{Bmatrix} + 1.22675 \begin{Bmatrix} -1.0 \\ -0.2 \end{Bmatrix}$$

$$\begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix}_1 = \begin{Bmatrix} 0.77325 \\ 1.75465 \end{Bmatrix}$$

$$\left.\begin{aligned} F(\vec{x}_1) &= 3.97061 \\ F(\vec{x}_0) &= 6.75 \end{aligned}\right\}$$

40% reduction in $F(x_1, x_2)$ in one iteration

MISSOURI S&T

Some Convergence Criteria in Iterative Optimization Algorithms:

① Using Objective Function values:

Absolute $\Rightarrow$ $\left| F(\vec{x}_k) - F(\vec{x}_{k-1}) \right| \leq Tol_1$

Relative $\Rightarrow$ $\dfrac{\left| F(\vec{x}_k) - F(\vec{x}_{k-1}) \right|}{\max\left\{ \left| F(\vec{x}_{k-1}) \right|, 10^{-10} \right\}} \leq Tol_2$

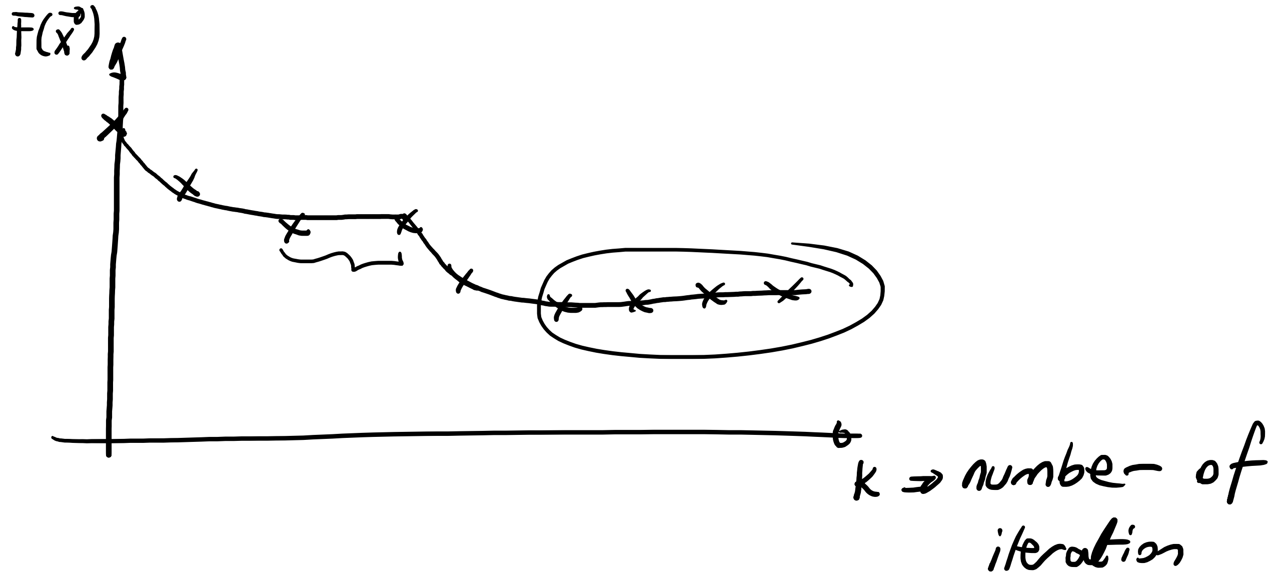
$10^{-10} \rightarrow$ very small value $\neq 0$

② Using Design Variable Vector:

Absolute $\Rightarrow$ $\| \vec{x}_k - \vec{x}_{k-1} \|_2 \leq Tol_3$

$\| \dots \|_2 \Rightarrow L_2$ norm

Relative $\Rightarrow$ $\dfrac{\| \vec{x}_k - \vec{x}_{k-1} \|_2}{\| \vec{x}_{k-1} \|_2} \leq Tol_4$



$$|F(\vec{x}_k) - F(\vec{x}_{k-1})| < Tol$$

make sure to satisfy the convergence criteria for two or three iterations