Lecture 08 – Solving Linear Systems (Part 2)

Prof. Brendan Kochunas

NERS/ENGR 570 - Methods and Practice of Scientific Computing (F20)

COLLEGE OF ENGINEERING
NUCLEAR ENGINEERING & RADIOLOGICAL SCIENCES
UNIVERSITY OF MICHIGAN

Outline

- Recap
- Convergence of Fixed point Methods
- Multigrid Methods
- Krylov Methods

Learning Objectives: By the end of Today's Lecture you should be able to

- (Knowledge) Determine the rate of convergence for fixed point methods
- (Knowledge) understand how multigrid exploits fixed point methods
- (Knowledge) explain what an eigendecomposition of a matrix is

Review/Recap

Basic Linear Algebra Operations

Residual and Norms of Vectors

$\mathbf{r} = \mathbf{A}\mathbf{x} - \mathbf{b}$ residual

$\|\mathbf{r}\|_1 = \sum_i |r_i|$ 1-norm

$\|\mathbf{r}\|_2 = \sqrt{\sum_i r_i^2}$ 2-norm (“average error”)

$\|\mathbf{r}\|_\infty = \max_i(|r_i|)$ ∞-norm (“max local error”)

$\|\mathbf{r}\|_p = \left(\sum_i |r_i|^p\right)^{1/p}$ *p*-norm

Inner/Dot Product (vector-vector multiply)

$$\mathbf{u}^T \cdot \mathbf{v} = \sum_i u_i v_i$$

Matrix-vector Multiply

$$\mathbf{A}\mathbf{x} = \mathbf{b} \rightarrow b_i = \sum_j a_{i,j} x_j$$

Matrix-Matrix Multiply

$$\mathbf{A}\mathbf{B} = \mathbf{C} \rightarrow c_{i,j} = \sum_k a_{i,k} b_{k,j}$$

Overview of Solution Methods

$$\boxed{Ax = b}$$

Iterative

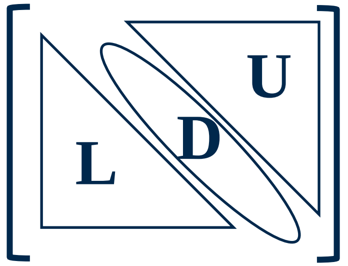
- ☆ Fixed Point (Classical Iterative Methods)
- Multigrid Methods (1970's-1980's)
- Krylov Methods
  - GMRES
  - Lanczos & Arnoldi (1950's)
  - Conjugate Gradient

Direct

- Gaussian Elim. / LU
- QR decomposition

Classical Iteration Schemes

$$\mathbf{A} = \boxed{(\mathbf{L} + \mathbf{D} + \mathbf{U})} =$$



Jacobi

$$x_i^{(\ell+1)} = -\frac{1}{a_{ii}}\left(\sum_{j=1}^{i-1} l_{ij} x_j^{(\ell)} + \sum_{j=i+1}^{n} u_{ij} x_j^{(\ell)}\right) x_i^{(\ell)} + b_i$$

$$\mathbf{x}^{(\ell+1)} = -\mathbf{D}^{-1}(\mathbf{L} + \mathbf{U})\mathbf{x}^{(\ell)} + \mathbf{D}^{-1}\mathbf{b}$$

$$\mathbf{F} = -\mathbf{D}^{-1}(\mathbf{L} + \mathbf{U})$$

$$\mathbf{c} = \mathbf{D}^{-1}\mathbf{b}$$

Gauss-Siedel

$$x_i^{(\ell+1)} = -\frac{1}{a_{ii}}\left(\sum_{j=1}^{i-1} l_{ij} x_j^{(\ell+1)} + \sum_{j=i+1}^{n} u_{ij} x_j^{(\ell)}\right) x_i^{(\ell)} + b_i$$

$$\mathbf{x}^{(\ell+1)} = \boxed{-(\mathbf{D} + \mathbf{L})^{-1}\mathbf{U}}\mathbf{x}^{(\ell)} + (\mathbf{D} + \mathbf{L})^{-1}\mathbf{b}$$

$$\mathbf{F} = -(\mathbf{D}^{-1} + \mathbf{L})\mathbf{U}$$

$$\mathbf{c} = (\mathbf{D}^{-1} + \mathbf{L})\mathbf{b}$$

$$\boxed{\mathbf{x}^{(\ell+1)} = \mathbf{F}\mathbf{x}^{(\ell)} + \mathbf{c}}$$

Definition of a Fixed-Point Iteration

Do they converge?

- Fixed-point iteration

$$\mathbf{x}^{(\ell+1)} = \mathbf{F}\mathbf{x}^{(\ell)} + \mathbf{c}$$

- Express iterate as combination of exact solution and error

$$\mathbf{x} + \boldsymbol{\varepsilon}^{(\ell+1)} = \mathbf{F}\left(\mathbf{x} + \boldsymbol{\varepsilon}^{(\ell)}\right) + \mathbf{c}$$

- If the method converges then:

$$\lim_{\ell \to \infty} \boldsymbol{\varepsilon}^{(\ell)} = 0$$

Convergence of Fixed-Point Iterative Methods

Prerequisite: Eigen-Decomposition of a Matrix

$A\underline{x} = \lambda \underline{x}$ unknowns

$x =$ eigenvector

$\lambda =$ eigenvalue

A must be square

A must be diagonalizable

$Ax_i = \lambda_i x_i$

$$A \begin{bmatrix} x_1 \,|\, x_2 \,|\, \ldots \end{bmatrix} = \begin{bmatrix} x_1 \,|\, x_2 \,|\, x_3 \ldots \end{bmatrix} \underbrace{\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \lambda_3 & \\ & & & \ddots \end{bmatrix}}_{\Lambda}$$

$AQ = Q\Lambda$

$AQQ^{-1} = Q\Lambda Q^{-1} \Rightarrow \underline{\underline{A = Q\Lambda Q^{-1}}}$

$x_i^T \cdot x_j = 0 \quad i \neq j$

$\|x_i\| = ? = 1$

$\boxed{x_i^T \cdot x_j = \delta_{ij}}$

Conditions for Convergence

$\cancel{X} \rightarrow \varepsilon^{(\ell+1)} = F\cancel{X} + F\varepsilon^{(0)} + \cancel{c}$

$\varepsilon^{(\ell+1)} = F\varepsilon^{(\ell)}$

$\varepsilon^{(1)} = F\varepsilon^{(0)}$

$\varepsilon^{(2)} = F\varepsilon^{(1)} = F^2\varepsilon^{(0)}$

$\varepsilon^{(\ell)} = F^{\ell}\varepsilon^{(0)}$

$\star \lim_{\ell\to\infty} \varepsilon^{(\ell)} = 0$

$\lim_{\ell\to\infty} F^{\ell} = 0$

$\lim_{\ell\to\infty} \Lambda^{\ell} \Rightarrow 0$

$\boxed{\max_i |\lambda_i| < 1}$

$F = Q\Lambda Q^{-1}$

$F^2 = (Q\Lambda \cancel{Q^{-1})(Q}\Lambda Q^{-1})$

$F^2 = Q\Lambda^2 Q^{-1}$

$F^{\ell} = Q\Lambda^{\ell} Q^{-1}$

$\star \boxed{\lim_{\ell\to\infty} \lambda_i^{\ell} = 0}$

$|\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \ldots$

$\boxed{\lim_{\ell\to\infty} \max_i |\lambda_i^{\ell}| = 0}$

More about the Spectral Radius

$\lambda_{max} = \max_i |\lambda_i| < 1$

λ_{max} = spectral radius ρ

$\rho(A) \equiv$ spectral radius of A

Condition for convergence $\rho(F) < 1$

Jacobi: $F = D^{-1}(L+U)$

GS: $F = (D+L)^{-1}U$

~~A~~ spectral radius determines the asymptotic rate of convergence of a fixed point method

$$\varepsilon^{(\ell+1)} = F\varepsilon^{(\ell)} \rightarrow \|\varepsilon^{(\ell+1)}\| \leq \|F\| \, \|\varepsilon^{(\ell)}\| \Rightarrow \frac{\|\varepsilon^{(\ell+1)}\|}{\|\varepsilon^{(\ell)}\|} \leq \|F\|$$

$$\frac{\|\varepsilon^{(\ell+1)}\|}{\|\varepsilon^{(\ell)}\|} \approx \rho(F) = \lim_{\ell \to \infty} \frac{\|\varepsilon^{(\ell+1)}\|}{\|\varepsilon^{(\ell)}\|}$$

Number of Iterations and Spectral Radius

$$\rho(F) \approx \frac{\| \varepsilon^{(\ell+1)} \|}{\| \varepsilon^{(\ell)} \|} \qquad x + \varepsilon^{(\ell)} = r^{(\ell)} = Ax^{(\ell)} - b$$

How many to converge?

$$\begin{cases} 0.99^{\ell} = 10^{-5} \Rightarrow \sim 1{,}100 \\ 0.3^{\ell} = 10^{-5} \Rightarrow \sim 10 \end{cases}$$

Summary of Classical Iteration Schemes

- Implementations are very simple
- Error properties are very well understood
- Generally slowly converging in practical problems
- Good for simple problems

Multigrid Methods

Briggs, Henson, and McCormick et al., A Multigrid Tutorial, 2nd Ed., SIAM Press (2000). https://doi.org/10.1137/1.9780898719505

Multigrid Methods

Images from: Briggs, Henson, and McCormick et al., A Multigrid Tutorial, 2nd Ed., SIAM Press (2000).

- Logical extension to classical methods that arises from error analysis.
  - Consider "shape" of error → frequency transform

Classical methods very good at "smoothing" high-frequency errors

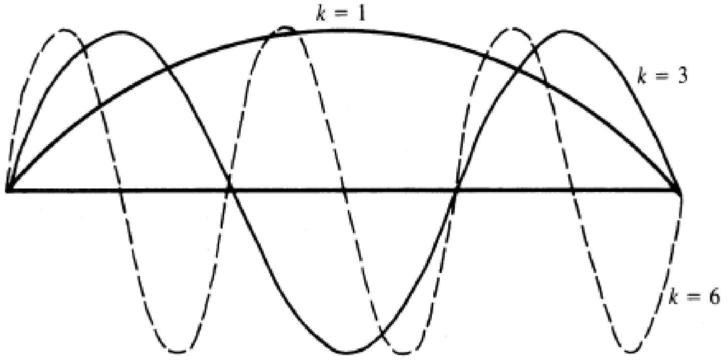


Figure 2.2: *The modes* $v_j = \sin\left(\frac{jk\pi}{n}\right)$, $0 \le j \le n$, *with wavenumbers* $k = 1, 3, 6$. *The kth mode consists of* $\frac{k}{2}$ *full sine waves on the interval.*

(a)

(b)

(c)

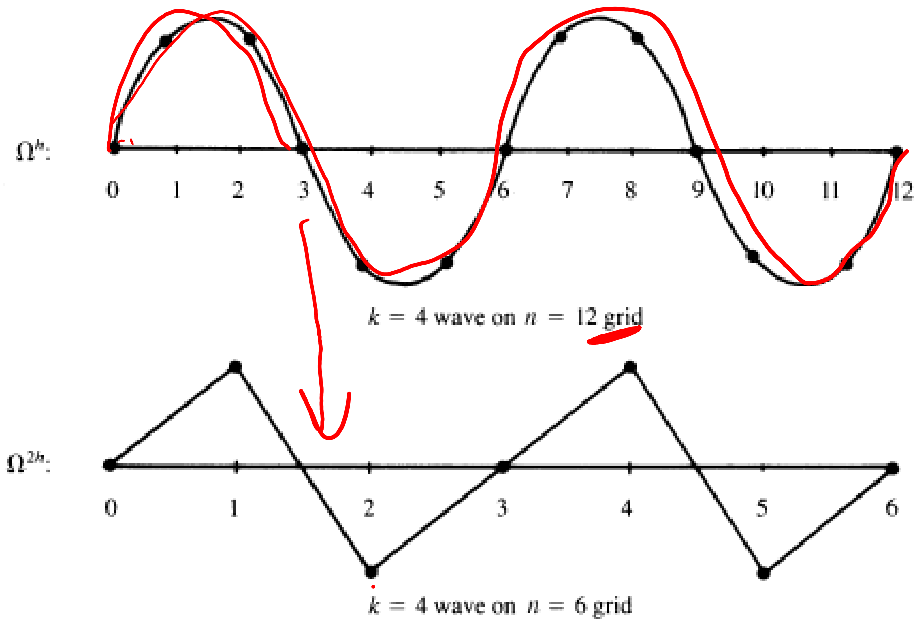
Figure 2.9: *Weighted Jacobi method with* $\omega = \frac{2}{3}$ *applied to the one-dimensional model problem with* $n = 64$ *points and with an initial guess consisting of* (a) $\mathbf{w}_3$, (b) $\mathbf{w}_{16}$, *and* (c) $(\mathbf{w}_2 + \mathbf{w}_{16})/2$. *The figures show the approximation after one iteration (left side) and after 10 iterations (right side).*

Multigrid Methods (2)

Images from: Briggs, Henson, and McCormick et al., A Multigrid Tutorial, 2nd Ed., SIAM Press (2000).

- Central idea of multigrid is to "map" errors onto coarser grids
  - A low-frequency error on a fine-grid is a high-frequency error on a coarse-grid!
- Recipe for Multigrid includes
  - How to map error from fine-grid to coarse-grid?
    - restriction operator (e.g. bi-linear average)
  - How to smooth error on each grid?
    - classical iteration scheme
  - How to correct error in fine-grid from coarse grid?
    - interpolation operator (e.g. linear interpolate)
  - How to traverse grids?

GS
J



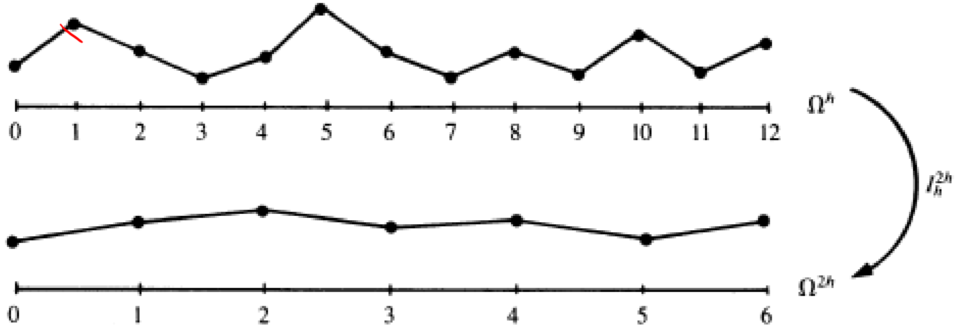
Multigrid: Restriction and Interpolation

Images from: Briggs, Henson, and McCormick et al., A Multigrid Tutorial, 2nd Ed., SIAM Press (2000).

Restriction

$$I_h^{2h}\mathbf{v}^h = \frac{1}{4}\underbrace{\begin{bmatrix} 1 & 2 & 1 & & & & \\ & & 1 & 2 & 1 & & \\ & & & & 1 & 2 & 1 \end{bmatrix}}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \end{bmatrix}_h = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}_{2h} = \mathbf{v}^{2h}$$

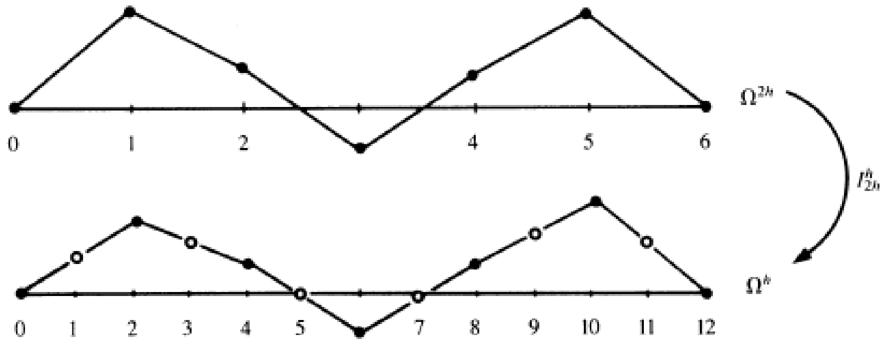
fine



coarse grid

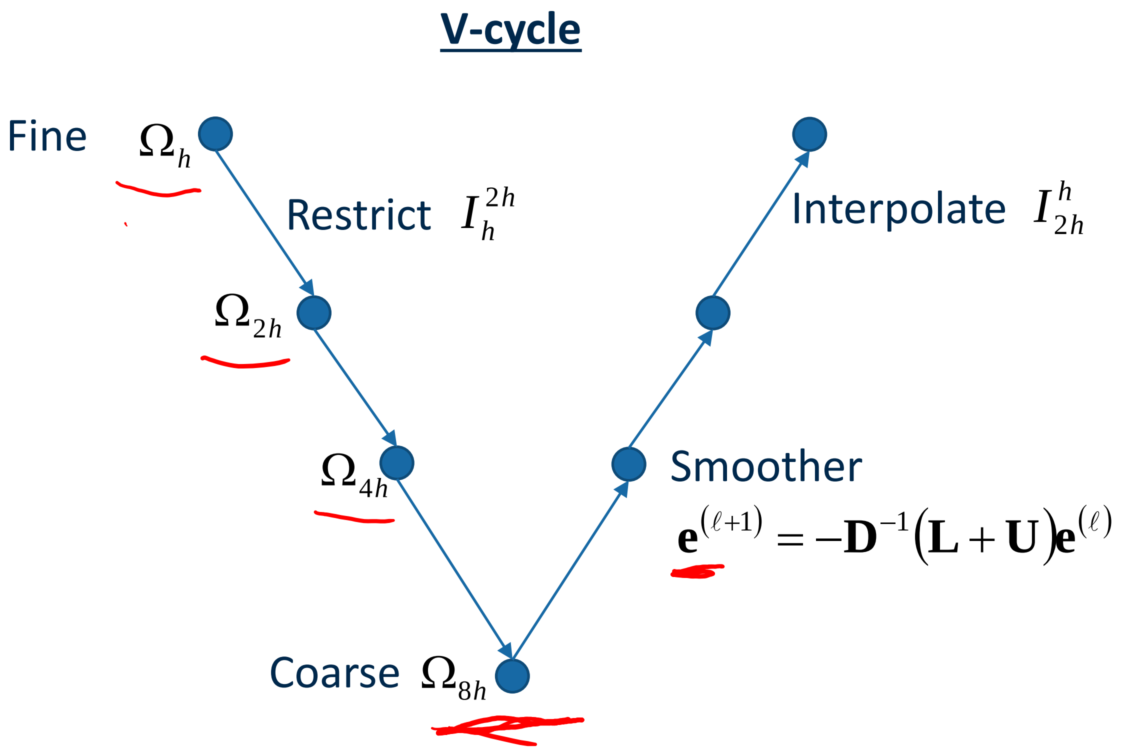
Interpolation

$$I_{2h}^h\mathbf{v}^{2h} = \frac{1}{2}\underbrace{\begin{bmatrix} 1 & & \\ 2 & & \\ 1 & 1 & \\ & 2 & \\ & 1 & 1 \\ & & 2 \\ & & 1 \end{bmatrix}}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}_{2h} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \end{bmatrix}_h = \mathbf{v}^h$$

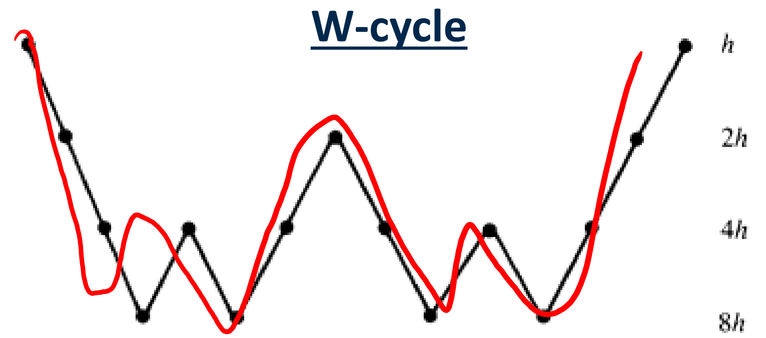


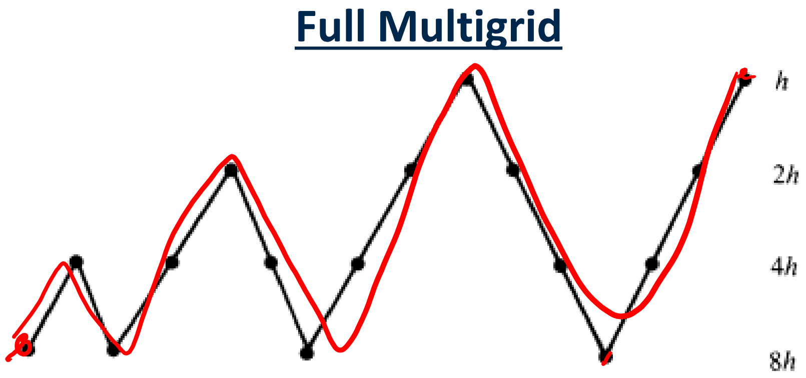
Multigrid: Traversing the Grids

Images from: Briggs, Henson, and McCormick et al., A Multigrid Tutorial, 2nd Ed., SIAM Press (2000).



$$\mathbf{e}^{(\ell+1)} = -\mathbf{D}^{-1}(\mathbf{L}+\mathbf{U})\mathbf{e}^{(\ell)} + \mathbf{D}^{-1}\mathbf{r}$$





Summary of Multigrid

- Very good for elliptic problems
- A type of fixed point iteration
  - May be analyzed via Fourier/Von Neumann Analysis for asymptotic convergence
- Builds on traditional classical fixed point iterative techniques
  - Uses same elements and adds a few more (interpolation/prolongation)
- Lots of parameters in the iteration that can be "tuned"
- Good for structured grids and finite differenced or finite volume "disc" (e.g. discretized operator is a stencil)
- Can be generalized to algebraic multi-grid (AMG)