# Numerical Computing

## Lecture 7: Iterative Methods for Linear Systems

Francisco Richter Mendoza

Università della Svizzera Italiana
Faculty of Informatics, Lugano, Switzerland

# Lecture Overview

- **Matrix Splitting Framework**
- **Classical Methods**: Jacobi, Gauss-Seidel, SOR
- **Convergence Theory**: Spectral radius analysis
- **Krylov Subspace Methods**: CG and GMRES
- **Preconditioning Techniques**
- **Practical Implementation**

# Matrix Splitting Framework

## Basic Idea

Split matrix $A = M - N$ where $M$ is easily invertible

$$Ax = b \tag{1}$$
$$(M - N)x = b \tag{2}$$
$$Mx = Nx + b \tag{3}$$
$$x = M^{-1}Nx + M^{-1}b \tag{4}$$

**Iterative scheme:** $x^{(k+1)} = M^{-1}Nx^{(k)} + M^{-1}b$
**Convergence condition:** $\rho(M^{-1}N) < 1$

# Classical Iterative Methods

Let $A = D - L - U$ where:

- $D$: diagonal part
- $L$: strictly lower triangular
- $U$: strictly upper triangular

## Method Definitions

$$\text{Jacobi:} \quad x^{(k+1)} = D^{-1}(L+U)x^{(k)} + D^{-1}b \tag{5}$$

$$\text{Gauss-Seidel:} \quad x^{(k+1)} = (D-L)^{-1}Ux^{(k)} + (D-L)^{-1}b \tag{6}$$

$$\text{SOR:} \quad x^{(k+1)} = (D-\omega L)^{-1}[(1-\omega)D + \omega U]x^{(k)} \tag{7}$$

$$+ \omega(D-\omega L)^{-1}b \tag{8}$$

# Classical Methods: Convergence Analysis

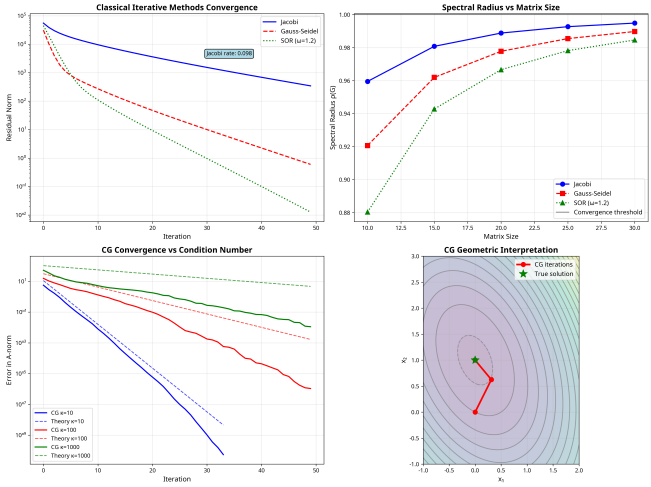


Figure: Comparison of classical iterative methods showing convergence rates, spectral radius behavior, SOR optimization, and geometric CG interpretation

# Convergence Theory

## Theorem (Convergence Condition)

*The iterative method* $x^{(k+1)} = Gx^{(k)} + c$ *converges for any initial guess if and only if* $\rho(G) < 1$.

## Key Results

- **Jacobi**: Converges if $A$ is strictly diagonally dominant
- **Gauss-Seidel**: Converges if $A$ is SPD or strictly diagonally dominant
- **SOR**: Optimal parameter $\omega_{opt} = \frac{2}{1+\sqrt{1-\rho(G_J)^2}}$

**Asymptotic convergence rate:** $R = -\ln(\rho(G))$

# Krylov Subspace Methods

## Krylov Subspace

$\mathcal{K}_k(A, r_0) = \text{span}\{r_0, Ar_0, A^2 r_0, \ldots, A^{k-1} r_0\}$

- **Conjugate Gradient (CG)**: For SPD matrices

$$\|x_k - x^*\|_A \leq 2 \left( \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^k \|x_0 - x^*\|_A \tag{9}$$

- **GMRES**: For general matrices

$$\|r_k\| = \min_{p \in \mathcal{P}_k} \|p(A) r_0\| \tag{10}$$

**Key advantage:** Optimal approximation in Krylov subspace

# Krylov Methods: CG vs GMRES Analysis

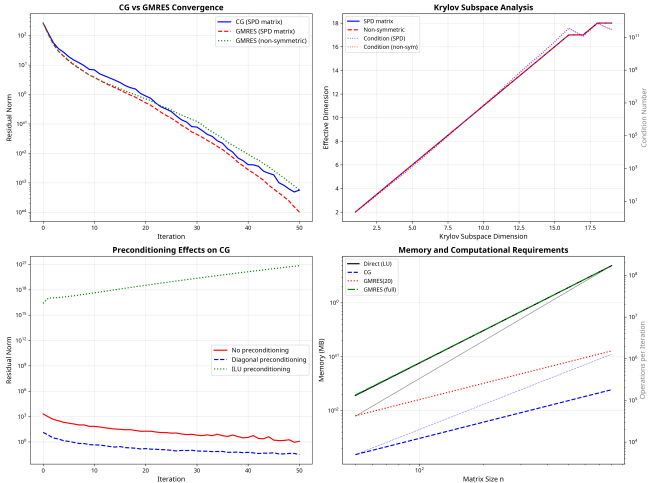


Figure: Comprehensive comparison of CG and GMRES methods showing convergence behavior, condition number effects, memory requirements, and subspace growth

# Conjugate Gradient Algorithm

## CG Algorithm for $Ax = b$ (A SPD)

1. $r_0 = b - Ax_0$, $p_0 = r_0$
2. For $k = 0, 1, 2, \ldots$ until convergence:

$$\alpha_k = \frac{r_k^T r_k}{p_k^T A p_k} \tag{11}$$

$$x_{k+1} = x_k + \alpha_k p_k \tag{12}$$

$$r_{k+1} = r_k - \alpha_k A p_k \tag{13}$$

$$\beta_k = \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k} \tag{14}$$

$$p_{k+1} = r_{k+1} + \beta_k p_k \tag{15}$$

**Memory:** Only 4 vectors of length $n$

# GMRES Algorithm

## GMRES for General Matrices

- Build orthonormal basis $\{v_1, v_2, \ldots, v_k\}$ for $\mathcal{K}_k(A, r_0)$
- Solve least squares problem:

$$\min_{y \in \mathbb{R}^k} \|\beta e_1 - H_k y\|_2 \tag{16}$$

  where $H_k$ is upper Hessenberg matrix
- Update: $x_k = x_0 + V_k y_k$

**Advantages:** Works for any matrix, minimizes residual norm
**Disadvantages:** Growing memory, restart needed

# Preconditioning

## Basic Idea

Solve $M^{-1}Ax = M^{-1}b$ instead of $Ax = b$

**Goals:**

- Reduce condition number: $\kappa(M^{-1}A) \ll \kappa(A)$
- $M$ should be easy to invert

## Common Preconditioners

- **Diagonal**: $M = \text{diag}(A)$
- **Incomplete LU**: $M \approx LU$ (sparse)
- **Multigrid**: Optimal for elliptic PDEs
- **Domain Decomposition**: Parallel-friendly

# Practical Implementation Guidelines

## Method Selection

- **SPD matrices**: Use CG with good preconditioner
- **General matrices**: Use GMRES with restart
- **Large sparse**: Classical methods with good ordering

## Stopping Criteria

$$\frac{\|r_k\|}{\|r_0\|} < \text{tol} \quad \text{or} \quad \frac{\|r_k\|}{\|b\|} < \text{tol} \tag{17}$$

**Typical tolerance:** $10^{-6}$ to $10^{-12}$

# Computational Complexity

## Cost per Iteration

- **Jacobi/Gauss-Seidel**: $O(n^2)$ for dense, $O(\text{nnz})$ for sparse
- **CG**: $O(n^2)$ + 1 matrix-vector product
- **GMRES**: $O(kn)$ + 1 matrix-vector product (k = iteration)

## Total Complexity

- **CG**: $O(n^{3/2})$ for well-conditioned SPD
- **GMRES**: $O(n^2)$ to $O(n^3)$ depending on restart
- **Multigrid**: $O(n)$ optimal complexity

# Key Takeaways

- **Matrix splitting** provides unified framework for classical methods
- **Spectral radius** determines convergence: $\rho(G) < 1$
- **Krylov methods** are optimal for their respective matrix classes
- **CG** is ideal for SPD systems with $O(\sqrt{\kappa})$ convergence
- **GMRES** handles general matrices but requires restarts
- **Preconditioning** is essential for practical performance
- **Method selection** depends on matrix structure and resources

**Next lecture:** Eigenvalue problems and power methods