Stationary Iterative Methods and Krylov Subspace Methods

Kui Du

kuidu@xmu.edu.cn

School of Mathematical Sciences, Xiamen University

https://kuidu.github.io

Outline

Stationary Iterative Methods

- Consider a linear system of equations

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{A} \in \mathbb{R}^{n\times n}, \quad \mathbf{b} \in \mathbb{R}^{n}.$$

- A *splitting* of $\mathbf{A} \in \mathbb{R}^{n\times n}$ is a decomposition

$$\mathbf{A} = \mathbf{M} - \mathbf{N},$$

  with $\mathbf{M}$ nonsingular.

- The equation

$$\mathbf{A}\mathbf{x} = (\mathbf{M} - \mathbf{N})\mathbf{x} = \mathbf{b}$$

  implies

$$\mathbf{x} = \mathbf{M}^{-1}\mathbf{N}\mathbf{x} + \mathbf{M}^{-1}\mathbf{b} := \mathbf{R}\mathbf{x} + \mathbf{c}.$$

  Given a starting vector $\mathbf{x}_0$, we obtain an iterative method

$$\mathbf{x}_{k+1} = \mathbf{R}\mathbf{x}_k + \mathbf{c}, \quad k = 0, 1, 2, \ldots$$

Stationary Iterative Methods

- Correction form

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{M}^{-1}\mathbf{r}_k,$$

where the residual

$$\mathbf{r}_k = \mathbf{b} - \mathbf{A}\mathbf{x}_k.$$

- Error recurrence

$$\mathbf{e}_k = \mathbf{x} - \mathbf{x}_k, \qquad \mathbf{e}_{k+1} = \mathbf{M}^{-1}\mathbf{N}\mathbf{e}_k$$

- Difference recurrence

$$\mathbf{d}_k = \mathbf{x}_{k+1} - \mathbf{x}_k, \qquad \mathbf{d}_{k+1} = \mathbf{M}^{-1}\mathbf{N}\mathbf{d}_k$$

- Residual recurrence

$$\mathbf{r}_{k+1} = (\mathbf{I} - \mathbf{A}\mathbf{M}^{-1})\mathbf{r}_k = \mathbf{N}\mathbf{M}^{-1}\mathbf{r}_k.$$

Stationary Iterative Methods

- Convergence criterion

  The iteration $\mathbf{x}_{k+1} = \mathbf{R}\mathbf{x}_k + \mathbf{c}$ converges to the solution of $\mathbf{A}\mathbf{x} = \mathbf{b}$ for all starting vectors $\mathbf{x}_0$ if and only if $\rho(\mathbf{R}) < 1$.

- Desirable splitting

  $\mathbf{R}\mathbf{v} = \mathbf{M}^{-1}\mathbf{N}\mathbf{v}$ and $\mathbf{c} = \mathbf{M}^{-1}\mathbf{b}$ are easy to evaluate, and $\rho(\mathbf{R})$ is small (< 1).

- Examples

  (1). Jacobi's method

  (2). Gauss–Seidel method

  (3). SOR(ω)

  (4). Domain decomposition methods

  (5). Multigrid methods

Outline

1. Stationary Iterative Methods
2. Krylov Subspace Methods
3. Preconditioning
4. Summary

Generalized Minimal Residual Method (GMRES)

- Idea of GMRES: Consider a nonsingular linear system

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{A} \in \mathbb{R}^{n\times n}, \quad \mathbf{b} \in \mathbb{R}^{n}.$$

For any initial guess $\mathbf{x}_0$, at step k, GMRES finds the kth approximate solution

$$\mathbf{x}_k = \underset{\mathbf{x}\in\mathbf{x}_0+\mathcal{K}_k(\mathbf{A},\mathbf{r}_0)}{\operatorname{argmin}} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2,$$

where $\mathbf{r}_0 := \mathbf{b} - \mathbf{A}\mathbf{x}_0$ and

$$\mathcal{K}_k(\mathbf{A}, \mathbf{r}_0) = \operatorname{span}\{\mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \ldots, \mathbf{A}^{k-1}\mathbf{r}_0\}.$$

For the residual $\mathbf{r}_k = \mathbf{b} - \mathbf{A}\mathbf{x}_k$, we have

$$\|\mathbf{r}_k\|_2 = \min_{\mathbf{x}\in\mathbf{x}_0+\mathcal{K}_k(\mathbf{A},\mathbf{r}_0)} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2 \quad \text{and} \quad \mathbf{r}_k \perp \mathbf{A}\mathcal{K}_k.$$

Conjugate Gradient (CG)

- Idea of CG: Consider an SPD linear system

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{A} \in \mathbb{R}^{n\times n}, \quad \mathbf{b} \in \mathbb{R}^{n}.$$

For initial guess $\mathbf{x}_0$, at step k, the conjugate gradient method finds an approximate solution

$$\mathbf{x}_k \in \mathbf{x}_0 + \mathcal{K}_k(\mathbf{A}, \mathbf{r}_0)$$

satisfying

$$\mathbf{r}_k := \mathbf{b} - \mathbf{A}\mathbf{x}_k \perp \mathcal{K}_k(\mathbf{A}, \mathbf{r}_0),$$

where

$$\mathcal{K}_k(\mathbf{A}, \mathbf{r}_0) := \operatorname{span}\{\mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \ldots, \mathbf{A}^{k-1}\mathbf{r}_0\}.$$

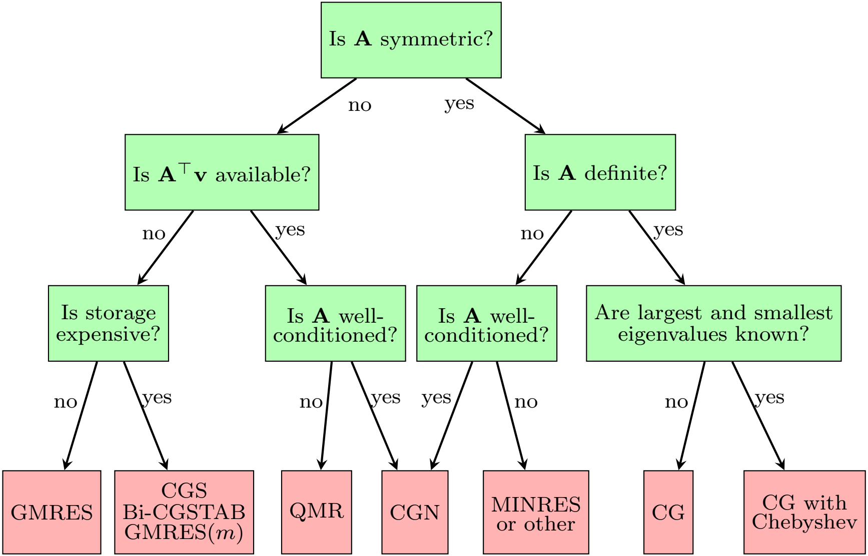
We have

$$\|\mathbf{x} - \mathbf{x}_k\|_{\mathbf{A}} = \min_{\mathbf{z}\in\mathbf{x}_0+\mathcal{K}_k(\mathbf{A},\mathbf{r}_0)} \|\mathbf{x} - \mathbf{z}\|_{\mathbf{A}}.$$

Krylov Subspace Methods

- Matrix-free implementation: $\mathbf{A}\mathbf{v}$ and $\mathbf{A}^\top\mathbf{v}$ are enough.
- GMRES = generalized minimal residual; MINRES
- CG = conjugate gradient; CGN; CGLS; LSQR
- Bi-CG = bi-conjugate gradient
  COCG = Bi-CG for complex symmetric systems
- CGS = conjugate gradients squares
  transpose free, but is twice as erratic
- Bi-CGSTAB = stabilized Bi-CG (via stabilizing CGS)
  Significantly smooths the convergence of Bi-CG, transpose free
- QMR = quasi-minimal residuals
  Pronounced effect on the smoothness of convergence
- TFQMR = tanspose-free QMR
  transpose free and smooth convergence of QMR
- ...

Decision Tree for Choosing Krylov Solvers



Outline

1. Stationary Iterative Methods
2. Krylov Subspace Methods
3. Preconditioning
4. Summary

Preconditioning

- To improve the convergence of Krylov subspace methods, it is important to have a preconditioner (suitable approximation for the original coefficient matrix $\mathbf{A}$), denoted by $\mathbf{M}$.
- Left preconditioning, i.e.,

$$\mathbf{M}^{-1}\mathbf{A}\mathbf{x} = \mathbf{M}^{-1}\mathbf{b}.$$

- Right preconditioning is often used, i.e.,

$$\mathbf{A}\mathbf{M}^{-1}\mathbf{z} = \mathbf{b}, \quad \mathbf{x} = \mathbf{M}^{-1}\mathbf{z},$$

  because it produces the unpreconditioned residual.
- If $\mathbf{M}$ is SPD, two-sided preconditioning:

$$\mathbf{M} = \mathbf{L}\mathbf{L}^\top, \quad (\mathbf{L}^{-1}\mathbf{A}\mathbf{L}^{-\top})\mathbf{L}^\top\mathbf{x} = \mathbf{L}^{-1}\mathbf{b}.$$

Preconditioning

- How to find a good preconditioner? It's problem dependent.

  Example. Let

$$\mathcal{A} = \begin{bmatrix} \mathbf{A} & \mathbf{B}^\top \\ \mathbf{C} & \mathbf{0} \end{bmatrix}$$

  and

$$\mathcal{M} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}\mathbf{A}^{-1}\mathbf{B}^\top \end{bmatrix},$$

  where $\mathbf{A} \in \mathbb{R}^{m\times m}$ is invertible, and $\mathbf{B}, \mathbf{C} \in \mathbb{R}^{n\times m}$ with $m \geq n$. Assume that $-\mathbf{C}\mathbf{A}^{-1}\mathbf{B}^\top$ is invertible.

  The preconditioned matrix $\mathcal{M}^{-1}\mathcal{A}$ is diagonalizable and has at most three distinct eigenvalues

$$1, \quad (1+\sqrt{5})/2, \quad (1-\sqrt{5})/2.$$

  GMRES converges within three steps.

Preconditioning

- Stationary iterative methods

  Any iterative technique can be used as a preconditioner.

- Algebraic preconditioning

  Incomplete Cholesky or LU factorization.

- Physical approximation problem preconditioning

  Constant-coefficient or symmetric approximation.

  Splitting of a multi-term operator.

  Dimensional splitting or ADI.

  Periodic or convolution approximation.

- Polynomial preconditioning

  PROXY-GMRES (SIMAX 2021).

- Multipreconditioning: Flexible GMRES

- ...

Preconditioning in Practice

- In many cases (e.g., multigrid methods and domain decomposition methods) the structure of $\mathbf{M}$ is unknown or $\mathbf{M}$ is expensive to compute.
- We never explicitly form $\mathbf{M}^{-1}$. Only the action of applying the preconditioner solve operation $\mathbf{M}^{-1}$ to a given vector is computed in iterative methods. So $\mathbf{M}^{-1}\mathbf{z}$ must be cheap.
- Example: we would like to use a stationary method

$$\mathbf{x}_k = \mathbf{M}^{-1}\mathbf{N}\mathbf{x}_{k-1} + \mathbf{M}^{-1}\mathbf{b}$$

as a preconditioner, but we do not know explicitly $\mathbf{M}$. If

$$\texttt{function} \quad \texttt{xk} = \texttt{stationary}(\texttt{A}, \texttt{b}, \texttt{x0}, \texttt{k}),$$

then

$$\mathbf{M}^{-1}\mathbf{z} = \texttt{stationary}(\texttt{A}, \texttt{z}, \texttt{0}, \texttt{1}),$$

$$\mathbf{M}^{-1}\mathbf{A}\mathbf{z} = \mathbf{z} - \texttt{stationary}(\texttt{A}, \texttt{0}, \texttt{z}, \texttt{1}).$$

Outline

1. Stationary Iterative Methods
2. Krylov Subspace Methods
3. Preconditioning
4. Summary

Summary

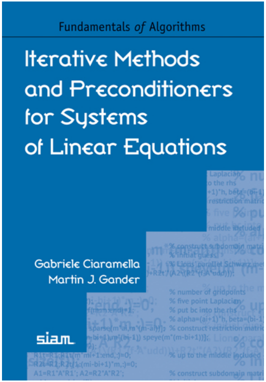
- Krylov subspace methods (with preconditioning) can be used in a matrix-free way.
- All the iterative methods like Jacobi, Gauss–Seidel, and SOR, should be used as preconditioners for a Krylov method. The Krylov method serves as an accelerator of convergence.
- We never explicitly form $\mathbf{M}^{-1}$. Only the action of applying the preconditioner solve operation $\mathbf{M}^{-1}$ to a given vector is computed in iterative methods.
- PETSc: the Portable, Extensible Toolkit for Scientific Computation
- Automatic selection of solver and preconditioner???

A Reference Book

- Iterative Methods and Preconditioners for Systems of Linear Equations

  Authors: Gabriele Ciaramella and Martin J. Gander

  SIAM, 2022



Thanks!