Lec 19 **Accelerating Convergence**

Lemma Consider the split $A = D + L + U$. Then $Ax = b$ may be written:

$$(D + \omega L)x = -[(\omega - 1)D + \omega U]x + \omega b. \quad (*)$$

Pf

$$Ax = b$$

$$\iff \omega A x = \omega b$$

$$\iff \omega(D + L + U)x = \omega b.$$

$$\iff \underbrace{(\omega D + \omega L)}_{\omega D + \omega L + D - D}x = -\omega U x + \omega b.$$

$$\iff (D + \omega L)x = -\omega D x + D x - \omega U x + \omega b$$

$$= -[(\omega - 1)D + \omega U]x + \omega b$$

$\boxtimes$

As before, we can solve $(*)$ iteratively via:

$$\boxed{(D + \omega L)x_{k+1} = -[(\omega - 1)D + \omega U]x_k + \omega b} \quad (**)$$

Note: This reduces to Gauss-Seidel when $\omega = 1$ (see $(**)$ on p3 of L18).

In fact, we can interpret (**) as averaging the current iterate x_k and the next iterate x_{k+1}. To see this, consider:

$$\boxed{x_{k+1} = \omega g_{GS}(x_k, x_{k+1}) + (1-\omega)x_k} \qquad \otimes$$

where (by (*b) on p3 of L18):

$$g_{GS}(x_k, x_{k+1}) = D^{-1}\left[-Lx_{k+1} - Ux_k + b\right]$$

Sub into $\otimes$, and premultiply by D, to get:

$$Dx_{k+1} = \omega\left[-Lx_{k+1} - Ux_k + b\right] + (1-\omega)Dx_k$$

Bring x_{k+1} to LHS to get:

$$(D + \omega L)x_{k+1} = \left[-\omega U + (1-\omega)D\right]x_k + \omega b$$

$$= -\left[(\omega-1)D + \omega U\right]x_k + \omega b$$

which is (**).

Next: Write RHS of (**) in terms of $(D+\omega L)x_k$:

$$\text{RHS} = (D+\omega L)x_k - \left[(\omega-1)D + \omega U + D + \omega L \right] x_k + \omega b$$

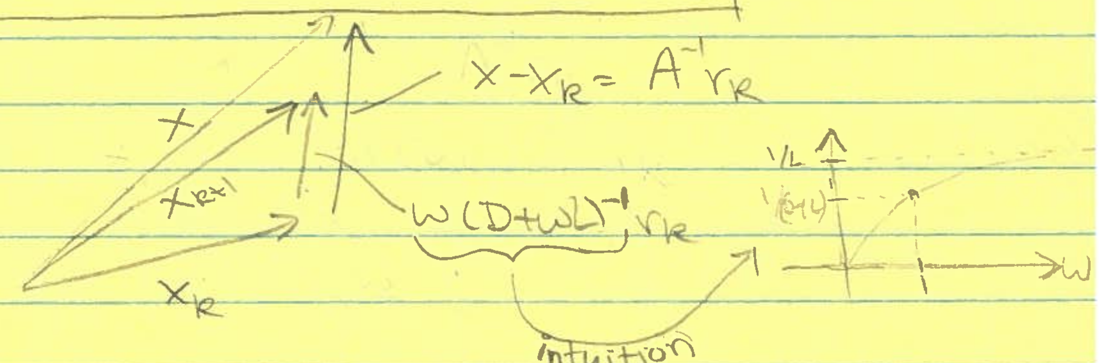
$$= (D+\omega L)x_k - [\omega D + \omega U + \omega L] x_k + \omega b$$

$$= (D+\omega L)x_k + \omega[\underbrace{b - Ax_k}_{r_k}]$$

r_k = residual @ k^{th} step
$(= A(x - x_k))$

Then:

(**) $\Rightarrow$ $$\boxed{x_{k+1} = x_k + \omega (D+\omega L)^{-1} r_k} \qquad (***)$$



We say that x_k is "relaxing" towards x and that when $\omega > 1$, x_k "over-relaxes" to x, ie. relaxes faster than for $\omega = 1$ (Gauss-Seidel). Thus (***) is a way to accelerate convergence relative to Gauss-Seidel. The method is called "Successive Over-Relaxation (SOR)".

COMPONENT FORM OF SOR

Recall (**):

$$(D + \omega L) x_{k+1} = -[(\omega-1)D + \omega U] x_k + \omega b$$

Given x_k, we want to compute x_{k+1}. This can be done by using forward substitution to solve the system of eqns above, ie.

$$\hat{A} x_{k+1} = \hat{b} \qquad (\Delta)$$

where

$$\hat{A} = D + \omega L = \text{lower triangular}$$

$$\hat{b} = -[(\omega-1)D + \omega U] x_k + \omega b$$

Forward subst. applied to (Δ) yields:

$$\hat{a}_{ii} x_i^{(k+1)} = -\sum_{j=1}^{i-1} \hat{a}_{ij} x_j^{(k+1)} + \hat{b}_i$$

In terms of D, L, U, we have:

$$a_{ii} x_i^{(k+1)} = -\sum_{j=1}^{i-1} \omega a_{ij} x_j^{(k+1)}$$

$$-\sum_{j=1}^{n} [(\omega-1) D_{ij} + \omega U_{ij}] x_j^{(k)} + \omega b_i.$$

$$(\Delta\Delta)$$

But:

$$D_{ij} = a_{ii}\delta_{ij}.$$

$$\Rightarrow \sum_{j=1}^{n} (\omega-1) D_{ij} x_j^{(k)} = (\omega-1) a_{ii} x_i^{(k)}$$

Similarly:

$$U_{ij} = \begin{cases} a_{ij} & j>i \\ 0 & j\leq i. \end{cases}$$

$$\Rightarrow \sum_{j=1}^{n} \omega U_{ij} x_j^{(k)} = \sum_{j=i+1}^{n} \omega a_{ij} x_j^{(k)}$$

Thus: $(\Delta\Delta) \Rightarrow$

$$a_{ii} x_i^{(k+1)} = -\sum_{j=1}^{i-1} \omega a_{ij} x_j^{(k+1)}$$

$$- (\omega-1) a_{ii} x_i^{(k)}$$

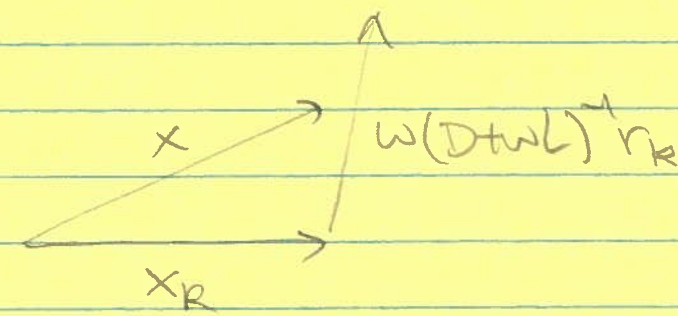
$$- \sum_{j=i+1}^{n} \omega a_{ij} x_j^{(k)}$$

$$+ \omega b_i.$$

or:

SOR $\quad x_i^{(k+1)} = (1-\omega) x_i^{(k)} + \frac{\omega}{a_{ii}} \left[b_i - \sum_{j<i} a_{ij} x_j^{(k+1)} - \sum_{j>i} a_{ij} x_j^{(k)} \right]$

One could imagine that for large ω, the vector $\omega(D+\omega L)^{-1}r_k$ might "over-shoot" the exact solution.



which could destroy the convergence.

In 1947, Ostrowski proved:

<u>Thm 1</u> If A is positive definite and $0<\omega<2$, then SOR converges for any starting point x_0.

But which value of ω in that range yields fastest convergence?

__Thm 2__ If A is positive definite and tridiagonal, then $\rho_{GS} = \rho_J^2 < 1$ and the optimal value of ω is

$$\omega^* = \frac{2}{1+\sqrt{1-\rho_J^2}}$$

for which

$$\rho_{\omega^*} = \omega^* - 1.$$

Example $A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$

Evidently, A is tridiagonal.

A is positive definite:

$$[x \ y]\begin{bmatrix}2 & -1\\ -1 & 2\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}$$

$$= [x \ y]\begin{bmatrix}2x-y\\ -x+2y\end{bmatrix}$$

$$= x(2x-y) + y(-x+2y)$$

$$= 2x^2 - xy - xy + 2y^2$$

$$= x^2 + y^2 + [x^2 - 2xy + y^2]$$

$$= x^2 + y^2 + (x-y)^2$$

$$\geq 0 \quad \forall x, y$$

Thus, the assumptions of Thm 2 hold

$x_{k+1} = Bx_k + c$ where

Jacobi: $B = S^{-1}T = -D^{-1}(L+U)$

$$= -\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$$

$$= -\begin{bmatrix} 0 & -\frac{1}{2} \\ -\frac{1}{2} & 0 \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}$$

$$\det(B - \lambda I) = 0 \Rightarrow \det\begin{pmatrix} -\lambda & \frac{1}{2} \\ \frac{1}{2} & -\lambda \end{pmatrix} = 0.$$

$$\Rightarrow \lambda^2 - \left(\tfrac{1}{2}\right)^2 = 0 \Rightarrow \lambda = \pm\tfrac{1}{2} \Rightarrow \rho_J = \tfrac{1}{2}.$$

Gauss-Seidel: $B = S^{-1}T = -(L+D)^{-1}U$

Now: $L+D = \begin{bmatrix} 2 & 0 \\ -1 & 2 \end{bmatrix}$

$$\Rightarrow (L+D)^{-1} = \begin{bmatrix} \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{2} \end{bmatrix}$$

$$\therefore B = -\begin{bmatrix} \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{2} \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} = -\begin{bmatrix} 0 & -\frac{1}{2} \\ 0 & -\frac{1}{4} \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{2} \\ 0 & \frac{1}{4} \end{bmatrix}$$

$$\therefore \det(B-\lambda I) = \begin{vmatrix} -\lambda & \frac{1}{2} \\ 0 & \frac{1}{4}-\lambda \end{vmatrix}$$

$$\Rightarrow -\lambda\left(\tfrac{1}{4}-\lambda\right) = 0$$

$$\Rightarrow \lambda = 0 \text{ or } 1/4 \Rightarrow \rho_{GS} = \tfrac{1}{4}.$$

Note: $\rho_{GS} = \rho_J^2$, as promised by Thm 2.

Thm 2 $\Rightarrow$ Optimal ω is:

$$\omega^* = \frac{2}{1+\sqrt{1-\rho_J^2}}$$

$$= \frac{2}{1+\sqrt{1-(\frac{1}{2})^2}}$$

$$= \frac{2}{1+\sqrt{3/4}}$$

$$= \frac{2}{1+\sqrt{3}/2} = \frac{4}{2+\sqrt{3}}$$

$$= \frac{4}{2+\sqrt{3}}\,\frac{2-\sqrt{3}}{2-\sqrt{3}} = \frac{8-4\sqrt{3}}{4-3}$$

$$= 8-4\sqrt{3}.$$

Thm 2 $\Rightarrow$ $\rho_{\omega^*} = \omega^* - 1 = 7-4\sqrt{3}. \cong 0.08$

Thus for the positive definite, tridiagonal matrix

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

we have:

$$\rho_{\omega^*} = 0.08 \ll \rho_{GS} = 0.25 < \rho_J = \tfrac{1}{2} < 1.$$

All three methods converge, but SOR is by far the fastest.