MSIE 4.0

Co-funded by the
Erasmus+ Programme
of the European Union

Advanced Optimization:
Techniques and Industrial Applications

Prince of Songkla University

Universidade do Minho

Curriculum Development
of Master's Degree Program in
Industrial Engineering for Thailand Sustainable Smart Industry

Co-funded by the Erasmus+ Programme of the European Union

Session 1.4: Non-linear Optimization

Curriculum Development of Master's Degree Program in Industrial Engineering for Thailand Sustainable Smart Industry

Optimality Conditions for Unconstrained Optimization

Consider an unconstrained minimization problem:

$$\min_{x \in R^n} f(\mathbf{x})$$

Then,

1. $\mathbf{x}^*$ is a *local* minimum, if there exists $\varepsilon > 0$ such that

$$f(\mathbf{x}) \geq f(\mathbf{x}^*) \text{ for any } \mathbf{x} \in N_\varepsilon(\mathbf{x}^*) = \left\{ \mathbf{x} \in R^n \middle| \sqrt{\sum_{i=1}^{n} (x_i - x_i^*)^2} < \varepsilon \right\}$$

   i.e., $\mathbf{x}^*$ is the best among its neighborhood.

2. $\mathbf{x}^*$ is a *strict* local minimum if $f(\mathbf{x}) > f(\mathbf{x}^*)$ $\forall \mathbf{x} \neq \mathbf{x}^*$ in the neighborhood of $\mathbf{x}^*$

3. $\nabla f(\mathbf{x})$ and $\nabla^2 f(\mathbf{x})$ are called the first order and the second order conditions

Co-funded by the Erasmus+ Programme of the European Union

Optimality Conditions for Unconstrained Optimization

Theoretical Review

Definition 1:

A square matrix **A** is said to be *positive definite* (or *negative definite*) if the quadratic form $\mathbf{x}^T\mathbf{A}\mathbf{x} = \sum_i \sum_j a_{ij}x_i x_j > 0$ (or < 0) for all $\mathbf{x} \neq \mathbf{0} \in R^n$.

Definition 2:

A square matrix A is said to be *positive semidefinite* (or *negative semidefinite*) if the quadratic form $\mathbf{x}^T\mathbf{A}\mathbf{x} = \sum_i \sum_j a_{ij}x_i x_j > 0$ (or ≤ 0) for all $\mathbf{x} \neq \mathbf{0} \in R^n$.

Definition 3:

$\lambda \in R$ is an eigen value of a square matrix **A** if $\exists \mathbf{x} \neq \mathbf{0} \in R^n$ such that $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$

Co-funded by the
Erasmus+ Programme
of the European Union

Optimality Conditions for Unconstrained Optimization

Properties:

1. λ is an *eigen value* of a square matrix **A** if and only if λ is a root of the following characteristic function of **A**: $\varphi_A(x) = det(\mathbf{A} - x\mathbf{I})$, which is a polynomial of degree *n*.
2. A square matrix **A** is *positive (negative) definite* if all its eigen values are *positive (negative).*
3. A square matrix **A** is *positive (negative) semidefinite* if all its eigen values are *nonnegative (nonpositive)*.

Co-funded by the Erasmus+ Programme of the European Union

Optimality Conditions for Unconstrained Optimization

4. Consider square matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

Define:

$$A_1 = |a_{11}| \quad A_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \quad \ldots \quad A_n = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

Co-funded by the Erasmus+ Programme of the European Union

Optimality Conditions for Unconstrained Optimization

- The matrix **A** will be *positive definite* if and only if all the values $A_1, A_2, \ldots, A_n$ are positive

- The matrix A will be *negative definite* if and only if the sign of A_j is $(-1)^j$ for $j = 1,2, \ldots, n$

- If some of the A_j are positive and the remaining are zero, the matrix A will be *positive semidefinite*

Co-funded by the Erasmus+ Programme of the European Union

Optimality Conditions for Unconstrained Optimization

Theorem: Second Order Necessary Conditions for Optimality

If $\mathbf{x}^*$ is a local minimum, then $\nabla f(\mathbf{x}^*) = \mathbf{0}$, and $\nabla^2 f(\mathbf{x}^*)$ is positive semidefinite.

$\nabla^2 f(\mathbf{x}) = \left\{ \frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} \middle| i, j = 1,2, \dots, n \right\}$ is called the Hessian matrix of $f(\mathbf{x})$

Co-funded by the Erasmus+ Programme of the European Union

Optimality Conditions for Unconstrained Optimization

Proof:
We have: $f(\mathbf{x}^* + d\mathbf{x}) - f(\mathbf{x}^*) \geq 0$.
1. By using Taylor's series expansion:

$$f\left(\mathbf{x}^* + d\mathbf{x}\right) = f\left(\mathbf{x}^*\right) + \nabla f\left(\mathbf{x}^*\right)^{\mathrm{T}} d\mathbf{x} + \frac{1}{2!}(d\mathbf{x})^{\mathrm{T}} \nabla^2 f\left(\mathbf{x}^*\right) d\mathbf{x} + \ldots$$

$$= f\left(\mathbf{x}^*\right) + \nabla f\left(\mathbf{x}^*\right)^{\mathrm{T}} d\mathbf{x} + O(.)$$

$$\Rightarrow f\left(\mathbf{x}^* + d\mathbf{x}\right) - f\left(\mathbf{x}^*\right) = \nabla f\left(\mathbf{x}^*\right)^{\mathrm{T}} d\mathbf{x} + O(.)$$

$$\Rightarrow \nabla f\left(\mathbf{x}^*\right)^{\mathrm{T}} d\mathbf{x} + O(.) \geq 0 \qquad (*)$$

Note that $d\mathbf{x}$ might be positive ($> \mathbf{0}$) or negative ($< \mathbf{0}$). Hence, (*) holds true if and only if $\nabla f(\mathbf{x}^*)^T d\mathbf{x} = 0$ or equivalently, $\nabla f(\mathbf{x}^*) = \mathbf{0}$.

Co-funded by the
Erasmus+ Programme
of the European Union

Optimality Conditions for Unconstrained Optimization

2. We also have:

$$f\left(\mathbf{x}^* + d\mathbf{x}\right) - f\left(\mathbf{x}^*\right) = \nabla f\left(\mathbf{x}^*\right)^{\mathrm{T}} d\mathbf{x} + \frac{1}{2!}\left(d\mathbf{x}\right)^{\mathrm{T}} \nabla^2 f\left(\mathbf{x}^*\right) d\mathbf{x} + \ldots$$

$$= \nabla f\left(\mathbf{x}^*\right)^{\mathrm{T}} d\mathbf{x} + \frac{1}{2!}\left(d\mathbf{x}\right)^{\mathrm{T}} \nabla^2 f\left(\mathbf{x}^*\right) d\mathbf{x} + O(.)$$

$$= \frac{1}{2!}\left(d\mathbf{x}\right)^{\mathrm{T}} \nabla^2 f\left(\mathbf{x}^*\right) d\mathbf{x} + O(.)$$

Hence, $(d\mathbf{x})^T \nabla^2 f(\mathbf{x}^*) d\mathbf{x} \geq 0 \Rightarrow \nabla^2 f(\mathbf{x}^*)$ is positive semidefinite.

Co-funded by the
Erasmus+ Programme
of the European Union

Optimality Conditions for Unconstrained Optimization

Note:

"$\nabla f(\mathbf{x}^*) = \mathbf{0}$ and $\nabla^2 f(\mathbf{x}^*)$ is positive semidefinite" does *not necessarily imply* that $\mathbf{x}^*$ is a local minimum. For example, consider $f(x) = x^3$ and $x^* = 0$. It can be checked that x^* is not a local minimum although $f'(x^*) = 3{x^*}^2 = 0$ and $f''(x^*) = 6x^* = 0$ is positive semidefinite.

Co-funded by the Erasmus+ Programme of the European Union

Optimality Conditions for Unconstrained Optimization

Theorem: Second Order Sufficient Conditions for Optimality

If satisfies $\nabla f(\mathbf{x}^*) = \mathbf{0}$ and $\nabla^2 f(\mathbf{x}^*)$ is positive definite, then $\mathbf{x}^*$ is a strict local minimum.

Proof:

$$f(\mathbf{x}^* + d\mathbf{x}) - f(\mathbf{x}^*) = \nabla f(\mathbf{x}^*)^T + \frac{1}{2!}(d\mathbf{x})^T \nabla^2 f(\mathbf{x}^*) d\mathbf{x} + O(.)$$

Due to $\nabla^2 f(\mathbf{x}^*)$ is positive definite: $(d\mathbf{x})^T \nabla^2 f(\mathbf{x}^*) d\mathbf{x} > 0$
Hence, $f(\mathbf{x}^* + d\mathbf{x}) - f(\mathbf{x}^*) > 0$. This means that $\mathbf{x}^*$ is a strict local minimum.

Co-funded by the Erasmus+ Programme of the European Union

Optimality Conditions for Unconstrained Optimization

Note:

"$\mathbf{x}^*$ is a strict local minimum" does not imply that "$\nabla f(\mathbf{x}^*) = \mathbf{0}$ and $\nabla^2 f(\mathbf{x}^*)$ is positive definite". For example, consider $f(x) = x_1^4 + x_2^4$ and $\mathbf{x}^* = \mathbf{0}$. It is easily to prove that $\mathbf{x}^* = \mathbf{0}$ is a strict local minimum (actually, a strict global minimum). However,

$$\nabla f(\mathbf{x}^*) = \begin{bmatrix} 4{x_1^*}^3 \\ 2x_2^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

But $\nabla^2 f(\mathbf{x}^*) = \begin{bmatrix} 12{x_1^*}^3 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$ is not positive definite!

Co-funded by the Erasmus+ Programme of the European Union

Optimality Conditions for Unconstrained Optimization
Convex Function and Global Optimality

Definition 4: $f(.)$ is convex over convex set S if for any $\mathbf{x}, \mathbf{y} \in S$

$$f(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) \leq \lambda f(\mathbf{x}) + (1-\lambda)f(\mathbf{y}) \quad \forall\lambda \in (0,1)$$

Theorem: $f(.)$ is convex over convex set S if and only if

$$f(\mathbf{y}) \geq f(\mathbf{y}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}) \qquad \forall\mathbf{x}, \mathbf{y} \in S$$

Co-funded by the
Erasmus+ Programme
of the European Union

Optimality Conditions for Unconstrained Optimization
Convex Function and Global Optimality

Proof:

1. Assume that $f(.)$ is convex over S. We have

$$f\big((1-\lambda)\mathbf{x}+\lambda\mathbf{y}\big) \le (1-\lambda)f(\mathbf{x})+\lambda f(\mathbf{y})$$

$$\Rightarrow \quad f\big(\mathbf{x}+\lambda(\mathbf{y}-\mathbf{x})\big)-f(\mathbf{x}) \le \lambda\big(f(\mathbf{y})-f(\mathbf{x})\big)$$

$$\Rightarrow \quad \nabla f(\mathbf{x})^T(\mathbf{y}-\mathbf{x}) = \lim_{\lambda\to 0}\frac{f\big(\mathbf{x}+\lambda(\mathbf{y}-\mathbf{x})\big)-f(\mathbf{x})}{\lambda} \le f(\mathbf{y})-f(\mathbf{x})$$

(Note that: $\lim_{\lambda\to 0}\frac{f(\mathbf{x}+\lambda\mathbf{d})-f(\mathbf{x})}{\lambda} = \nabla f(\mathbf{x})^T\mathbf{d}$)

Hence, $\quad f(\mathbf{y}) \ge f(\mathbf{y})+\nabla f(\mathbf{x})^T(\mathbf{y}-\mathbf{x})$

Co-funded by the Erasmus+ Programme of the European Union

Optimality Conditions for Unconstrained Optimization
Convex Function and Global Optimality

2. Assume that $f(\mathbf{y}) \geq f(\mathbf{y}) + \nabla f(\mathbf{x})^T(\mathbf{y}-\mathbf{x}) \quad \forall \mathbf{x}, \mathbf{y} \in S$

Let $\mathbf{z} = \lambda\mathbf{x} + (1-\lambda)\mathbf{y}$. By the assumption, we have

$$f(\mathbf{x}) \geq f(\mathbf{z}) + \nabla f(\mathbf{z})^T(\mathbf{x}-\mathbf{z}) \text{ and } f(\mathbf{y}) \geq f(\mathbf{z}) + \nabla f(\mathbf{z})^T(\mathbf{y}-\mathbf{z})$$

Multiply the two inequalities by λ and $(1-\lambda)$, respectively, and summing them up, we have

$$\lambda f(\mathbf{x}) + (1-\lambda)f(\mathbf{y}) \geq f(\mathbf{z}) + \nabla f(\mathbf{z})^T[\lambda(\mathbf{x}-\mathbf{z}) + (1-\lambda)(\mathbf{y}-\mathbf{z})]$$

or

$$\lambda f(\mathbf{x}) + (1-\lambda)f(\mathbf{y}) \geq f(\mathbf{z}) = f(\lambda\mathbf{x} + (1-\lambda)\mathbf{y})$$

Hence, $f(.)$ is convex.

Co-funded by the Erasmus+ Programme of the European Union

Optimality Conditions for Unconstrained Optimization
Convex Function and Global Optimality

Theorem: Global Optimality for Convex Function

For a convex function $f(\mathbf{x})$, if $\mathbf{x}^*$ satisfies $\nabla f(\mathbf{x}^*) = \mathbf{0}$, then $\mathbf{x}^*$ is a global minimum. Conversely, if $\mathbf{x}^*$ is a global minimum, then $\nabla f(\mathbf{x}^*) = \mathbf{0}$.

Proof:

Consider $\forall \mathbf{x} \in S$ (S : convex set), assume that $f(.)$ is a differentiable convex function, we have

$$f(\mathbf{x}) - f(\mathbf{x}^*) \geq f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^T(\mathbf{x} - \mathbf{x}^*)$$

If $\nabla f(\mathbf{x}^*) = \mathbf{0}$ then $f(\mathbf{x}) - f(\mathbf{x}^*) \geq \mathbf{0}$, i.e., $\mathbf{x}^*$ is a global minimum

Co-funded by the Erasmus+ Programme of the European Union

Optimality Conditions for Constrained Optimization

Consider the problem

(P) Min $z = f(\mathbf{x})$

s.t. $g_i(\mathbf{x}) \geq 0 \qquad i = 1,2,\dots,m$

The Lagrangian function of the problem is defined as:

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x})$$

In which λ_i's ($i = 1,2,\dots,m$) are Lagrange multipliers.

Co-funded by the Erasmus+ Programme of the European Union

Optimality Conditions for Constrained Optimization
The Karush-Kuhn-Tucker (KKT) Necessary Condition

If $\mathbf{x}^*$ is a local minimum of (**P**), then there should exist $\boldsymbol{\lambda}^*$ such that $\mathbf{x}^*$ and $\boldsymbol{\lambda}^*$ are the solutions of:

$$\lambda_i^* \geq 0 \qquad i = 1,2,\ldots,m$$

$$\nabla f(\mathbf{x}^*) - \sum_{i=1}^{m} \lambda_i^* \nabla g_i(\mathbf{x}^*) = \mathbf{0}$$

$$\lambda_i^* g_i(\mathbf{x}^*) = 0 \qquad i = 1,2,\ldots,m$$

$$g_i(\mathbf{x}) \geq 0 \qquad i = 1,2,\ldots,m$$

Co-funded by the Erasmus+ Programme of the European Union

Optimality Conditions for Constrained Optimization
The Karush-Kuhn-Tucker (KKT) Necessary Condition

Notes:

- If the constraint is $g_i(\mathbf{x}) \geq 0$ then the condition of λ_i^* is: $\lambda_i^* \leq 0$
- If the constraint is $g_i(\mathbf{x}) \geq 0$ then λ_i^* is unrestricted in sign

The KKT condition is also sufficient if program (**P**) is a convex program, i.e., if $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ satisfies the KKT conditions then $\mathbf{x}^*$ is a *global minimum* for (**P**)

Co-funded by the Erasmus+ Programme of the European Union

Optimality Conditions for Constrained Optimization

Convex Program

Consider the program:

$$
\begin{array}{llll}
\textbf{(P)} & \text{Min} & z = f(\mathbf{x}) & \\
 & \text{s.t.} & g_i(\mathbf{x}) \leq 0 & i = 1,2,\dots,r \\
 & & g_i(\mathbf{x}) \geq 0 & i = r+1,\dots,p \\
 & & g_i(\mathbf{x}) = 0 & i = p+1,\dots,m
\end{array}
$$

The above program is a convex program if

- $f(\mathbf{x})$ is a convex function
- $g_i(\mathbf{x})$ $(i = 1,2,\dots,r)$ are convex functions
- $g_i(\mathbf{x})$ $(i = r+1,\dots,p)$ are concave functions
- $g_i(\mathbf{x})$ $(i = p+1,\dots,m)$ are linear functions

Co-funded by the Erasmus+ Programme of the European Union

Optimality Conditions for Constrained Optimization

Example: Consider the problem

$$
\text{(P)} \quad \text{Min} \quad f(\mathbf{x}) = x_1^2 + x_2^2 + x_3^2
$$

$$
\begin{aligned}
\text{s.t.} \quad & g_1(\mathbf{x}) = 2x_1 + x_2 - 5 \le 0 \\
& g_2(\mathbf{x}) = x_1 + x_3 - 2 \le 0 \\
& g_3(\mathbf{x}) = -x_1 + 1 \le 0 \\
& g_4(\mathbf{x}) = -x_2 + 2 \le 0 \\
& g_5(\mathbf{x}) = -x_3 \le 0
\end{aligned}
$$

The above program is a convex program.

Co-funded by the Erasmus+ Programme of the European Union

Optimality Conditions for Constrained Optimization

The KKT conditions are given as:

$$\lambda_1,\lambda_2,\lambda_3,\lambda_4,\lambda_5 \le 0 \qquad \begin{bmatrix}2x_1 & 2x_2 & 2x_3\end{bmatrix} - \begin{bmatrix}\lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 & \lambda_5\end{bmatrix}\begin{bmatrix}2 & 1 & 0\\ 1 & 0 & 1\\ -1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & -1\end{bmatrix} = \mathbf{0}$$

$$\begin{aligned}\lambda_1(2x_1+x_2-5) &= 0 \\ \lambda_2(x_1+x_3-2) &= 0 \\ \lambda_3(-x_1+1) &= 0 \\ \lambda_4(-x_2+2) &= 0 \\ \lambda_5 x_3 &= 0\end{aligned} \qquad \begin{aligned}2x_1+x_2 &\le 5 \\ x_1+x_3 &\le 2 \\ x_1 \ge 1; x_2 \ge 2&; x_3 \ge 0\end{aligned}$$

The solution (global minimum) is: $x_1 = 1, x_2 = 2,\ x_3 = 0\ (\lambda_1 = \lambda_2 = \lambda_5 = 0,\ \lambda_3 = -2,\ \lambda_4 = -4)$

Co-funded by the
Erasmus+ Programme
of the European Union