Nov 6, 2020

$Ax = b$ $\quad x_0, x_1, \dots, x_k, \dots$ hope: $x_k \to A^{-1}b$

One key idea: only need "$y = Ax$" (not A_{ij})

Today: splitting / stationary methods

$A = M - N$ $\quad Ax = b$ $\quad (M-N)x = b \iff Mx_{k+1} = Nx_k + b$

$$x_{k+1} = M^{-1}Nx_k + M^{-1}b = Rx_k + c$$

(special case of fixed point iteration: $x_{k+1} = F(x_k)$)

Convergence analysis

$$\begin{aligned} Mx_{k+1} &= Nx_k + b \\ - \quad Mx &= Nx + b \\ M(\underbrace{x_{k+1} - x}_{\hat{e}_{k+1}}) &= N(\underbrace{x_k - x}_{\hat{e}_k}) \end{aligned}$$

$\|\hat{e}_{k+1}\| \le \|R\| \, \|\hat{e}_k\|$

$\le \|R\|^{k+1} \|\hat{e}_0\|$

Want: $\|R\| < 1$

$\hat{e}_{k+1} = R^{k+1} e_0$ $\qquad \hat{e}_{k+1} = R\hat{e}_k$ $\quad (R = M^{-1}N)$

Convergence

Any operator norm $\|\cdot\|$ s.t. $\|R\| < 1 \Rightarrow \hat{e}_k \to 0$ $x_k \to A^{-1}b$

$$\max_k |\lambda_k(R)| = \rho(R) \leq \|R\|$$

$\exists \overset{\text{op norm} \Rightarrow \text{submult}}{\|\cdot\|_*}$ s.t. $\|R\|_* \leq \rho(R) + \varepsilon$ (See Demmel Ch. 6)

$\rho(R) < 1 \Rightarrow$ choose $\|\cdot\|_*$ so that $\|R\|_* < 1$ (convergence)

$\rho(R) \geq 1 \Rightarrow$ choose $\hat{e}_0 = x_0 - x$ with $R\hat{e}_0 = \lambda \hat{e}_0$ $|\lambda| = \rho(R)$

$$\|\hat{e}_{k+1}\| = \|R^{k+1}\hat{e}_0\| = \|\lambda^{k+1}\hat{e}_0\| = \rho(R)^{k+1}\|\hat{e}_0\| \geq \|\hat{e}_0\|$$

Goals: (1) small $\|R\|$ (2) $Mx_{k+1} = Nx_k + b$ easy to solve

(1):

$M = A \quad N = 0$

$R = M^{-1}N = 0 \quad \|R\| = 0$

$Ax_{k+1} = 0x_k + b$

(2):

$M = cI \quad N = cI - A$

$$x_{k+1} = \frac{1}{c}(cI - A)x_k + \frac{1}{c}b$$

$$\boxed{x_{k+1} = x_k - \frac{1}{c}(Ax_k - b)}$$

Richardson iteration

$$A = D - L - U$$

Jacobi: $M = D \quad N = L + U \quad D x_{k+1} = (L+U) x_k + b$

"damped" $\omega < 1 \quad \omega D x_{k+1} = (L + U + (1-\omega) D) x_k + b$

Gauss-Seidel: $M = D - L, \; N = U \quad (D-L) x_{k+1} = U x_k + b$

Sweeps: one equation at a time

$$(D-L) y = U x + b \qquad \begin{matrix} x = x_k \\ y = x_{k+1} \end{matrix}$$

overwrite x with y

$$d_{11} \overset{x_1}{\cancel{y_1}} = \sum_{j=2}^{n} u_{1j} x_j + b_1$$

$$-l_{21} \overset{x_1}{\cancel{y_1}} + d_{22} \overset{x_2}{\cancel{y_2}} = \sum_{j=3} u_{2j} x_j + b_2$$

$$\vdots$$

$$-\sum_{i=1}^{k-1} l_{ki} \overset{x_i}{\cancel{y_i}} + d_{kk} \overset{x_k}{\cancel{y_k}} = \sum_{j=k+1}^{n} u_{kj} x_j + b_k$$

$$d_{kk} x_k = \sum_{i=1}^{k-1} l_{ki} x_i + \sum_{j=k+1}^{n} u_{kj} x_j + b_k$$

$$A_{kk} x_k = -\sum_{j \neq k} A_{kj} + b_k$$

Convergence

$A = M - N \qquad M x_{k+1} = N x_k + b \qquad \hat{e}_{k+1} = M^{-1} N \hat{e}_k$

Richardson

$$x_{k+1} = x_k - \frac{1}{c}(A x_k - b)$$

$$M = cI \qquad N = cI - A$$

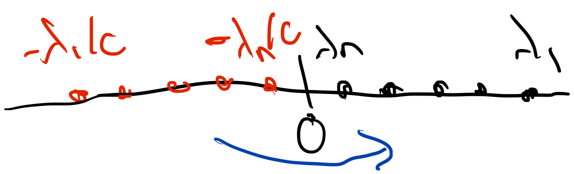
$$M^{-1}N = R = I - \frac{1}{c} A$$

$$\lambda_k(R) = 1 - \lambda_k(A)/c$$

$$c \to \infty \Rightarrow |\lambda(R)| < 1$$

$$A \text{ spd} \quad \lambda_1 \geq \dots \geq \lambda_n > 0$$

$$\|R\|_2 = \max_k \left|1 - \frac{\lambda_k}{c}\right|$$



$$\|R\|_2 = \max\left(\left|1 - \frac{\lambda_1}{c}\right|, \left|1 - \frac{\lambda_n}{c}\right|\right)$$

$$-\left(1 - \frac{\lambda_1}{c}\right) = 1 - \frac{\lambda_n}{c} \Rightarrow 2c = \lambda_1 + \lambda_n$$

$$\|R\|_2 = 1 - \frac{\lambda_n}{\frac{1}{2}(\lambda_1 + \lambda_n)} \approx 1 - 2\frac{\lambda_n}{\lambda_1} = 1 - \frac{2}{\kappa_2(A)}$$

Jacobi

$$x_{k+1} = D^{-1}(L+U) x_k + D^{-1} b$$

$$M = D, \quad N = L + U$$

$$R = D^{-1}(L+U)$$

row diag. dominant

$$\|D^{-1}(L+U)\|_\infty = \max_i \left|\frac{1}{d_{ii}}\right| \sum_{\substack{j=1 \\ j \neq i}}^{n} |A_{ij}| < 1$$

Gauss-Seidel

$$x_{k+1} = (D-L)^{-1} U x_k + (D-L)^{-1} b$$

Thm (Demmel Ch. 6)

If A row diag. dom.

$$\|R_{GS}\|_\infty \leq \|R_J\| < 1$$

Splitting as optimization

A symm. $\quad \min_x f(x) \quad f(x) = \frac{1}{2} x^T A x - x^T b$

$$\nabla f(x) = Ax - b \Rightarrow \nabla f(x) = 0 \iff Ax = b$$

Gradient descent: $x_{k+1} = x_k - \alpha \nabla f(x_k)$

$= x_k - \alpha (A x_k - b)$ Richardson! $(\alpha = \frac{1}{c})$

Coordinate descent: $\min_\alpha f(x + \alpha e_k)$

$$f(x + \alpha e_k) = \frac{1}{2}(x + \alpha e_k)^T A (x + \alpha e_k) - (x + \alpha e_k)^T b$$

$$= f(x) + \frac{1}{2} \alpha^2 A_{kk} + \alpha e_k^T A x - \alpha b_k$$

$$\frac{\partial}{\partial \alpha} = \alpha A_{kk} + e_k^T A x - b_i$$

Gauss-Seidel!

$0 =$

$$\alpha A_{kk} = - \sum_{j=1}^{n} A_{kj} x_j + b_k$$

$$\alpha (A_{kk} + x_k) = - \sum_{j \neq k} A_{kj} x_j + b_k$$