Dec 2, 2020

Last time: $f(x) = \frac{1}{2}x^T A x - x^T b \quad \nabla f(x) = Ax - b$

GD: $x_{k+1} = x_k - \alpha_k g_k \quad g_k = \nabla f(x_k) \qquad \left(1 - \frac{1}{\kappa_2(A)}\right)^{k/2}$

CG: $x_{k+1} = \operatorname{argmin} f(x) \text{ s.t. } x_k \in \mathcal{K}_k \qquad \left(1 - \frac{1}{\sqrt{\kappa_2(A)}}\right)^{k/2}$

How is CG a gradient method? residual $\iff$ -gradient

CG + Lanczos

$q_0 = 0, \ \beta_0 = \|b\|_2, \ q_1 = b/\beta_0, \ x_0 = 0, \ w_0 = 0$

for $k = 1, 2, \ldots$

- $\alpha_k = q_k^T A q_k$
- search direction: $w_k = q_k - \ell_{k-1} w_{k-1}$
- CG step: $x_k = x_{k-1} + \eta_k w_k$
- $\beta_k q_{k+1} = A q_k - \alpha_k q_k - \beta_{k-1} q_{k-1}$

$Q_k^T A Q_k = T_k$

$T_k = L_k D_k L_k^T$

$W_k L_k^T = Q_k$

Claim 1: $g_k = -r_k \propto q_{k+1}$

Proof: $g_k = Ax_k - b$

$x_k, b \in K_k(A, b)$

$g_k \in K_{k+1}(A, b)$

$-g_k = r_k \perp K_k(A, b)$ (HW 6)

$g_k \propto q_{k+1}$

Corollary 2:
residuals/gradients orthog

f(x)

Claim 3: w_k are "A-orthogonal"

- $w_i^T A w_j = 0 \quad i \neq j$

Proof: $W_k^T A W_k$

$= L_k^{-1} Q_k^T A Q_k L_k^{-T}$

$= L_k^{-1} T_k L_k^{-T}$

$= L_k^{-1} L_k D_k L_k^T L_k^{-T}$

$= D_k$

"linear"