October 17, 2013

$\boxed{\text{CG:}}$ $Ax = b$, A is SPD

$r_0 = b - Ax_0$, $p_0 = r_0$ orthogonal to residual (steepest descent).

for $k = 0, 1, \ldots$

$$\alpha_k = \frac{(r_k, r_k)}{(Ar_k, p_k)}$$

step size

search direction

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

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

$$\beta_k = \frac{(r_{k+1}, r_{k+1})}{(r_k, r_k)}$$

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

end

recall : $\phi(x) = \frac{1}{2} x^T A x - x^T b$

$\min \phi(x)$

$\nabla \phi(x) = Ax - b = -r$

so $-\nabla \phi(x) = r$ is the direction of steepest descent

$x_{k+1} = x_k - \alpha_k \nabla \phi(x_k)$

↳ gradient vector is always orthogonal to the contour lines

CG:

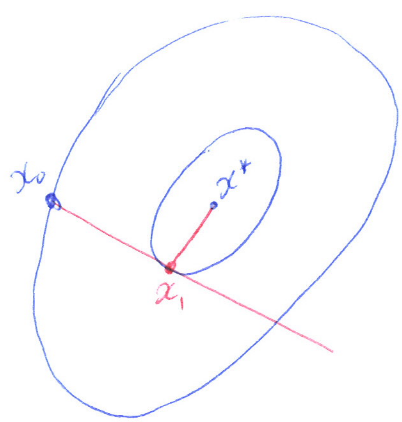
Choose a different (not steepest) direction p_k:

$$x_{k+1} = x_k + \alpha p_k$$

choose α to minimize

$$\phi(x_k + \alpha p_k)$$

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



$r_1 = b - Ax_1$

$\quad = A(x^* - x_1)$

so

$$p_0^T A \underbrace{(x^* - x_1)}_{\alpha p_1} = 0 \longrightarrow p_0^T A p_1 = 0 \quad p_0, p_1 \text{ are } A\text{-conjugate}$$

Could also derive CG from a pure linear algebra point of view:

(Nice geometric description: LeVeque's book, painless CG)

↳ this description: Demmel, Golub + Van Loan, Saad

$$\begin{cases} T_{kk} z = \rho e_1 \\ x_k = x_0 + Q_k z \end{cases}$$

$$x_{k+1} = x_0 + Q_k T_{kk}^{-1} (\rho e_1)$$

$$T_{kk} = L_k D_k L_k^T$$

$$= L_k U_k = \begin{pmatrix} 1 & & 0 \\ \alpha_2 & \ddots & \\ 0 & \ddots \alpha_k & 1 \end{pmatrix} \begin{pmatrix} \eta_1 & \beta_2 & & \\ & \ddots & \ddots & \\ & & & \beta_k \\ & & & \eta_k \end{pmatrix}$$

$$P_k = Q_k U_k^{-1}$$

What does $P_k^T A P_k$ look like?

$$\underbrace{P_k^T A P_k}_{\text{symmetric}} = (Q_k U_k^{-1})^T A (Q_k U_k^{-1})$$

$$= U_k^{-T} \underbrace{Q_k^T A Q_k}_{T_{kk} = L_k U_k} U_k^{-1}$$

$$= \underbrace{U_k^{-T} L_k}_{\text{lower triangular}} = D_k$$

note: $P_k = Q_k U_k^{-1}$

$\rightarrow P_k U_k = Q_k$

$\rightarrow P_k \begin{pmatrix} \eta_1 & \beta_2 & & \\ & \ddots & \ddots & \\ & & & \beta_k \\ & & & \eta_k \end{pmatrix} = Q_k$

$\rightarrow q_k = \beta_k p_{k-1} + \eta_k p_k$

$\rightarrow p_k = \frac{1}{\eta_k}(q_k - \beta_k p_{k-1})$

short recurrance.

Convergence:

- clustering of eigenvalues → primary factor
- $\|e_k\|_A \leq 2 \left(\frac{\sqrt{\kappa_2(A)} - 1}{\sqrt{\kappa_2(A)} + 1} \right)^k \|e_0\|_A$

$\kappa_2(A)$: condition number

↳ good if no specific distribution is known

Symm: optimality + short recurrences : CG, MINRES

non-symm: optimality + long " : GMRES, FOM ← restarted GMRES comprimises optimality for better storage (a comprimise)

nonoptimal + short : BICG

$\boxed{\text{Multigrid:}}$

- powerpoint available on website
- after a few iterations of damped Jacobi or Gauss-Seidel, error is still large, but smoother
- So, non-smooth (oscillatory) modes are taken care of nicely by these methods

Note: damped J is a good smoother with $\omega < 1$ ($\omega = 2/3$ in 1D eg).
Jacobi is a bad smoother ($\omega = 1$)