Nov 30, 2020

Krylov subspace $K_k(A,b) = \{b, Ab, \dots, A^{k-1}b\}$

$\|w\|_A = \sqrt{w^T A w}$

$A \in \mathbb{R}^{n\times n}$ SPD $\Rightarrow$ CG $\quad Ax=b$

$$x_k = \arg\min_x \underbrace{\tfrac{1}{2}x^T A x - b^T x}_{f(x)} \iff x_k = \arg\min_x \|x - \underbrace{A^{-1}b}_{x^*}\|_A$$

s.t. $x \in K_k(A,b)$ $\qquad$ s.t. $x_k \in K_k$

First approach: GD on $f(x)$ $\quad x_0, x_1, \dots, x_k, \dots$

$\nabla f(x) = Ax - b$, $\quad g_k = \nabla f(x_k)$ $\quad$ GD: $x_{k+1} = x_k - \alpha_k g_k$

$$f(x_k - \alpha_k g_k) = \tfrac{1}{2}(x_k - \alpha_k g_k)^T A (x_k - \alpha_k g_k) - b^T(x_k - \alpha_k g_k)$$

$$\star \qquad = f(x_k) + \tfrac{1}{2}\alpha_k^2 g_k^T A g_k - \alpha_k g_k^T A x_k + \alpha_k b^T g_k$$

$$\frac{\partial}{\partial \alpha_k} = 0 \Rightarrow \alpha_k g_k^T A g_k - x_k^T A g_k + b^T g_k$$

$$\Rightarrow \alpha_k^* = (\underbrace{Ax_k - b}_{g_k})^T g_k / g_k^T A g_k = \frac{g_k^T g_k}{g_k^T A g_k}$$

$$f(x_{k+1}) = f(x_k) + \frac{1}{2}\left(\frac{g_k^T g_k}{g_k^T \cancel{A g_k}}\right)^2 g_k^T \cancel{A} g_k - \frac{g_k^T g_k}{g_k^T A g_k} g_k^T g_k$$

$$☆ = f(x) + \frac{g_k^T g_k}{g_k^T A g_k}\left[\frac{1}{2} g_k^T g_k - g_k^T g_k\right] = -\frac{1}{2} g_k^T g_k \qquad 2f(x_k)$$

$$\underline{g_k^T A^{-1} g_k} = (Ax_k - b)^T A^{-1}(Ax_k - b) = (Ax_k - b)^T (x_k - A^{-1}b) = \overbrace{x_k^T A x_k - 2b^T x_k}^{2f(x_k)} + b^T A^{-1} b$$

$$\underline{f(A^{-1}b)} = \frac{1}{2}(A^{-1}b)^T A (A^{-1}b) - b^T A^{-1} b = -\frac{1}{2} b^T A^{-1} b$$

$$★\ f(x_{k+1}) = f(x_k) - \underbrace{\frac{g_k^T g_k}{g_k^T A g_k}\,\frac{g_k^T g_k}{g_k^T A^{-1} g_k}}_{c_k}\left[f(x_k) + \underbrace{\frac{1}{2} b^T A^{-1} b}_{= -f(A^{-1}b)}\right]$$

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

$$\underbrace{f(x_{k+1}) - f(A^{-1}b)}_{\frac{1}{2}\|e_{k+1}\|_A^2} = f(x_k) - f(A^{-1}b) - c_k\left(f(x_k) - f(A^{-1}b)\right)$$

$$\frac{1}{2}\|e_{k+1}\|_A^2 = \frac{1}{2}\|e_k\|_A^2 - c_k \frac{1}{2}\|e_k\|_A^2$$

$$\|e_{k+1}\|_A = \sqrt{1 - c_k}\,\|e_k\|_A \qquad \left|\frac{1}{c_k}\right| = \left|\frac{g_k^T A g_k}{g_k^T g_k}\,\frac{g_k^T A^{-1} g_k}{g_k^T g_k}\right| \le \kappa_2(A)$$

GD: (with line search) $\|e_{k+1}\|_A \le \left(1 - \frac{1}{\kappa_2(A)}\right)^{k/2} \|e_0\|_A$

$$x_k = \arg\min_x \|A^{-1}b - x\|_A^2 \quad \text{s.t.} \quad x \in K_k(A, b)$$

$x = \sum_{j=0}^{k-1} (A^j b) c_j$

$$= \arg\min_{p_k} \|A^{-1}b - p_k(A) b\|_A^2 \quad \text{s.t.} \quad p_k \in P_{k-1}$$

$b = Ax_*$

$$= \| x^* - p_k(A) \cdot Ax^* \|_A^2$$

$$= \arg\min_{q_k} \| q_k(A) x^* \|_A^2 \quad q_k \in P_k \quad q_k(0) = 1$$

$A = Q \Lambda Q^T$

$$x_*^T Q q_k(\Lambda) \cancel{Q^T Q} \Lambda \cancel{Q^T Q} q_k(\Lambda) Q^T x_*$$

$y = Q^T x_*$

$$y^T q_k(\Lambda) \Lambda q_k(\Lambda) y$$

$$= \sum y_i^2 \lambda_i q_k(\lambda_i)^2$$

$$\le \left(\max_{\lambda_i} q_k(\lambda_i)^2\right) \sum y_i^2 \lambda_i$$

$= \text{error } k = 0$

$q_0(z) = 1$

$$\|e_k\|_A^2 \le \left(\min_{\substack{q_k \in P_k \\ q_k(0) = 1}} \max_{\lambda_i} q_k(\lambda_i)^2\right) \|e_0\|_A^2$$

$$\min_{q_k \in P_k} \max_i \left[q_k(\lambda_i)\right]^2$$

s.t. $q_k(0) = 1$

Chebyshev polynomials

$$T_0(z) = 1$$

$$T_1(z) = z$$

$$T_{n+1}(z) = 2z\,T_n(z) - T_{n-1}(z)$$

fact:

$$T_j(z) = \cos(j \cdot \arccos(z)) \quad |z| \le 1$$

$$|T_j(z)| \le 1 \quad |z| \le 1$$

$$q_k(x) = \frac{T_k\left(\gamma - \frac{2x}{\lambda_1 - \lambda_n}\right)}{T_k(\gamma)} \qquad \gamma = \frac{\lambda_n + \lambda_1}{\lambda_1 - \lambda_n}$$

$$q_k(0) = T_k(\gamma) / T_k(\gamma) = 1 \checkmark$$

$$x \in [\lambda_1, \lambda_n] \Rightarrow -1 \le \gamma - \frac{2x}{\lambda_1 - \lambda_n} \le 1$$

numerator bounded by 1 when $x \in [\lambda_1, \lambda_n]$

How big is $T_k(\gamma)$ or small is $1/T_k(\gamma)$?

Fact: $T_k(z) = \frac{1}{2}\left[\left(z+\sqrt{z^2-1}\right)^k + \left(z+\sqrt{z^2-1}\right)^{-k}\right]$

$|z| > 1$

$$\frac{\lambda_1+\lambda_n}{\lambda_1-\lambda_n} = \frac{\lambda_1/\lambda_n+1}{\lambda_1/\lambda_n-1} = \frac{\kappa_2(A)+1}{\kappa_2(A)-1}$$

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

$$1/T_k(\gamma) \sim \left(\frac{\sqrt{c}+1}{\sqrt{c}-1}\right)^k = \left(1-\frac{2}{\sqrt{c}-1}\right)^k$$

CG: $\|e_k\|_A \leq \left(1-\frac{2}{\sqrt{\kappa_2(A)}-1}\right)^{k/2}\|e_0\|_A$