Lecture 5

An Overview of Static Optimization – II

Prof. Radhakant Padhi
Dept. of Aerospace Engineering
Indian Institute of Science - Bangalore

Topics

• Constrained optimization with inequality constraints

• Numerical Optimization

• Numerical examples
Constrained Optimization with Inequality Constraints

Prof. Radhakant Padhi
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 \( x_{i_{\text{min}}} \leq x_i \leq x_{i_{\text{max}}} \) (Important for control problems)

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

Consider \( \alpha_i \) as a free variable.

Note: This approach does not work in general.
Optimization with Inequality Constraints

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

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

Solution: First, introduce "slack variables" \( \mu_1, \ldots, \mu_m \) to convert inequality constraints to equality constraints as follows:

\[
\begin{aligned}
& f_s(X, \mu) \triangleq \begin{bmatrix} g_1(X) + \mu_1^2 \\ \vdots \\ g_m(X) + \mu_m^2 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \\
& \lambda + \mu = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_m \\ \mu_1 \\ \vdots \\ \mu_m \end{bmatrix} 
\end{aligned}
\]

Then follow the routine procedure for the equality constraints.

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

Necessary Conditions:

\[
\begin{align*}
\frac{\partial \bar{J}}{\partial x_i} &= \frac{\partial J}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i} = 0, \quad i = 1, \ldots, n \\
\frac{\partial \bar{J}}{\partial \lambda_j} &= g_j(X) + \mu_j^2 = 0, \quad j = 1, \ldots, m \\
\frac{\partial \bar{J}}{\partial \mu_j} &= 2 \lambda_j \mu_j = 0, \quad j = 1, \ldots, m
\end{align*}
\]
**Optimization with Inequality Constraints**

\[ \frac{\partial T}{\partial \lambda_j} = g_j(X) + \mu_j^2 = 0 \]

\[ g_j(X) = -\mu_j^2 \]

\[ \lambda_j g_j = -\mu_j \lambda_j \mu_j \]

But \[ \frac{\partial T}{\partial \mu_j} = 2\lambda_j \mu_j = 0 \]

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.

---

**Optimization with Inequality Constraints (single variable case)**

Let us assume:

(i) a maximization problem

(ii) \( g_j \) is active, i.e. \( g_j = 0 \)

Let

(i) \( X \) is a point where a maximum occurs

(ii) \( \Delta x_i \) (a small change) that causes \( g_j < 0 \)
Optimization with Inequality Constraints (single variable case)

In this case,

(i) \( \frac{dJ}{dx_i} \Delta x_i < 0 \) (since \( J \) is a maximum)

(ii) \( \frac{dg_j}{dx_i} \Delta x_i < 0 \)

Hence,

(i) If \( \Delta x_i > 0 \), then \( \frac{\partial J}{\partial x_i} < 0 \) & \( \frac{\partial g_j}{\partial x_i} < 0 \) (both negative)

(ii) If \( \Delta x_i < 0 \), then \( \frac{\partial J}{\partial x_i} > 0 \) & \( \frac{\partial g_j}{\partial x_i} > 0 \) (both positive)

Optimization with Inequality Constraints (single variable case)

Necessary condition:

\[
\frac{\partial J}{\partial x_i} + \frac{\partial}{\partial x_i} (\lambda_j g_j) = 0
\]

\[
\frac{\partial J}{\partial x_i} = -\frac{\partial}{\partial x_i} (\lambda_j g_j)
\]

\[
\frac{\partial J}{\partial x_i} = -\lambda_j \left( \frac{\partial g_j}{\partial x_i} \right)
\]

But \( \frac{\partial J}{\partial x_i} \) & \( \frac{\partial g_j}{\partial x_i} \) are either both positive or both negative

Hence, \( \lambda_j < 0 \) for maximization!
Necessary Conditions: Karush-Kuhn-Tucker (KKT) Conditions

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

\[ \lambda_j g_j(X) = 0, \quad j = 1, \ldots, m \quad (m \text{ equations}) \]

For \( J(X) \) to be MINIMUM
\[
\begin{align*}
\text{if } g_j(X) &\leq 0 \quad \text{then } \lambda_j \geq 0 \\
\text{if } g_j(X) &\geq 0 \quad \text{then } \lambda_j \leq 0
\end{align*}
\]
(opposite sign)

For \( J(X) \) to be MAXIMUM
\[
\begin{align*}
\text{if } g_j(X) &\leq 0 \quad \text{then } \lambda_j \leq 0 \\
\text{if } g_j(X) &\geq 0 \quad \text{then } \lambda_j \geq 0
\end{align*}
\]
(same sign)

Comments on Karush-Kuhn-Tucker (KKT) Conditions

- One should explore all possibilities in the Karush-Kuhn-Tucker conditions to arrive at an appropriate conclusion
- Karush-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**.

![Diagram of Convex/Concave Function](image)

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

<table>
<thead>
<tr>
<th>Definition</th>
<th>Eigenvalues</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strictly convex</td>
<td>$\lambda_i &gt; 0, \forall i$</td>
</tr>
<tr>
<td>Convex</td>
<td>$\lambda_i \geq 0, \forall i$</td>
</tr>
<tr>
<td>Strictly concave</td>
<td>$\lambda_i &lt; 0, \forall i$</td>
</tr>
<tr>
<td>Concave</td>
<td>$\lambda_i \leq 0, \forall i$</td>
</tr>
<tr>
<td>No classification</td>
<td>Some $\lambda_i &gt; 0$. Rest are $\leq 0$</td>
</tr>
</tbody>
</table>
Conditions for which Kuhn-Tucker Conditions are also Sufficient

<table>
<thead>
<tr>
<th>Condition</th>
<th>( J(X) )</th>
<th>All ( g_j(X) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum</td>
<td>Strictly concave</td>
<td>Convex</td>
</tr>
<tr>
<td>Minimum</td>
<td>Strictly convex</td>
<td>Convex</td>
</tr>
</tbody>
</table>

Example

Problem: Minimize: \( J(X) = (x_1^2 + x_2^2) \)
Subject to: \( (x_1 - x_2) \leq 5 \)
\( (x_1 - x_2) \geq 1 \)

Solution: \( g_1(X) = (x_1 - x_2 - 5) \leq 0 \)
\( g_2(X) = (-x_1 + x_2 + 1) \leq 0 \)
\( \mathcal{J} = (x_1^2 + x_2^2) + \lambda_1 (x_1 - x_2 - 5) + \lambda_2 (-x_1 + x_2 + 1) \)
Example: Karush-Kuhn-Tucker Conditions

\[
\begin{align*}
\frac{\partial J}{\partial x_1} &= 2x_1 + \lambda_1 - \lambda_2 = 0 \\
\frac{\partial J}{\partial x_2} &= 2x_2 - \lambda_1 + \lambda_2 = 0 \\
\lambda_1 (x_1 - x_2 - 5) &= 0 \\
\lambda_2 (-x_1 + x_2 + 1) &= 0 \\
(x_1 - x_2 - 5) &\leq 0 \\
(-x_1 + x_2 + 1) &\leq 0 \\
\lambda_1 &\geq 0 \\
\lambda_2 &\geq 0
\end{align*}
\]

Note: \( x_2 = -x_1 \)

All possible solutions should be investigated

Feasible Solution of Karush-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, Karush-Kuhn-Tucker conditions are both Necessary and Sufficient.

Moreover, \( \frac{\partial^2 J}{\partial X^2} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \geq 0 \) and it does not depend on the value of \( X \).

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

---

**Numerical Optimization**

**Prof. Radhakant Padhi**

Dept. of Aerospace Engineering

Indian Institute of Science - Bangalore
Basic Philosophy

1. Start with a meaningful initial guess value $X^1$
2. Find a search direction $p^k, \ k = 1, 2, \ldots$
3. Update the guess value $X^{k+1} = X^k + \alpha p^k, \ \alpha > 0$
4. Repeat Steps 2 & 3 until convergence, i.e.
$$\left\| J(X^{k+1}) - J(X^k) \right\| < \text{tol}$$

Unconstrained Optimization: Steepest Descent Search

$$J(X^{k+1}) = J(X^k) + \left[ \nabla J(X^k) \right]^T (X^{k+1} - X^k) + \text{HOT}$$
$$J(X^{k+1}) - J(X^k) \approx \left[ \nabla J(X^k) \right]^T \frac{(X^{k+1} - X^k)}{\alpha p^k}$$
$$= \alpha \left[ \nabla J(X^k) \right]^T p^k$$

Hence, if $p^k = -\nabla J(X^k)$ (steepest descent direction)
$$[J(X^{k+1}) - J(X^k)] = -\alpha \left[ \nabla J(X^k) \right]^T \left[ \nabla J(X^k) \right] \ (\alpha > 0)$$
$$< 0$$
Unconstrained Optimization:
Pictorial Representation

\[ f(x^j) = J(X) \]
\[ f(x^j) = J(X^k) \]
\[ p^j = p^k \]

Note:

- Search along \( \nabla f(x^j) \) until the minimum is obtained
  - Find three guess values of \( p \), such that there is an up-down-up behaviour
  - Fit a quadratic curve (parabola) for these three points
  - Minimum of this quadratic curve is the updated value
- Find a new direction at this point
- Repeat the procedure

Reference: R. D. Robinett III, D. G. Wilson, G. R. Eisler and J. E. Hurtado, 
Optimal Control, Guidance and Estimation

Unconstrained Optimization:
Pictorial Representation of Line Search

Reference: R. D. Robinett III, D. G. Wilson, G. R. Eisler and J. E. Hurtado,

Unconstrained Optimization:
Newton’s Method

\[
\nabla J(X^{k+1}) = \nabla J(X^k) + \left[ \nabla^2 J(X^k) \right] (X^{k+1} - X^k) + \text{HOT}
\]

(at extremum point)

\[
0 = \nabla J(X^k) + \alpha \left[ \nabla^2 J(X^k) \right] p^k
\]

\[
p^k = -\left( \frac{1}{\alpha} \right) \left[ \nabla J(X^k) \right]^{-1} \nabla J(X^k)
\]

\[
p^k = -\beta \left[ \nabla^2 J(X^k) \right]^{-1} \nabla J(X^k), \quad \beta > 0
\]

Advantage: Fast convergence
Drawback: Computation of \( \left[ \nabla^2 J(X^k) \right]^{-1} \) is not trivial and can be computationally intensive
Constrained Optimization: Equality Constraint

**Problem:** Minimize \( J(X) \in \mathbb{R} \left( X \in \mathbb{R}^n \right) \)
Subject to \( f(X) = 0 \)
where, \( f(X) = [f_1(X) \cdots f_m(X)]^T \in \mathbb{R}^m \)

**Solution Procedure:**
Formulate an augmented cost function
\[
\tilde{J}(X, \lambda) \triangleq J(X) + \lambda^T f(X)
\]

Constrained Optimization: Steepest Descent Search

\[
\begin{align*}
\tilde{J}(X^{k+1}, \lambda^{k+1}) & = \tilde{J}(X^k, \lambda^k) + \left( \frac{\partial \tilde{J}}{\partial X} \right)^T_{X^k, \lambda^k} (X^{k+1} - X^k) + \left( \frac{\partial \tilde{J}}{\partial \lambda} \right)^T_{X^k, \lambda^k} (\lambda^{k+1} - \lambda^k) \\
& = \tilde{J}(X^k, \lambda^k) + \nabla J(X^k) + (\lambda^k)^T \nabla f(X^k) (X^{k+1} - X^k) + \left( \frac{f(X^k)}{0} \right)^T (\lambda^{k+1} - \lambda^k)
\end{align*}
\]

\[
\tilde{J}(X^{k+1}, \lambda^{k+1}) - \tilde{J}(X^k, \lambda^k) = \alpha \left( \nabla J(X^k) + (\lambda^k)^T \nabla f(X^k) \right)^T p^k \leq 0
\]

This suggests:
\[
p^k = - \left( \nabla J(X^k) + (\lambda^k)^T \nabla f(X^k) \right)
\]
Constrained Optimization: Pictorial Representation

At optimum point:

\[ f(X) = J(X) \]
\[ f(X^*) = J(X^k) \]
\[ g(X) = f(X) \]
\[ p^* = p^k \]


MATLAB Function: *fmincon*

Problem: Minimize: \[ f(X) \in \mathbb{R} \quad (X \in \mathbb{R}^n) \]

Subject to:

\[ h(X) \leq 0 \in \mathbb{R}^m \]
\[ g(X) = 0 \]
\[ X_l \leq X \leq X_u \]
\[ DX \leq b \]
\[ D_{eq} X = b_{eq} \]
References


Thanks for the Attention....!!

Questions ... ??