Usually $\underline{\phi(x) = x^T A x}$, where A is $\underline{\underline{\text{sym}}}$ pos def.

Assuming A is <u>not</u> sym. But it is PD.

$$A = \underbrace{\frac{A + A^T}{2}}_{B} + \underbrace{\frac{A - A^T}{2}}_{C}$$

$B \rightarrow ?$ $B^T = B$
sym

$C \rightarrow C^T = -C$
skew sym

$$\underbrace{x^T A x}_{\text{scalar}} = x^T B x + x^T C x$$

scalar $\xrightarrow{T}$ $(x^T A x)^T = \underbrace{x^T A x}$ $= (Ax)^T x = \underbrace{x^T A^T x}$

$$\cancel{x^T B x} + x^T C x = \cancel{x^T B^T x} + x^T C^T x \Rightarrow 2x^T C x = 0$$

$$\Rightarrow x^T A x = x^T B x.$$

Conjugate Gradient Methods.

Ch 5 of NW.

↳ System of eqns $\underbrace{Ax = b}$

eg. 1 → $\phi_1(x) = \|Ax - b\|^2 = x^T A^T A x + \dots$

eg. 2 → $\phi_2(x) = \frac{1}{2} x^T A x - b^T x$

→ $\nabla \phi_1(x) = Ax - b$

when $\|\nabla \phi_1(x^*)\| = 0$, x^* is a stationary pt

$\Rightarrow Ax^* = b$

In $\phi_1 \to$ matrix is $A^T A \to K^2$ ✗ | A can be anything

In $\phi_2 \to$ matrix is $A \to K$ ✓ | A has to be PD.

Higher is K, worse is progress/accuracy.

Here on, assume A is sym P.D.

$$\phi(x) = \underbrace{\frac{1}{2} x^T A x - b^T x}_{} \Rightarrow \nabla \phi(x) = Ax - b$$

$= r(x)$

gradient/residual

$$A = \underbrace{U}_{\text{Cols are eig vecs}} \overbrace{\Lambda}^{\text{diag matrix}} U^T$$

Cols are eig vecs

diag matrix.

$$AU = U\Lambda \underbrace{U^T U}_{I} = \begin{bmatrix} u_1\lambda_1 & u_2\lambda_2 & \cdots \\ | & | & \end{bmatrix}$$

$$\begin{bmatrix} u_1 & u_2 & u_n \\ | & | & | \end{bmatrix}$$

Properties → ① $\underline{U^T = U^{-1}}$ $\quad U^T U \rightarrow$

$$\begin{bmatrix} \text{---} \\ \text{---} \\ \text{---} \end{bmatrix} \begin{bmatrix} | & | & | & | \end{bmatrix} = \begin{bmatrix} I \end{bmatrix}$$

② $\|Uq\| = \|q\|$

③ $(Uq, Ut) = (q, t)$

Visualize quadratic forms

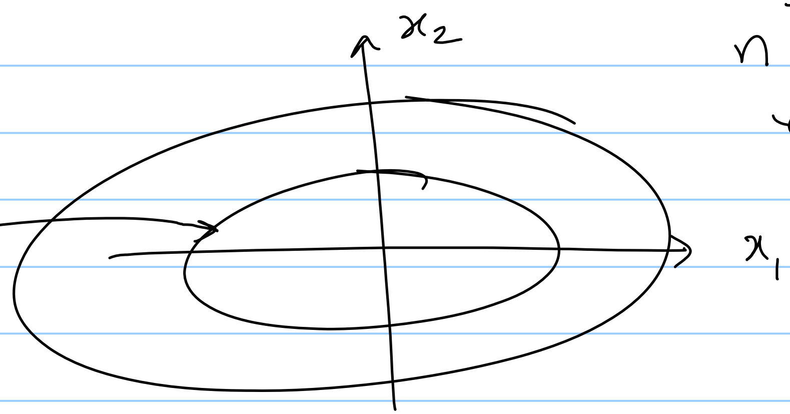
$q(x) = x^T A x$

Assume A is diagonal.

$$= \begin{bmatrix} x_1 & \quad \end{bmatrix} A \begin{bmatrix} x_1 \\ \\ \end{bmatrix} = \sum_{i=1}^{n} \underbrace{x_i^2 A_{ii}}_{\text{n dimensional ellipsoid.}}$$

2D $\quad 5x_1^2 + 2x_2^2$

Contours

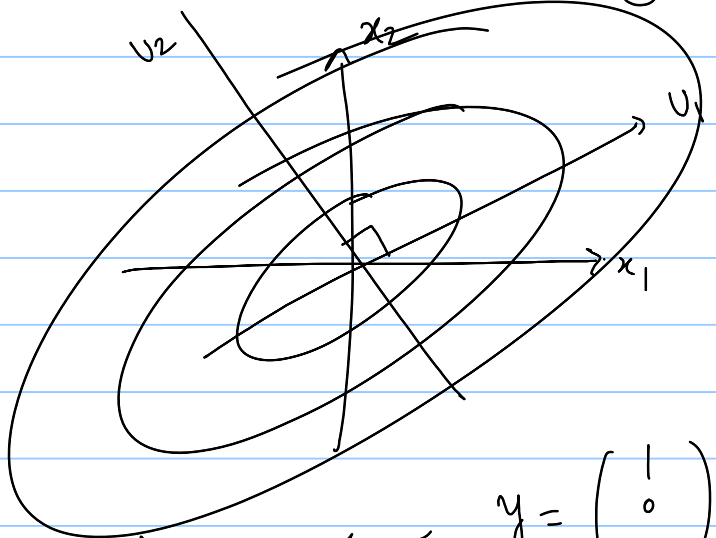


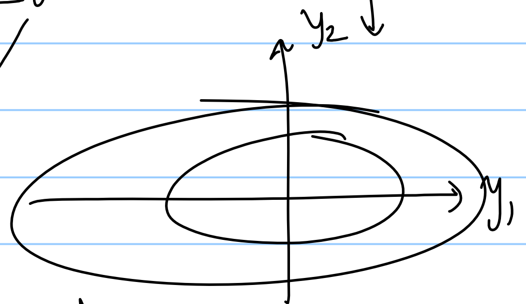
A is not diagonal.

$$q(x) = x^T A x = \underbrace{x^T U}\Lambda\underbrace{U^T x}$$

Say $\underbrace{U^T x = y} \Rightarrow y^T = x^T U$

$$\underline{x = Uy} \qquad = y^T \Lambda y$$





$x_1 = v_1 \leftarrow y = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \leftarrow$ Cannonical basis

↳ Orthogonality → $p_i, p_j \Rightarrow p_i^T p_j = 0 \quad i \neq j$

Conjugacy → $\boxed{p_i^T A p_j = 0 \quad , i \neq j}$

We say $\{p_0, \ldots, p_{n-1}\}$ is conjugate w.r.t P.D matrix A.

↳ These p_i's are linearly independent.

Proof: By Contradiction. $(p_1, \ldots, p_\ell)$

$$p_k = \sum_{\substack{i=1 \\ i \neq k}}^{\ell} \alpha_i p_i$$

Left Multiply $p_k^T A$

$$\underbrace{p_k^T A p_k} = \sum_{\substack{i=1 \\ i \neq k}}^{\ell} \alpha_i \, p_k^T A p_i = 0$$

$$\underbrace{A} = U \Lambda U^T = \sum_{i=1}^{n} \lambda_i \underbrace{u_i u_i^T}$$ (outer product notation)

$$\sum_{i=1}^{n} \lambda_i \underbrace{(p_k^T u_i)}\underbrace{(u_i^T p_k)} = \sum_{i=1}^{n} \underbrace{\lambda_i}_{> 0} \underbrace{\| u_i p_k^T \|^2}_{> 0} = 0$$

A contradiction.

Conjugate Directions Method (CDM)
(ancestor).

Iterative

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

Need not be a descent direction.

however

2 requirements

① The p_i's must be conjugate w.r.t. A

② Step length is an exact minimizer of $\phi(x)$ along the p_k direction, $\frac{d}{d\alpha}\phi(x_k + \alpha p_k) = 0$

$$\left(\alpha_k = - \frac{r_k^T p_k}{p_k^T A p_k} \right), \quad r_k = A x_k - b.$$

Result: Starting from x_0, the sequence $\{x_k\}$ generated as per (1) & (2) above, converges to x^* ($Ax^* = b$) in at most n steps!

Proof: We are at step k, assume $k \leq n$

$(p_0, p_1, \ldots, p_{n-1}) \rightarrow$ Set of CDs.

$\rightarrow$ lin indepn

$$\left[\, \overbrace{x^* - x_0} = \sum_{i=0}^{n-1} \sigma_i p_i \,\right] \quad (\text{basis exp}$$

L. M. by $p_k^T A \rightarrow p_k^T A (x^* - x_0) = \sigma_k \, p_k^T A p_k$

$$\Rightarrow \quad \sigma_k = \frac{p_k^T A (x^* - x_0)}{p_k^T A p_k} \quad \longleftarrow$$

$$\longrightarrow \quad x_k = \underbrace{\underbrace{x_0 + \alpha_0 p_0}_{x_1} + \alpha_1 p_1}_{x_2} + \cdots + \alpha_{k-1} p_{k-1}$$

$$\longrightarrow \quad \underbrace{x_k - x_0} = \sum_{i=0}^{k-1} \alpha_i p_i \quad \longleftarrow$$

$$\longrightarrow \quad p_k^T A (x_k - x_0) = 0 \quad \longleftarrow$$

$$\alpha_k = \frac{p_k^T A (x^* - x_0)}{p_k^T A p_k} = \frac{p_k^T A (x^* - x_k + x_k - x_0)}{p_k^T A p_k}$$

$$= \frac{p_k^T A (x^* - x_k)}{p_k^T A p_k} = \frac{p_k^T (b - A x_k)}{p_k^T A p_k}$$

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

$$\text{———} \times \text{———}$$