18-660: Numerical Methods for Engineering Design and Optimization

Xin Li
Department of ECE
Carnegie Mellon University
Pittsburgh, PA 15213

Electrical & Computer ENGINEERING

Overview

- Constrained Optimization
  - Linear equality constraint
  - Lagrange multiplier

Constrained Nonlinear Optimization

$$\min_{X} \quad f(X)$$
$$\text{S.T.} \quad \begin{cases} g_1(X) \le 0 \\ g_2(X) \le 0 \\ \vdots \end{cases}$$

- Equality constraint can be written as two inequality constraints

$$\begin{aligned} &\min_{X} \quad f(X) \\ &\text{S.T.} \quad g(X) = 0 \end{aligned} \quad \Rightarrow \quad \begin{aligned} &\min_{X} \quad f(X) \\ &\text{S.T.} \quad \begin{cases} g(X) \le 0 \\ -g(X) \le 0 \end{cases} \end{aligned}$$

Linear Equality Constraint

$$\begin{aligned} \min_{X} \quad & f(X) \\ \text{S.T.} \quad & AX = B \end{aligned}$$

- Linear equality constraint can be efficiently handled by a number of optimization algorithms
  - We do not write AX = B as two inequality constraints
  - It can be directly solved with high efficiency

Subspace Reduction

- Eliminate linear equality constraint

$$AX = B \quad \Rightarrow \quad X = FZ + D \quad \left(\forall Z \in R^{N-P}\right)$$

A: P×N, B: P×1, F: N×(N–P), Z: (N–P)×1

- X = FZ+D is the (non-unique) solution of under-determined linear equation AX = B
- For any Z value, AX is equal to B

A Simple Example

$$AX = B \quad \Rightarrow \quad X = FZ + D$$

$$x_1 + x_2 = 1$$

$$x_1 = z$$
$$x_2 = -z + 1$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \cdot z + \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Subspace Reduction

$$\min_{X} \quad f(X)$$
$$\text{S.T.} \quad AX = B$$

$$X = FZ + D$$

$$\Rightarrow \quad \min_{Z} \quad f(FZ + D)$$

- Solve the optimal value Z by unconstrained optimization – minimizing f(FZ+D)

- Calculate the optimum X = FZ+D

Lagrange Multiplier

- Equality constraint can also be handled by Lagrange multiplier

- If X* is a local minimum of

$$\min_{X} \quad f(X)$$
$$\text{S.T.} \quad g_i(X) = 0 \quad (i = 1, 2, \cdots, P)$$

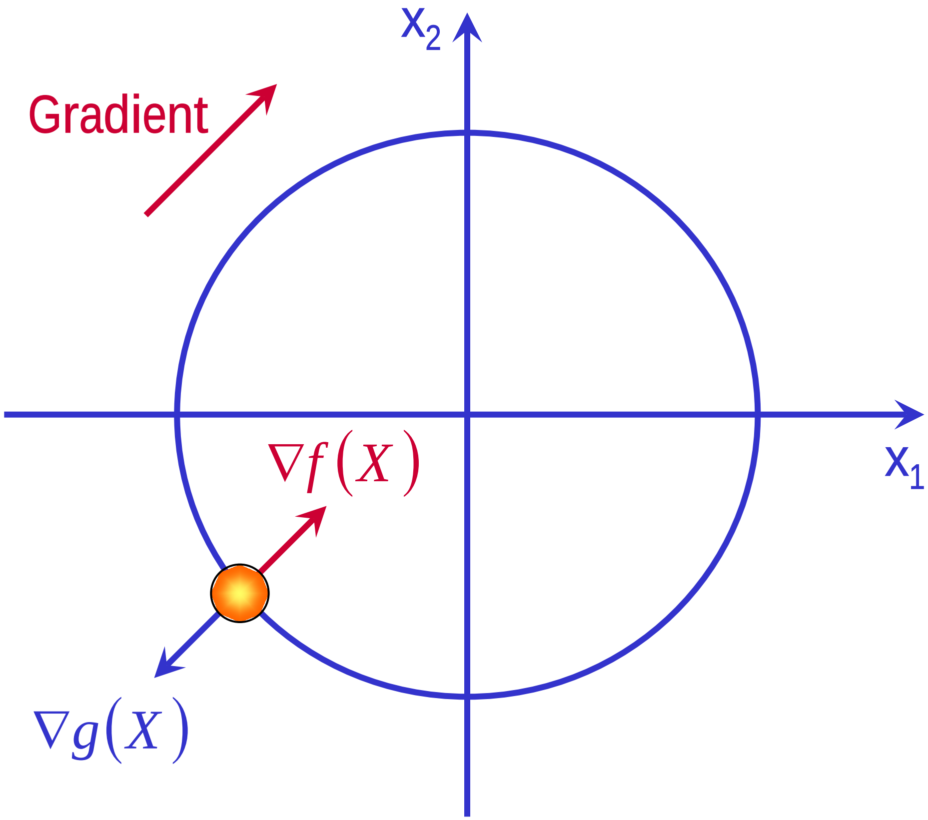
  - there exist λ_1, λ_2, ..., λ_P, called Lagrange multipliers, such that

$$\nabla f(X^*) + \sum_{i=1}^{P} \lambda_i \cdot \nabla g_i(X^*) = 0$$

A Simple Example

$$\min_{x_1, x_2} \quad f(X) = x_1 + x_2$$

$$\text{S.T.} \quad g(X) = x_1^2 + x_2^2 - 2 = 0$$



$$X^* = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

$$\nabla f(X^*) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\nabla g(X^*) = \begin{bmatrix} -2 \\ -2 \end{bmatrix}$$

Lagrange Multiplier

$$\min_{X} \quad f(X)$$
$$\text{S.T.} \quad g_i(X) = 0 \quad (i = 1, 2, \cdots, P)$$

$$\Rightarrow \quad \nabla f(X^*) + \sum_{i=1}^{P} \lambda_i \cdot \nabla g_i(X^*) = 0$$

- Optimality condition for linear constraints

$$\min_{X} \quad f(X)$$
$$\text{S.T.} \quad AX = B$$

$$g_i(X) = A(i,:) \cdot X - B(i) = 0$$

$$\begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix} \cdot X - \begin{bmatrix} \\ B \\ \\ \end{bmatrix} = 0$$

i-th row

$$\nabla g_i(X) = A(i,:)^T$$

Lagrange Multiplier

$$\min_X \quad f(X)$$
$$\text{S.T.} \quad g_i(X) = 0 \quad (i = 1, 2, \cdots, P)$$

$$\Rightarrow \quad \nabla f(X^*) + \sum_{i=1}^{P} \lambda_i \cdot \nabla g_i(X^*) = 0$$

- Optimality condition for linear constraints

$$\nabla g_i(X) = A(i,:)^T$$

$$\sum_{i=1}^{P} \lambda_i \cdot \nabla g_i(X) = \sum_{i=1}^{P} \lambda_i \cdot A(i,:)^T = A^T V$$

$$\nabla f(X^*) + A^T V = 0$$

$$\underbrace{\begin{bmatrix} A(1,:)^T & A(2,:)^T & \cdots \end{bmatrix}}_{A^T} \cdot \underbrace{\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \end{bmatrix}}_{V}$$

Linear Equality Constrained Quadratic Programming

$$\min_X \quad f(X) \qquad \text{S.T.} \quad AX = B \quad \Rightarrow \quad \nabla f(X^*) + A^T V = 0$$

- We first consider quadratic cost function
  - Any smooth nonlinear cost function can be locally approximated as a quadratic function (2nd-order Taylor expansion)

$$\min_X \quad f(X) = \frac{1}{2} X^T Q X + R^T X + C$$
$$\text{S.T.} \quad AX = B$$

$$\nabla f(X) = QX + R$$

$$QX^* + R + A^T V = 0$$

Linear Equality Constrained Quadratic Programming

$$\min_{X} \quad \frac{1}{2} X^T Q X + R^T X + C$$
$$\text{S.T.} \quad AX = B$$

- Optimality condition for quadratic programming

$$QX^* + R + A^T V = 0$$

$$AX^* = B$$

$$\begin{bmatrix} Q & A^T \\ A & 0 \end{bmatrix} \cdot \begin{bmatrix} X^* \\ V \end{bmatrix} = \begin{bmatrix} -R \\ B \end{bmatrix}$$

Solve linear equation to determine X* (optimal solution) and V (Lagrange multipliers)

Linear Equality Constrained Nonlinear Programming

$$\begin{aligned} &\min_X \quad \frac{1}{2} X^T Q X + R^T X + C \\ &\text{S.T.} \quad AX = B \end{aligned} \qquad \Rightarrow \qquad \begin{bmatrix} Q & A^T \\ A & 0 \end{bmatrix} \cdot \begin{bmatrix} X^* \\ V \end{bmatrix} = \begin{bmatrix} -R \\ B \end{bmatrix}$$

- Minimize nonlinear function f(X) given linear constraint AX = B

$$\begin{aligned} &\min_X \quad f(X) \\ &\text{S.T.} \quad AX = B \end{aligned}$$

$$f\left[X^{(k+1)}\right] \approx \frac{1}{2} \cdot \Delta X^T \cdot \nabla^2 f\left[X^{(k)}\right] \cdot \Delta X + \nabla f\left[X^{(k)}\right]^T \cdot \Delta X + f\left[X^{(k)}\right]$$

$$A \cdot \Delta X = A \cdot \left[X^{(k+1)} - X^{(k)}\right] = A \cdot X^{(k+1)} - AX^{(k)} = B - AX^{(k)}$$

Linear Equality Constrained Nonlinear Programming

$$\min_{X} \quad \frac{1}{2} X^T Q X + R^T X + C$$
$$\text{S.T.} \quad AX = B$$

$$\Rightarrow \quad \begin{bmatrix} Q & A^T \\ A & 0 \end{bmatrix} \cdot \begin{bmatrix} X^* \\ V \end{bmatrix} = \begin{bmatrix} -R \\ B \end{bmatrix}$$

$$f\left[X^{(k+1)}\right] \approx \frac{1}{2} \cdot \Delta X^T \cdot \nabla^2 f\left[X^{(k)}\right] \cdot \Delta X + \nabla f\left[X^{(k)}\right]^T \cdot \Delta X + f\left[X^{(k)}\right]$$

$$A \cdot \Delta X = B - AX^{(k)}$$

$$\begin{bmatrix} \nabla^2 f\left[X^{(k)}\right] & A^T \\ A & 0 \end{bmatrix} \cdot \begin{bmatrix} \Delta X \\ V \end{bmatrix} = \begin{bmatrix} -\nabla f\left[X^{(k)}\right] \\ B - AX^{(k)} \end{bmatrix}$$

Solve linear equation to determine ΔX

$$X^{(k+1)} = X^{(k)} + \Delta X$$

Update $X^{(k+1)}$

Linear Equality Constrained Nonlinear Programming

$$\min_X \quad f(X)$$
$$\text{S.T.} \quad AX = B$$

$$\begin{bmatrix} \nabla^2 f\left[X^{(k)}\right] & A^T \\ A & 0 \end{bmatrix} \cdot \begin{bmatrix} \Delta X \\ V \end{bmatrix} = \begin{bmatrix} -\nabla f\left[X^{(k)}\right] \\ B - AX^{(k)} \end{bmatrix}$$

- If $X^{(k)}$ is a feasible solution
  - We can start from an initial solution $X^{(0)}$ that is feasible
  - Even if $X^{(0)}$ is not feasible, $X^{(1)}$ is feasible after one iteration

$$AX^{(k)} = B$$

$$B - AX^{(k)} = 0$$

$$\begin{bmatrix} \nabla^2 f\left[X^{(k)}\right] & A^T \\ A & 0 \end{bmatrix} \cdot \begin{bmatrix} \Delta X \\ V \end{bmatrix} = \begin{bmatrix} -\nabla f\left[X^{(k)}\right] \\ 0 \end{bmatrix}$$

A Simple Example

$$\min_{x_1,x_2} \quad x_1^4 + x_2^4$$
$$\text{S.T.} \quad x_1 + x_2 = 1$$

$$X^{(0)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

(Feasible solution)

$$\min_{X} \quad f(X)$$
$$\text{S.T.} \quad AX = B$$

$$f(X) = x_1^4 + x_2^4$$
$$A = \begin{bmatrix} 1 & 1 \end{bmatrix}$$
$$B = 1$$

$$\nabla^2 f(X) = \begin{bmatrix} 12x_1^2 & 0 \\ 0 & 12x_2^2 \end{bmatrix}$$
$$\nabla f(X) = \begin{bmatrix} 4x_1^3 \\ 4x_2^3 \end{bmatrix}$$

A Simple Example

$$\min_{x_1, x_2} \quad x_1^4 + x_2^4$$
$$\text{S.T.} \quad x_1 + x_2 = 1$$

$$\begin{bmatrix} \nabla^2 f\left[X^{(k)}\right] & A^T \\ A & 0 \end{bmatrix} \cdot \begin{bmatrix} \Delta X \\ V \end{bmatrix} = \begin{bmatrix} -\nabla f\left[X^{(k)}\right] \\ 0 \end{bmatrix}$$

$$X^{(0)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\nabla^2 f\left[X^{(0)}\right] = \begin{bmatrix} 12 & 0 \\ 0 & 0 \end{bmatrix} \quad \nabla f\left[X^{(0)}\right] = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 12 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} \Delta X \\ V \end{bmatrix} = \begin{bmatrix} -4 \\ 0 \\ 0 \end{bmatrix}$$

$$\nabla^2 f(X) = \begin{bmatrix} 12x_1^2 & 0 \\ 0 & 12x_2^2 \end{bmatrix}$$

$$\nabla f(X) = \begin{bmatrix} 4x_1^3 \\ 4x_2^3 \end{bmatrix}$$

$$\Delta X = \begin{bmatrix} -0.33 \\ 0.33 \end{bmatrix} \quad X^{(1)} = \begin{bmatrix} 0.67 \\ 0.33 \end{bmatrix}$$

A Simple Example

$$\min_{x_1, x_2} \quad x_1^4 + x_2^4$$
$$\text{S.T.} \quad x_1 + x_2 = 1$$

$$\begin{bmatrix} \nabla^2 f\left[X^{(k)}\right] & A^T \\ A & 0 \end{bmatrix} \cdot \begin{bmatrix} \Delta X \\ V \end{bmatrix} = \begin{bmatrix} -\nabla f\left[X^{(k)}\right] \\ 0 \end{bmatrix}$$

$$X^{(1)} = \begin{bmatrix} 0.67 \\ 0.33 \end{bmatrix}$$

$$\nabla^2 f\left[X^{(1)}\right] = \begin{bmatrix} 5.39 & 0 \\ 0 & 1.31 \end{bmatrix} \quad \nabla f\left[X^{(1)}\right] = \begin{bmatrix} 1.20 \\ 0.14 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 5.39 & 0 & 1 \\ 0 & 1.31 & 1 \\ 1 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} \Delta X \\ V \end{bmatrix} = \begin{bmatrix} -1.20 \\ -0.14 \\ 0 \end{bmatrix}$$

$$\nabla^2 f(X) = \begin{bmatrix} 12x_1^2 & 0 \\ 0 & 12x_2^2 \end{bmatrix}$$

$$\Delta X = \begin{bmatrix} -0.16 \\ 0.16 \end{bmatrix} \quad X^{(2)} = \begin{bmatrix} 0.51 \\ 0.49 \end{bmatrix}$$

$$\nabla f(X) = \begin{bmatrix} 4x_1^3 \\ 4x_2^3 \end{bmatrix}$$

Linear Equality Constrained Nonlinear Programming

- Linear equality constraints can be efficiently handled by subspace reduction or Lagrange multiplier

- Nonlinear equality constraints and inequality constraints must be handled by a different algorithm
  - Interior point method (also referred to as barrier method)
  - More details in future lectures

Summary

- Constrained optimization
  - Linear equality constraint
  - Lagrange multiplier