18-660: Numerical Methods for Engineering Design and Optimization

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

Electrical & Computer ENGINEERING

Overview

- Conjugate Gradient Method (Part 3)
  - Conjugate gradient method
  - Convergence rate

Conjugate Search Direction

- Important equations about conjugate search direction

$$AX = B$$ → Linear equation

$$\min_{X} \quad f(X) = \frac{1}{2} X^T A X - B^T X + C$$ → Equivalent optimization

$$R^{(k)} = B - AX^{(k)}$$ → Residual definition

$$\nabla f\left[X^{(k)}\right] = AX^{(k)} - B = -R^{(k)}$$ → Residual vs. gradient

Conjugate Search Direction

- Important equations about conjugate search direction

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

$$\mu^{(k)} = \frac{D^{(k)T} R^{(k)}}{D^{(k)T} A D^{(k)}}$$

→ Iteration scheme

$$D^{(i)T} A D^{(j)} = 0$$

→ Conjugate search directions

$$D^{(k)T} R^{(k+1)} = 0$$

→ Orthogonal residual

Conjugate Gradient Method

- Step 1: start from an initial guess $X^{(0)}$, and set k = 0
- Step 2: calculate

$$D^{(0)} = R^{(0)} = B - AX^{(0)}$$

- Step 3: update solution

$$X^{(k+1)} = X^{(k)} + \mu^{(k)} D^{(k)} \quad \text{where} \quad \mu^{(k)} = \frac{D^{(k)T} R^{(k)}}{D^{(k)T} A D^{(k)}}$$

- Step 4: calculate residual

$$R^{(k+1)} = B - AX^{(k+1)}$$

- Step 5: determine search direction

$$D^{(k+1)} = R^{(k+1)} + \sum_{i=0}^{k} \beta_{k+1,i} D^{(i)} \quad \text{where} \quad \beta_{k+1,i} = -\frac{D^{(i)T} A R^{(k+1)}}{D^{(i)T} A D^{(i)}}$$

- Step 6: set k = k + 1 and go to Step 3

Conjugate Gradient Method

- Step 1: start from an initial guess $X^{(0)}$, and set k = 0
- Step 2: calculate

$$D^{(0)} = R^{(0)} = B - AX^{(0)}$$

- Step 3: update solution

$$X^{(k+1)} = X^{(k)} + \mu^{(k)} D^{(k)} \quad \text{where} \quad \mu^{(k)} = \frac{D^{(k)T} R^{(k)}}{D^{(k)T} A D^{(k)}}$$

- Step 4: calculate residual

$$R^{(k+1)} = B - AX^{(k+1)}$$

- Step 5: determine search direction

$$D^{(k+1)} = R^{(k+1)} + \sum_{i=0}^{k} \beta_{k+1,i} D^{(i)} \quad \text{where} \quad \beta_{k+1,i} = -\frac{D^{(i)T} A R^{(k+1)}}{D^{(i)T} A D^{(i)}}$$

- Step 6: set k = k + 1 and go to Step 3

Conjugate Gradient Method

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

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

$$R^{(k+1)} = B - AX^{(k+1)}$$

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

$$R^{(k+1)} = B - AX^{(k)} - \mu^{(k)} AD^{(k)}$$

$$R^{(k+1)} = R^{(k)} - \mu^{(k)} AD^{(k)}$$

Conjugate Gradient Method

- Step 1: start from an initial guess $X^{(0)}$, and set k = 0
- Step 2: calculate

$$D^{(0)} = R^{(0)} = B - AX^{(0)}$$

- Step 3: update solution

$$X^{(k+1)} = X^{(k)} + \mu^{(k)} D^{(k)} \quad \text{where} \quad \mu^{(k)} = \frac{D^{(k)T} R^{(k)}}{D^{(k)T} \underline{AD^{(k)}}}$$

- Step 4: calculate residual

$$R^{(k+1)} = R^{(k)} - \mu^{(k)} \underline{AD^{(k)}}$$

- Step 5: determine search direction

$$D^{(k+1)} = R^{(k+1)} + \sum_{i=0}^{k} \beta_{k+1,i} D^{(i)} \quad \text{where} \quad \beta_{k+1,i} = -\frac{D^{(i)T} AR^{(k+1)}}{D^{(i)T} AD^{(i)}}$$

- Step 6: set k = k + 1 and go to Step 3

Conjugate Gradient Method

- Step 1: start from an initial guess $X^{(0)}$, and set k = 0
- Step 2: calculate

$$D^{(0)} = R^{(0)} = B - AX^{(0)}$$

- Step 3: update solution

$$X^{(k+1)} = X^{(k)} + \mu^{(k)} D^{(k)} \quad \text{where} \quad \mu^{(k)} = \frac{D^{(k)T} R^{(k)}}{D^{(k)T} A D^{(k)}}$$

- Step 4: calculate residual

$$R^{(k+1)} = R^{(k)} - \mu^{(k)} A D^{(k)}$$

- Step 5: determine search direction

$$D^{(k+1)} = R^{(k+1)} + \sum_{i=0}^{k} \beta_{k+1,i} D^{(i)} \quad \text{where} \quad \beta_{k+1,i} = -\frac{D^{(i)T} A R^{(k+1)}}{D^{(i)T} A D^{(i)}}$$

- Step 6: set k = k + 1 and go to Step 3

Conjugate Gradient Method

- Orthogonal residuals

$$R^{(k+1)} = R^{(k)} - \mu^{(k)} A D^{(k)} \qquad D^{(i)T} A D^{(j)} = 0 \qquad D^{(k)T} R^{(k+1)} = 0$$

$$D^{(k)T} R^{(k+1)} = 0$$

$$D^{(k)T} R^{(k+2)} = D^{(k)T} \cdot \left[R^{(k+1)} - \mu^{(k+1)} A D^{(k+1)}\right] = D^{(k)T} R^{(k+1)} - \mu^{(k+1)} D^{(k)T} A D^{(k+1)} = 0$$

$$D^{(k)T} R^{(k+3)} = D^{(k)T} \cdot \left[R^{(k+2)} - \mu^{(k+2)} A D^{(k+2)}\right] = D^{(k)T} R^{(k+2)} - \mu^{(k+2)} D^{(k)T} A D^{(k+2)} = 0$$

$$\vdots$$

$$D^{(i)T} R^{(j)} = 0 \quad (i < j)$$

Conjugate Gradient Method

- Orthogonal residuals

$$D^{(i)T}R^{(j)} = 0 \quad (i < j) \qquad span\{D^{(0)}, D^{(1)}, \cdots, D^{(k)}\} = span\{R^{(0)}, R^{(1)}, \cdots, R^{(k)}\}$$

$$R^{(k)} \perp span\{D^{(0)}, D^{(1)}, \cdots, D^{(k-1)}\}$$

$$R^{(k)} \perp span\{R^{(0)}, R^{(1)}, \cdots, R^{(k-1)}\}$$

$$R^{(i)T}R^{(j)} = 0 \quad (i < j)$$

$$R^{(i)T}R^{(j)} = 0 \quad (i \neq j)$$

$R^{(i)}$ and $R^{(j)}$ are orthogonal

Conjugate Gradient Method

$$D^{(k+1)} = R^{(k+1)} + \sum_{i=0}^{k} \beta_{k+1,i} D^{(i)} \quad \text{where} \quad \beta_{k+1,i} = -\frac{D^{(i)T} A R^{(k+1)}}{D^{(i)T} A D^{(i)}}$$

$$R^{(k+1)} = R^{(k)} - \mu^{(k)} A D^{(k)}$$

$$R^{(i+1)T} R^{(k+1)} = \left[R^{(i)} - \mu^{(i)} A D^{(i)}\right]^T \cdot R^{(k+1)} = R^{(i)T} R^{(k+1)} - \mu^{(i)} D^{(i)T} A R^{(k+1)}$$

$$\mu^{(i)} D^{(i)T} A R^{(k+1)} = R^{(i)T} R^{(k+1)} - R^{(i+1)T} R^{(k+1)}$$

Conjugate Gradient Method

$$D^{(k+1)} = R^{(k+1)} + \sum_{i=0}^{k} \beta_{k+1,i} D^{(i)} \quad \text{where} \quad \beta_{k+1,i} = -\frac{D^{(i)T} A R^{(k+1)}}{D^{(i)T} A D^{(i)}}$$

$$\mu^{(i)} D^{(i)T} A R^{(k+1)} = R^{(i)T} R^{(k+1)} - R^{(i+1)T} R^{(k+1)} \qquad R^{(i)T} R^{(j)} = 0$$

- i = k

$$\mu^{(i)} D^{(i)T} A R^{(k+1)} = -R^{(k+1)T} R^{(k+1)}$$

- i < k

$$\mu^{(i)} D^{(i)T} A R^{(k+1)} = 0$$

$$D^{(k+1)} = R^{(k+1)} + \beta_{k+1,k} D^{(k)} \quad \text{where} \quad \beta_{k+1,k} = -\frac{D^{(k)T} A R^{(k+1)}}{D^{(k)T} A D^{(k)}} = \frac{R^{(k+1)T} R^{(k+1)}}{\mu^{(k)} D^{(k)T} A D^{(k)}}$$

Conjugate Gradient Method

- Step 1: start from an initial guess $X^{(0)}$, and set k = 0
- Step 2: calculate

$$D^{(0)} = R^{(0)} = B - AX^{(0)}$$

- Step 3: update solution

$$X^{(k+1)} = X^{(k)} + \mu^{(k)} D^{(k)} \quad \text{where} \quad \mu^{(k)} = \frac{D^{(k)T} R^{(k)}}{D^{(k)T} A D^{(k)}}$$

- Step 4: calculate residual

$$R^{(k+1)} = R^{(k)} - \mu^{(k)} A D^{(k)}$$

- Step 5: determine search direction

$$D^{(k+1)} = R^{(k+1)} + \beta_{k+1,k} D^{(k)} \quad \text{where} \quad \beta_{k+1,k} = \frac{R^{(k+1)T} R^{(k+1)}}{\mu^{(k)} D^{(k)T} A D^{(k)}}$$

- Step 6: set k = k + 1 and go to Step 3

Conjugate Gradient Method

- Step 1: start from an initial guess $X^{(0)}$, and set k = 0
- Step 2: calculate

$$D^{(0)} = R^{(0)} = B - AX^{(0)}$$

- Step 3: update solution

$$X^{(k+1)} = X^{(k)} + \mu^{(k)} D^{(k)} \quad \text{where} \quad \mu^{(k)} = \frac{D^{(k)T} R^{(k)}}{D^{(k)T} A D^{(k)}}$$

- Step 4: calculate residual

$$R^{(k+1)} = R^{(k)} - \mu^{(k)} A D^{(k)}$$

- Step 5: determine search direction

$$D^{(k+1)} = R^{(k+1)} + \beta_{k+1,k} D^{(k)} \quad \text{where} \quad \beta_{k+1,k} = \frac{R^{(k+1)T} R^{(k+1)}}{\mu^{(k)} D^{(k)T} A D^{(k)}}$$

- Step 6: set k = k + 1 and go to Step 3

Conjugate Gradient Method

$$D^{(k+1)} = R^{(k+1)} + \beta_{k+1,k} D^{(k)} \quad \text{where} \quad \beta_{k+1,k} = \frac{R^{(k+1)T} R^{(k+1)}}{\mu^{(k)} D^{(k)T} A D^{(k)}}$$

$$\mu^{(k)} = \frac{D^{(k)T} R^{(k)}}{D^{(k)T} A D^{(k)}}$$

$$\mu^{(k)} D^{(k)T} A D^{(k)} = D^{(k)T} R^{(k)}$$

$$\beta_{k+1,k} = \frac{R^{(k+1)T} R^{(k+1)}}{\mu^{(k)} D^{(k)T} A D^{(k)}} = \frac{R^{(k+1)T} R^{(k+1)}}{D^{(k)T} R^{(k)}}$$

Conjugate Gradient Method

- Step 1: start from an initial guess $X^{(0)}$, and set k = 0
- Step 2: calculate

$$D^{(0)} = R^{(0)} = B - AX^{(0)}$$

- Step 3: update solution

$$X^{(k+1)} = X^{(k)} + \mu^{(k)} D^{(k)} \quad \text{where} \quad \mu^{(k)} = \frac{D^{(k)T} R^{(k)}}{D^{(k)T} A D^{(k)}}$$

- Step 4: calculate residual

$$R^{(k+1)} = R^{(k)} - \mu^{(k)} A D^{(k)}$$

- Step 5: determine search direction

$$D^{(k+1)} = R^{(k+1)} + \beta_{k+1,k} D^{(k)} \quad \text{where} \quad \beta_{k+1,k} = \frac{R^{(k+1)T} R^{(k+1)}}{D^{(k)T} R^{(k)}}$$

- Step 6: set k = k + 1 and go to Step 3

Convergence Rate

- Mathematical analysis of convergence is quite complex
  - We will directly show the results, but not detailed analysis

- Convergence rate of conjugate gradient method

$$AX = B$$

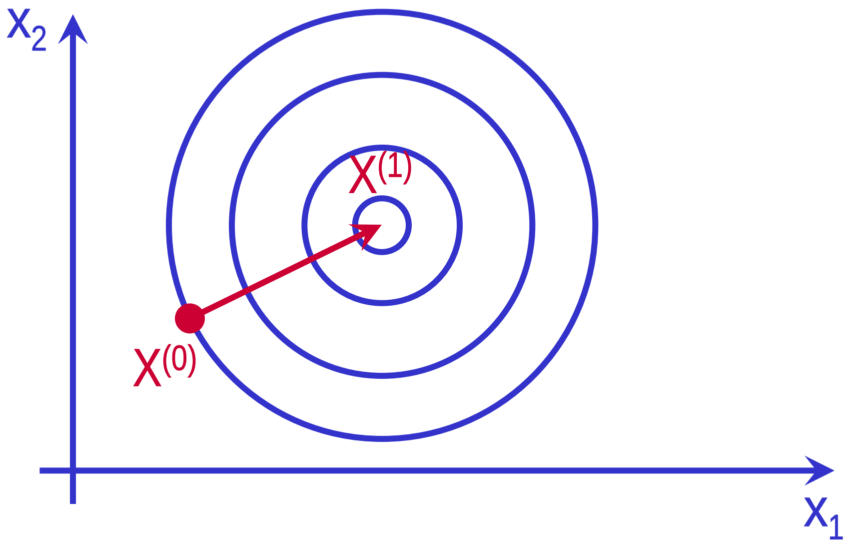
$$\left\| X^{(k+1)} - X \right\| \le \left[\frac{\sqrt{\kappa(A)} - 1}{\sqrt{\kappa(A)} + 1} \right]^k \cdot \left\| X^{(0)} - X \right\|$$

Exact solution (X)   Condition number ($\kappa(A)$)

Convergence Rate

$$\left\| X^{(k+1)} - X \right\| \le \left[\frac{\sqrt{\kappa(A)} - 1}{\sqrt{\kappa(A)} + 1} \right]^k \cdot \left\| X^{(0)} - X \right\|$$

- Property 1: Converge by one iteration if **κ**(A) = 1
  - E.g., A is an identity matrix



Convergence Rate

$$\left\| X^{(k+1)} - X \right\| \le \left[\frac{\sqrt{\kappa(A)} - 1}{\sqrt{\kappa(A)} + 1} \right]^k \cdot \left\| X^{(0)} - X \right\|$$

- Property 2: slowly converge if **κ**(A) is large
  - I.e., AX = B is ill-conditioned

- In this case, we want to improve convergence rate by pre-conditioning
  - Scale the matrix A to reduce its condition number

Summary

- Conjugate gradient method (Part 3)
  - Conjugate gradient method
  - Convergence rate