Lecture – 23

Static Optimization: An Overview

Dr. Radhakant Padhi
Asst. Professor
Dept. of Aerospace Engineering
Indian Institute of Science - Bangalore

Topics

- Unconstrained optimization
- Constrained optimization with equality constraints
- Constrained optimization with inequality constraints
- Numerical examples

Unconstrained Optimization

Dr. Radhakant Padhi
Asst. Professor
Dept. of Aerospace Engineering
Indian Institute of Science - Bangalore

Static Optimization

Observation for: $J_1(x)$

Point 1: Local maximum

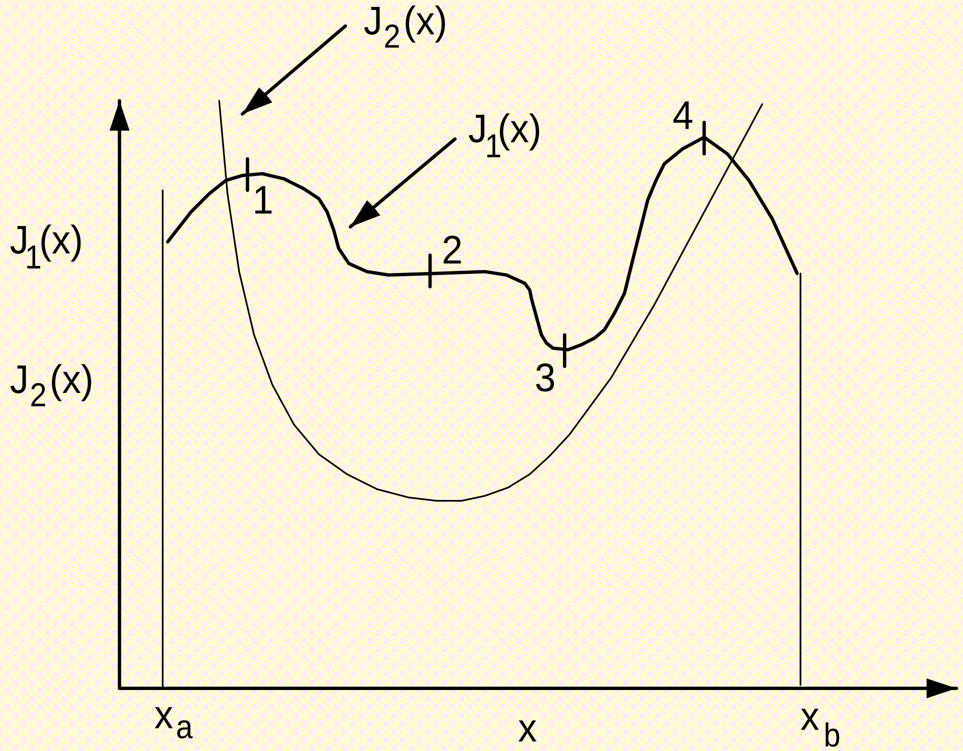
Point 2: Point of inflexion

Point 3: Local minimum

Point 4: Local maximum

Point 3: Global minimum

Point 4: Global maximum



At all minima/maxima: $$\frac{dJ(x)}{dx} = 0$$

Necessary and Sufficient Conditions for Optimality

Scalar Case:

Performance Index $J(x)$: An analytic function x of

Taylor series:

$$\left[J\left(x^* + \Delta x\right) - J\left(x^*\right)\right] = \left.\frac{dJ}{dx}\right|_{x=x^*} \Delta x + \frac{1}{2!}\left.\frac{d^2 J}{dx^2}\right|_{x=x^*} \left(\Delta x\right)^2 + \cdots$$

Necessary Condition:

If $J(x^*)$ is a minimum irrespective of the sign of Δx,

then

$$\left.\frac{dJ}{dx}\right|_{x=x^*} = 0$$

Necessary and Sufficient Conditions for Optimality

Sufficient Condition:

$$\left[J\left(x^* + \Delta x\right) - J\left(x^*\right) \right] = \frac{1}{2!} \left. \frac{d^2 J}{dx^2} \right|_{x=x^*} \left(\Delta x\right)^2 + \text{HOT}$$

$$\left[J\left(x^* + \Delta x\right) > J\left(x^*\right) \right], \quad \text{irrespective of the sign of } \Delta x$$

if $\left. \dfrac{d^2 J}{dx^2} \right|_{x=x^*} > 0$ **(sufficiency condition for local minimum)**

Similarly, if $\left. \dfrac{d^2 J}{dx^2} \right|_{x=x^*} < 0,$ it leads to a local maximum

Necessary and Sufficient Conditions for Optimality

Q-1: What if $\left.\frac{dJ}{dx}\right|_{x=x^*} = \left.\frac{d^2 J}{dx^2}\right|_{x=x^*} = 0 \;\; ?$

Answer:

$$J\left(x^* + \Delta x\right) - J\left(x^*\right) = \frac{1}{3!}\left.\frac{d^3 J}{dx^3}\right|_{x=x^*} \left(\Delta x\right)^3 + \frac{1}{4!}\left.\frac{d^4 J}{dx^4}\right|_{x=x^*} \left(\Delta x\right)^4 + \cdots$$

Necessary condition $\left.\frac{d^3 J}{dx^3}\right|_{x=x^*} = 0$

Sufficient condition $\left.\frac{d^4 J}{dx^4}\right|_{x=x^*} > 0$ (for minimization)

Necessary and Sufficient Conditions for Optimality

Q-2: **What if** $\left.\dfrac{dJ}{dx}\right|_{x=x^*} = \left.\dfrac{d^2J}{dx^2}\right|_{x=x^*} = 0$ but $\left.\dfrac{d^3J}{dx^3}\right|_{x=x^*} \neq 0$?

Then $x = x^*$ is a point of inflexion

Example – 1: $J = x^4$

$$dJ/dx = 4x^3 = 0$$

$$x^* = 0, 0, 0$$

$$\left.\frac{d^2J}{dx^2}\right|_{x^*=0} = 12x^{*2} = 0, \quad \left.\frac{d^3J}{dx^3}\right|_{x^*=0} = 24x^* = 0, \quad \left.\frac{d^4J}{dx^4}\right|_{x^*=0} = 24 > 0$$

minimum

Necessary and Sufficient Conditions for Optimality

Example – 2: $J = x^3$

$$dJ/dx = 3x^2 = 0$$

$$\Rightarrow \quad x^* = 0, 0$$

$$\left.\frac{d^2J}{dx^2}\right|_{x^*=0} = 6x^* = 0\,, \qquad \left.\frac{d^3J}{dx^3}\right|_{x^*=0} = 6 \neq 0$$

Hence, x^* is a point of inflexion.

Necessary and Sufficient Conditions for Optimality

Vector case

Minimize $J(X) \in \mathbb{R}$ where $X \in \mathbb{R}^n$

By definition,

$$\frac{\partial J}{\partial X} \triangleq \begin{bmatrix} \dfrac{\partial J}{\partial x_1} \\ \vdots \\ \dfrac{\partial J}{\partial x_n} \end{bmatrix} \qquad \frac{\partial^2 J}{\partial X^2} \triangleq \begin{bmatrix} \dfrac{\partial^2 J}{\partial x_1^2} & \cdots & \dfrac{\partial^2 J}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 J}{\partial x_n \partial x_1} & \cdots & \dfrac{\partial^2 J}{\partial x_n^2} \end{bmatrix}$$

Necessary and Sufficient Conditions for Optimality

$$J(X) = J(X^* + \Delta X)$$

$$= J\left(X^*\right) + \left(\left.\frac{\partial J}{\partial X}\right|_{X^*}\right)\Delta X + \frac{1}{2!}(\Delta X)^T\left(\left.\frac{\partial^2 J}{\partial X^2}\right|_{X^*}\right)\Delta X + \ldots$$

For minimization,

$$J\left(X^* + \Delta X\right) - J\left(X^*\right) > 0 \quad \left(\text{irrespective of sign of } \Delta X\right)$$

Necessary Condition: $\left.\left[\frac{\partial J}{\partial X}\right]\right|_{X^*} = 0$

Sufficient Condition: $\left.\left[\frac{\partial^2 J}{\partial X^2}\right]\right|_{X^*} > 0$ (positive definite)

Remark: Further Conditions are difficult to use in practice!

Necessary and Sufficient Conditions for Optimality

Example – 1: $$J(X) = \frac{1}{2}\left(x_1^2 + x_2^2\right)$$

Necessary Condition $$\left[\frac{\partial J}{\partial X}\right]\Bigg|_{X^*} = 0$$

$$\begin{bmatrix} \dfrac{\partial J}{\partial x_1} \\ \dfrac{\partial J}{\partial x_2} \end{bmatrix}_{X^*} = \begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Necessary and Sufficient Conditions for Optimality

Sufficient Condition:

$$\left[\frac{\partial^2 J}{\partial X^2}\right]_{X^*} = \begin{bmatrix} \dfrac{\partial^2 J}{\partial x_1^2} & \dfrac{\partial^2 J}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 J}{\partial x_2 \partial x_1} & \dfrac{\partial^2 J}{\partial x_2^2} \end{bmatrix}_{X^*} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Eigenvalues: 1,1 at $X = X^*$

$\left[\frac{\partial^2 J}{\partial X^2}\right]\Bigg|_{X^*} > 0$ (positive definite). So $X^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is a minimum point

Necessary and Sufficient Conditions for Optimality

Example – 2: $J(x) = \frac{1}{2}\left(x_1^2 - x_2^2\right)$

Solution:

$$\frac{\partial J}{\partial X} = 0 \quad \Rightarrow \quad X^* = \begin{bmatrix} x_1^* \\ -x_2^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\frac{\partial^2 J}{\partial X^2} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad \text{Eigenvalues: } \; 1, -1$$

$$i.e. \; \frac{\partial^2 J}{\partial X^2} \text{ is neither positive definite, nor negetive definite}$$

Hence $X = 0$ is a 'saddle point'.

Constrained Optimization with Equality Constraints

Dr. Radhakant Padhi

Asst. Professor

Dept. of Aerospace Engineering

Indian Institute of Science - Bangalore

Constrained Optimization: Equality Constraint

Problem: Minimize $J(X) \in \mathbb{R} \quad (X \in \mathbb{R}^n)$

Subject to $f(X) = 0$

where, $f(X) = \begin{bmatrix} f_1(X) & \cdots & f_m(X) \end{bmatrix}^T \in \mathbb{R}^m$

Solution Procedure:

Formulate an augmented cost function

$$\bar{J}(X, \lambda) \triangleq J(X) + \lambda^T f(X)$$

Constrained Optimization: Equality Constraint

Necessary Conditions:

$$\frac{\partial \bar{J}}{\partial X} = \frac{\partial J}{\partial X} + \left[\frac{\partial f}{\partial X}\right]^T \lambda = 0 \qquad \Leftarrow n \;\; \text{equations}$$

$$\frac{\partial \bar{J}}{\partial \lambda} = f(X) = 0 \qquad \Leftarrow m \;\; \text{equations}$$

Hence, it lead to $(n+m)$ equations

with $(n+m)$ variables. Solve it!

Constrained Optimization with Equality Constraint: An Example

Minimize $J(X) = \frac{1}{2}\left(x_1^2 + x_2^2\right)$

Subjected to: $f(X) = x_1 + x_2 - 2 = 0$

Solution: $\bar{J}(X) = \frac{1}{2}\left(x_1^2 + x_2^2\right) + \lambda\left(x_1 + x_2 - 2\right)$

$$\begin{bmatrix} \partial \bar{J} / \partial x_1 \\ \partial \bar{J} / \partial x_2 \\ \partial \bar{J} / \partial \lambda \end{bmatrix} = \begin{bmatrix} x_1^* + \lambda^* \\ x_2^* + \lambda^* \\ x_1^* + x_2^* - 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} -\lambda^* \\ -\lambda^* \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Constrained Optimization with Equality Constraint: Another Example

Minimize $J(X) = \frac{1}{2}\left[\left(\frac{x_1}{a}\right)^2 + \left(\frac{x_2}{b}\right)^2\right]$

Subject to $x_1 + m x_2 - c = 0$

where a, b, m, c are Constants

Solution:

$$\bar{J} = \frac{1}{2}\left[\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2}\right] + \lambda\left(x_1 + m x_2 - c\right)$$

Constrained Optimization with Equality Constraint: Another Example

Solve:

$$\begin{bmatrix} \dfrac{\partial \bar{J}}{\partial x_1} \\ \dfrac{\partial \bar{J}}{\partial x_2} \\ \dfrac{\partial \bar{J}}{\partial \lambda} \end{bmatrix}_{\left(x_1^*, x_2^*, \lambda^*\right)} = \begin{bmatrix} \dfrac{x_1^*}{a^2} + \lambda^* \\ \dfrac{x_2^*}{b^2} + \lambda^* \\ x_1^* + m x_2^* - c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$x_1^* = \left(\frac{a^2 c}{a^2 + m^2 b^2} \right), x_2^* = \left(\frac{b^2 m c}{a^2 + m^2 b^2} \right), \lambda^* = \left(\frac{-c}{a^2 + m^2 b^2} \right)$$

Remark: λ^* *has no physical meaning. It only helps to solve the problem.*

Constrained Optimization with Equality Constraint: Sufficiency Condition

If the equation

$$\left| \begin{bmatrix} \left[\dfrac{\partial^2 \bar{J}}{\partial X^2} - I\sigma \right] & \left[\dfrac{\partial f}{\partial X} \right]^T \\ \left[\dfrac{\partial f}{\partial X} \right] & 0 \end{bmatrix} \right| = 0$$

has only positive roots $\sigma_i \Rightarrow$ Minimum

has only negative roots $\sigma_i \Rightarrow$ Maximum

Example – 1

Problem: $J = \frac{1}{2}\left(x^2{}_1 + x_2^2\right), \quad f(x_1, x_2) = x_1 - x_2 - 5 = 0$

Solution: $\bar{J} = \frac{1}{2}\left(x^2{}_1 + x_2^2\right) + \lambda\left(x_1 - x_2 - 5\right)$

Necessary condition: $\begin{bmatrix} x_1 + \lambda \\ x_2 - \lambda \\ x_1 - x_2 - 5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

$$\lambda = -\tfrac{5}{2}, \; x_1 = \tfrac{5}{2}, \; x_2 = -\tfrac{5}{2}$$

Example – 1

Sufficient condition:

$$\frac{\partial^2 \bar{J}}{\partial X^2} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \frac{\partial f}{\partial X} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\det\left(\begin{bmatrix} 1-\sigma & 0 & 1 \\ 0 & 1-\sigma & -1 \\ 1 & -1 & 0 \end{bmatrix}\right) = 0 \quad \Rightarrow \quad \boxed{\sigma = 1 > 0}$$

The Solution $x_1 = \frac{5}{2},\ x_2 = -\frac{5}{2}$ is a minimum.

Example – 1: Some Remarks

In this example, x_1, x_2 and λ do not appear in the equation for σ. Moreover, the solution is the 'only solution'. Hence, the result is **'global'.**

In general, however σ will be a function of X and λ. Hence, various conclusions have to be derived on case-to-case basis.

Example – 2

Problem: $J = x_1 - x_2^2 \,, \quad f(X) = x_1^2 + x_2^2 - 1 = 0$

Solution: $\bar{J} = x_1 - x_2^2 + \lambda\left(x_1^2 + x_2^2 - 1\right)$

Necessary condition:

$$\begin{bmatrix} 1 + 2\lambda x_1 \\ 2x_2(\lambda - 1) \\ x_1^2 + x_2^2 - 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Sufficient condition:

$$\frac{\partial^2 \bar{J}}{\partial X^2} = \begin{bmatrix} 2\lambda & 0 \\ 0 & 2(\lambda - 1) \end{bmatrix}, \frac{\partial f}{\partial X} = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}$$

Example – 2

$$\det\left(\begin{bmatrix} 2\lambda-\sigma & 0 & 2x_1 \\ 0 & 2\lambda-2-\sigma & 2x_2 \\ 2x_1 & 2x_2 & 0 \end{bmatrix}\right)=0$$

x_1	x_2	λ	σ	Conclusion
1	0	-1/2	- 3	Maximum
-1	0	1/2	-1	Maximum
-1/2	1.73	1	3/2	Minimum
-1/2	-1.73	1	3/2	Minimum

Constrained Optimization with Inequality Constraints

Dr. Radhakant Padhi
Asst. Professor
Dept. of Aerospace Engineering
Indian Institute of Science - Bangalore

Constrained Optimization with Inequality Constraints: A naïve approach

Remark: ***One way of dealing with inequality constraints for the variables is as follows:***

Let $x_{i_{\min}} \le x_i \le x_{i_{\max}}$ ***(Important for control problems)***

Replace: $x_i = x_{i_{\min}} + \left(x_{i_{\max}} - x_{i_{\min}} \right) \sin^2 \alpha_i$

Consider α_i ***as a free variable.***

Optimization with Inequality Constraints

Problem: Maximize / Minimize: $J(X) \in \mathbb{R},\ X \in \mathbb{R}^n$

Subject to: $$g(X) \triangleq \begin{bmatrix} g_1(X) \\ \vdots \\ g_m(X) \end{bmatrix} \le \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

Solution: First, introduce "slack variables" $\mu_1, \ldots, \mu_m$ to convert ineqaulity constraints to equality constraints as follows:

$$f_g(X,\mu) \triangleq \begin{bmatrix} g_1(X) + \mu_1^2 \\ \vdots \\ g_m(X) + \mu_m^2 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

Then follow the routine procedure for the equality constraints.

Optimization with Inequality Constraints

Augmented PI: $\bar{J}(X,\lambda,\mu) = J(X) + \sum_{j=1}^{m}\left[\lambda_j g_j(X) + \lambda_j \mu_j^2\right]$

Necessary Conditions:

$$\frac{\partial \bar{J}}{\partial x_i} = \frac{\partial J}{\partial x_i} + \sum_{j=1}^{m}\lambda_j \frac{\partial g_j}{\partial x_i} = 0, \quad i = 1,\ldots,n \qquad (n \text{ equations})$$

$$\frac{\partial \bar{J}}{\partial \lambda_j} = g_j(X) + \mu_j^2 = 0, \qquad j = 1,\ldots,m \qquad (m \text{ equations})$$

$$\frac{\partial \bar{J}}{\partial \mu_j} = 2\lambda_j \mu_j = 0, \qquad j = 1,\ldots,m \qquad (m \text{ equations})$$

Optimization with Inequality Constraints

$$\frac{\partial \bar{J}}{\partial \lambda_j} = g_j\left(X\right) + \mu_j^2 = 0$$

$$g_j\left(X\right) = -\mu_j^2$$

$$\lambda_j g_j = -\mu_j\left(\lambda_j \mu_j\right)$$

$$\frac{\partial \bar{J}}{\partial \mu_j} = 2\lambda_j \mu_j = 0$$

Hence, $\lambda_j g_j = 0$

This leads to the conclusion that either $\lambda_j = 0$ or $g_j = 0$

i.e.

If a constraint is strictly an inequality constraint, then the problem can be solved without considering it. Otherwise, the problem can be solved by considering it as an equality constraint.

Necessary Conditions (Kuhn-Tucker Conditions)

$$\frac{\partial \bar{J}}{\partial x_i} = \frac{\partial J}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i} = 0, \quad i = 1, \ldots, n \qquad (n \text{ equations})$$

$$\lambda_j g_j(X) = 0, \qquad j = 1, \ldots, m \qquad (m \text{ equations})$$

For $J(X)$ to be MINIMUM

$$\text{if } g_j(X) \le 0 \quad \text{then} \quad \lambda_j \ge 0$$

$$\text{if } g_j(X) \ge 0 \quad \text{then} \quad \lambda_j \le 0$$

(opposite sign)

For $J(X)$ to be MAXIMUM

$$\text{if } g_j(X) \le 0 \quad \text{then} \quad \lambda_j \le 0$$

$$\text{if } g_j(X) \ge 0 \quad \text{then} \quad \lambda_j \ge 0$$

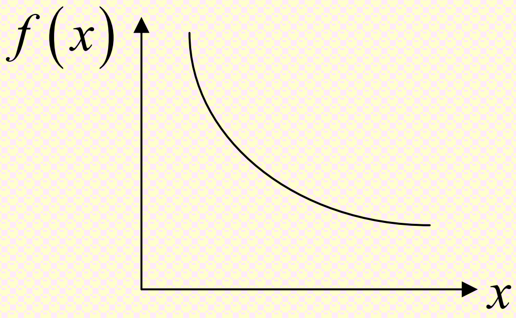
(same sign)

Optimization with Inequality Constraints: Comments

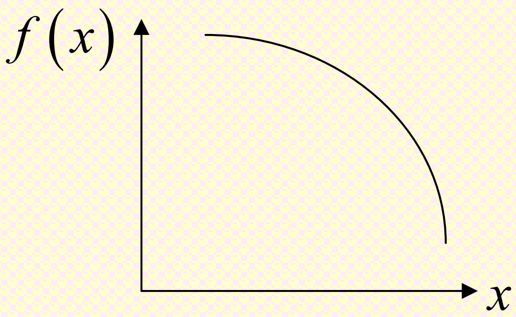
- One should explore all possibilities in the Kuhn-Tucker conditions to arrive at an appropriate conclusion
- Kuhn-Tucker conditions are only “necessary conditions”
- Sufficiency check demands the concept of “convexity”

Convex/Concave Function $f(x)$

- A function is called **convex**, if a straight line drawn between any two points on the surface generated by the function lies completely above or on the surface.
- If the line lies strictly above the surface, then the function is called **strictly convex**.
- If the line lies below the surface, then the function is called a **concave**.



Strictly convex



Strictly concave

Result for Local Convexity/Concavity of $f(X)$ at X^*

Definition	$\left[\dfrac{\partial^2 f}{\partial X^2}\right]_{X^*}$	Eigenvalues
Strictly convex	Positive definite	$\lambda_i > 0, \quad \forall i$
Convex	Positive Semi-definite	$\lambda_i \geq 0, \quad \forall i$
Strictly concave	Negative definite	$\lambda_i < 0, \quad \forall i$
Concave	Negative Semi-definite	$\lambda_i \leq 0, \quad \forall i$
No classification	Indefinite	Some $\lambda_i > 0$. Rest are ≤ 0

Conditions for which Kuhn-Tucker Conditions are also Sufficient

Condition	$J(X)$	All $g_j(X)$
Maximum	Strictly concave	Convex
Minimum	Strictly convex	Convex

Example

Problem: Minimize: $J(X) = \left(x_1^2 + x_2^2\right)$

Subject to: $(x_1 - x_2) \le 5$

$(x_1 - x_2) \ge 1$

Solution: $g_1(X) = (x_1 - x_2 - 5) \le 0$

$g_2(X) = (-x_1 + x_2 + 1) \le 0$

$$\bar{J} = \left(x_1^2 + x_2^2\right) + \lambda_1 (x_1 - x_2 - 5) + \lambda_2 (-x_1 + x_2 + 1)$$

Example: Kuhn-Tucker Conditions

$$\frac{\partial \bar{J}}{\partial x_1} = 2x_1 + \lambda_1 - \lambda_2 = 0$$

$$\frac{\partial \bar{J}}{\partial x_2} = 2x_2 - \lambda_1 + \lambda_2 = 0$$

$$\lambda_1 \left(x_1 - x_2 - 5 \right) = 0$$

$$\lambda_2 \left(-x_1 + x_2 + 1 \right) = 0$$

Note: $x_2 = -x_1$

$$\left(x_1 - x_2 - 5 \right) \le 0$$

$$\left(-x_1 + x_2 + 1 \right) \le 0$$

$$\lambda_1 \ge 0$$

$$\lambda_2 \ge 0$$

All possible solutions should be investigated

Feasible Solution of Kuhn-Tucker Conditions

- Case – 1: $\lambda_1 = 0, \; \lambda_2 \neq 0,$ Feasible: $x_1 = \frac{1}{2}, x_2 = -\frac{1}{2}$
- Case – 2: $\lambda_1 = 0, \; \lambda_2 = 0,$ Not Feasible: $x_1 = x_2 = 0$
- Case – 3: $\lambda_1 \neq 0, \; \lambda_2 = 0,$ Not Feasible: $x_1 = \frac{5}{2}, x_2 = -\frac{5}{2}$
- Case – 4: $\lambda_1 \neq 0, \; \lambda_2 \neq 0,$ Not Feasible: *No Solution*!

Sufficiency condition

$J(X) = (x_1^2 + x_2^2)$ is strictly convex. $g_1(X),\ g_2(X)$ are also convex.

Hence, the Kuhn-Tucker conditions are both Necessary and Sufficient.

Moreover, $\dfrac{\partial^2 J}{\partial X^2} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} > 0$ and it doesnot depend on the value of X.

Hence, $X^* = \begin{bmatrix} 1/2 & -1/2 \end{bmatrix}^T$ is the GLOBAL minimum!

References

- T. F. Elbert, *Estimation and Control Systems*, Von Nostard Reinhold, 1984.

- S. S. Rao, *Optimization Theory and Applications*, Wiley, Second Edition, 1984.

Thanks for the Attention...!