Expanding Subspace Theorem.

Result: Say CDM $\to \{x_k\}$ starting from x_0. (minimize ϕ) then

① $r_k^T p_i = 0$ for $i \in [0, k-1]$. ✓

② In an affine space $\{x \mid x_0 + \text{span}\{p_0, \dots, p_{k-1}\}\}$
x_k is the minimizer of $\phi(x)$. ?

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

Let's say $\hat{x} = x_0 + \sum_{i=0}^{k-1} \sigma_i p_i$. We want: σ's s.t $\hat{x}$ minimizes $\phi(x)$.

$$\frac{\partial}{\partial \sigma_i} \phi(x_0 + \Sigma \sigma_i p_i) = 0 \quad \text{for } i \in [0, k-1]$$

$\phi \to$ sy pos. def. $\Rightarrow$ strictly convex $\Rightarrow$ unique minimizer $\hat{}$

$$= \nabla\phi(x_0 + \Sigma \sigma_i p_i)^T p_i = 0 \qquad i \in [0, k-1]$$

$$\Rightarrow r(\hat{x})^T p_i = 0$$

This is actually an iff statement!

1) If $\hat{x}$ is a minimizer, then $r(\hat{x})^T p_i = 0$ for $i \in [0, k-1]$.

2) If $\boxed{r(\hat{x})^T p_i = 0}$ for $i \in [0, k-1]$ then

$\Rightarrow \quad \nabla\phi(x_0 + \sum \sigma_i p_i)^T p_i = 0$ " "

$= \frac{\partial \phi(x_0 + \sum \sigma_i p_i)}{\partial \sigma_i}^T p_i = 0$ ''

$\Rightarrow$ $\hat{x}$ is a stationary point

$\Rightarrow$ $\hat{x}$ is the minimizer.

First result by induction.

<u>base</u> $\quad k = 1 \quad \Rightarrow \quad r_1^T p_0 \stackrel{?}{=} 0$

$x_1 = x_0 + \alpha_0 p_0 \quad = \alpha_0$ is an exact min

$$\Rightarrow \quad \frac{d}{d\alpha}\phi(x_0 + \alpha p_0) = \underbrace{\nabla\phi(x_0 + \alpha p_0)}^T p_0 = 0 \text{ at } \alpha = \alpha_0$$

$$\underline{r(x_1)^T p_0 = 0}.$$

<u>Hypothesis</u> : Assume the result for $k-1$.

$$\boxed{\Rightarrow \quad r_{k-1}^T p_i = 0 \text{ for } i \in [0, k-2]}$$

$\rightarrow \quad r_k = A x_k - b$

$r_{k+1} - r_k = A(x_{k+1} - x_k) = \alpha_k A p_k$

$k-1$

$r_k = r_{k-1} + \alpha_{k-1} A p_{k-1}$

left multiply by p_i^T

$$p_i^T r_k = \underbrace{p_i^T r_{k-1}}_{\substack{\| \\ 0 \text{ by} \\ \text{ind hyp.}}} + \alpha_{k-1} \underbrace{p_i^T A p_{k-1}}_{\substack{\| \\ 0 \\ \text{by conjugacy.}}}, \quad i \in [0, k-2]$$

$$r_k^T p_i = 0 \quad \text{for } i \in [0, k-2]$$

left multiply by p_{k-1}^T

$$p_{k-1}^T r_k = \underline{p_{k-1}^T r_{k-1} + \alpha_{k-1} \underline{p_{k-1}^T A p_{k-1}}}$$

recall: $\alpha_k = -\dfrac{r_k^T p_k}{p_k^T A p_k}$ $= 0$ by defn of α.

$$\boxed{\Rightarrow \quad r_k^T p_i = 0 \quad \text{for } i \in [0, k-1]}$$

Complete proof by induction.

— x —

Conjugate directions: eigvectors or Gram-Schmidt

both $O(n^3)$ X

Conjugate Gradient Method

$p_k \to p_{k-1}$ AND $-r_k$ ONLY.

$$p_k = -r_k + \beta_k p_{k-1}$$

1) $$p_{k-1}^T A p_k = 0 = -p_{k-1}^T A r_k + \beta_k p_{k-1}^T A p_{k-1}$$

$$\Rightarrow \quad \beta_k = \frac{r_k^T A p_{k-1}}{p_{k-1}^T A p_{k-1}}$$

2) $p_0 = -r_0$ initial choice.

3) $x_{k+1} = x_k + \alpha_k p_k$, α_k exact min.

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

(2 matrix-vector products
4 vector-vector products.)

→ Some more algebra: $r_{k+1} - r_k = \alpha_k A p_k$

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

$p_k^T A p_k$.

2) $$\beta_{k+1} = \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k}$$

(1 matrix-vector product
3 vector-vector products.) To compute.

<u>To store:</u> $\quad p_k = -r_k + \beta_k p_{k-1}$

1) To get $p_k \rightarrow r_k, p_{k-1}, r_{k-1}$

2) To $x_{k+1} \rightarrow x_k, p_k$

[Need to store only k, $k-1$]

Only need to compute Ap_k.

Rate of convergence. (CG)

$\rightarrow$ Linear rate of conv.

[A has $\lambda_1 \leq \lambda_2 \cdots \leq \lambda_n \Rightarrow \kappa(A) = \dfrac{\lambda_n}{\lambda_1}$]

1) If I know all eigenvalues:

$$\| x_{k+1} - x^* \|_A \leq \left(\lambda_{n-k} - \lambda_1 \right) \| x_0 - x^* \|_A$$

$$\lambda_{n-k} + \lambda_1$$

2) If I know only κ

$$\| x_{k+1} - x^* \| \leq 2 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^k \| x_0 - x^* \|_A$$

↳ faster than SD $\sqrt{\kappa}$ instead κ