Lec 25.

Convergence of SD.

$$x_k \longrightarrow x_{k+1}$$

$$\underbrace{\alpha_k}_{\text{step size}} \underbrace{P_k}_{\text{direction}} = r_k = b - Ax_k$$

$$\varphi(x_{k+1}) = \frac{1}{2} x_{k+1}^T A x_{k+1} - x_{k+1}^T \underset{\substack{\| \\ Ax_*}}{b}$$

$$\inf_{\alpha} \varphi(x_k + \alpha_k p_k)$$

$$\varphi(x_{k+1}) + \frac{1}{2} x_*^T A x_*$$

$$= \frac{1}{2} (x_{k+1} - x_*)^T A (x_{k+1} - x_*)$$

$$:= \frac{1}{2} \| x_{k+1} - x_* \|_A^2 \qquad A > 0.$$

$$\|x_{k+1} - x_*\|_A^2$$

$$= \|x_k + \alpha_k (\overset{A x_*}{\overset{\parallel}{b}} - A x_k) - x_*\|_A^2$$

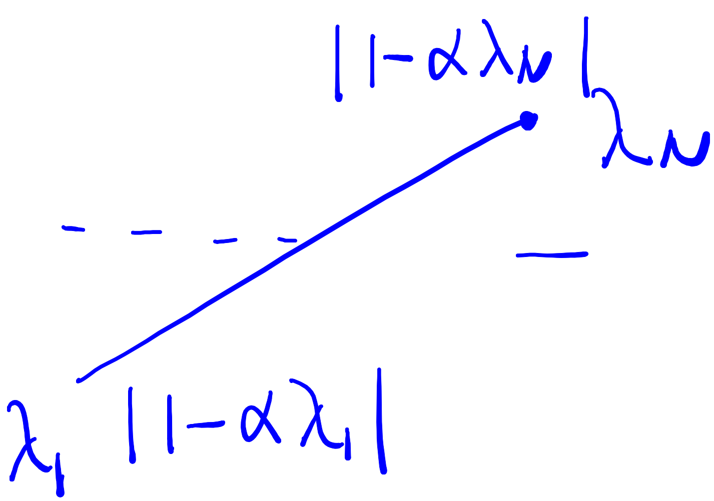
$$= \|(I - \alpha_k A)(x_k - x_*)\|_A^2$$

$$\left(= (x_k - x_*)^T (I - \alpha_k A)^2 A \, (x_k - x_*) \right)$$

$$\leq \left(\max_{\lambda_1 \leq \lambda \leq \lambda_N} |1 - \alpha_k \lambda|^2 \right) \| x_k - x_* \|_A^2$$

SD :

$$\min_{\alpha} \max_{\lambda_1 \leq \lambda \leq \lambda_N} |1 - \alpha \lambda|^2$$



$$1 - \alpha \lambda_1 = \alpha \lambda_N - 1$$

$$\Rightarrow \alpha = \frac{2}{\lambda_1 + \lambda_N}$$

$$\max_{\lambda_1 \le \lambda \le \lambda_N} |1 - \alpha \lambda|^2$$

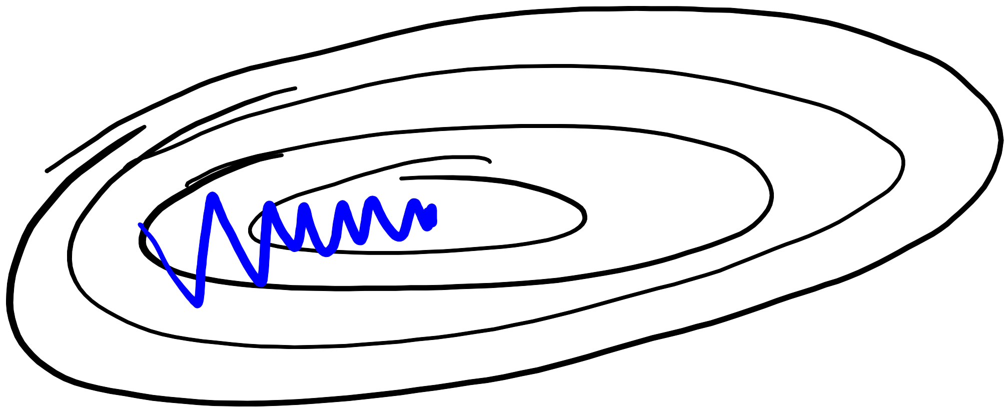
$$= \left(\frac{\lambda_N - \lambda_1}{\lambda_N + \lambda_1} \right)^2 = \left(\frac{\frac{\lambda_N}{\lambda_1} - 1}{\frac{\lambda_N}{\lambda_1} + 1} \right)^2 := \left(\frac{K - 1}{K + 1} \right)^2$$

$$\|x_{k+1} - x_*\|_A \le \left(\frac{\kappa-1}{\kappa+1}\right) \|x_k - x_*\|_A$$

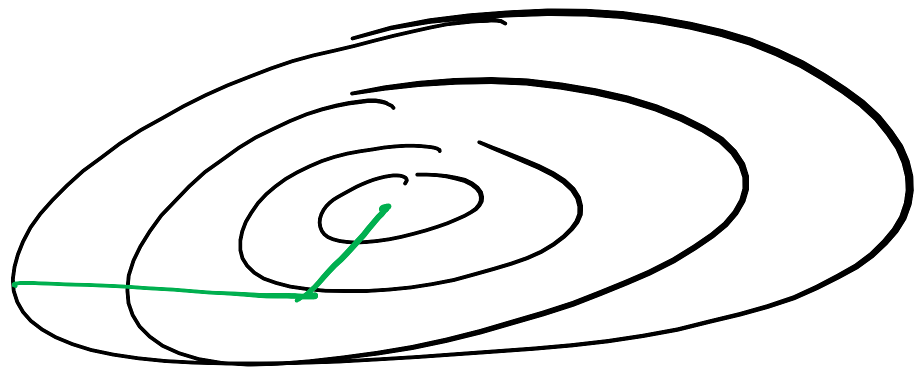
$$\le \left(\frac{\kappa-1}{\kappa+1}\right)^{k+1} \|x_0 - x_*\|_A$$

$$< \epsilon$$

$$\Rightarrow \quad k > C \kappa \log(1/\epsilon) .$$



SD



① 2D

CG

② arbitrary large

$$\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^{k}$$

$$k > C\sqrt{\kappa}\log(1/\epsilon).$$

$$\inf_{\alpha} \varphi(x_k + \alpha p_k)$$

$$\parallel$$

$$\frac{1}{2}(x_k + \alpha p_k)^T A (x_k + \alpha p_k)$$

$$- b^T (x_k + \alpha p_k)$$

$$\Rightarrow \alpha = \frac{p_k^T r_k}{p_k^T A p_k}$$

2D.

1 st step: SD.

$x_0 . \Rightarrow r_0 . \Rightarrow p_0 = r_0$

$\alpha_0 . \Rightarrow X_1$

2nd step. $p_1^T A p_0 = 0$. A-ortho to p_0.

(in 2D) uniquely fixes p_1 (up to scaling)

$\Rightarrow \alpha_1 \Rightarrow x_2 = x_* \Leftrightarrow r_2 = 0$.

$\Leftrightarrow p_0^T r_2 = 0, \; p_1^T r_2 = 0$.

Now compute.

$$r_2 = b - Ax_2 = b - A(x_1 + \alpha_1 p_1)$$

$$= r_1 - \alpha_1 A p_1$$

$$= b - A(x_0 + \alpha_0 r_0) - \alpha_1 A p_1$$

$$= (I - \alpha_0 A) r_0 - \alpha_1 A p_1$$

$$p_0^T r_2 = \left(r_0^T r_0 - \cancel{\alpha_0 r_0^T A r_0}\right) - \alpha_1 \cancel{p_0^T A p_1}$$

$$p_1^T r_2 = \underbrace{\left(p_1^T r_0 - \alpha_1 p_1^T A p_1\right)}_{\shortparallel} - \alpha_0 \cancel{p_1^T A p_0}$$

$$p_1^T(r_0 - r_1) = p_1^T\left(r_0 - b + A(x_0 + \alpha_0 r_0)\right)$$

$$= p_1^T A p_0 \cdot (+\alpha_0) = 0.$$

$$\Rightarrow \quad r_2 = 0 \qquad \square.$$

Generally. $\{x_k, p_k, r_k\}$

$$
\begin{cases}
x_{k+1} = x_k + \alpha_k p_k, \\
r_{k+1} = b - A x_{k+1}. \\
p_{k+1} = r_{k+1} + \beta_k p_k
\end{cases}
$$

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

(minimization)

$$\beta_k = \frac{p_k^T A r_{k+1}}{p_k^T A p_k}$$

(A-orthogonality)

P_{k+1} : A-ortho to P_k.

(non-trivial fact):

$P_{k+1} \perp_A \operatorname{span}\{P_0, \cdots, P_k\}$.