Convergence Rate (steepest descent)

Assume convex & quadratic objective fn.

$\Rightarrow$ exact line search gives closed form expr.

$$f(x) = \frac{1}{2} x^T Q x - b^T x \qquad \nabla f_k = Q x_k - b \quad (1)$$

Q is sym, pos. def.

$$x_{k+1} = x_k - \alpha_k \nabla f_k .$$

$$\phi(\alpha) = f(x_k - \alpha_k \nabla f_k)$$

$$x_{k+1} = x_k - \left(\frac{\nabla f_k^T \nabla f_k}{\nabla f_k^T Q \nabla f_k} \right) \nabla f_k$$

want measure of $\| x_k - x^* \|_2$

$$\| a \|_Q^2 = x^T Q x . \qquad Q x^* = b \quad (2)$$

$$\frac{1}{2} \| x_k - x^* \|_Q^2 = \frac{1}{2} (x_k - x^*)^T Q (x_k - x^*)$$

$$= \frac{1}{2} \Big[x_k^T Q x_k - \underbrace{x^{*T} Q}_{b^T} x_k + \underbrace{x^{*T} Q x^*}_{-f(x^*)} - x_k^T \underbrace{Q x^*}_{b} \Big]$$

$$= \frac{f(x_k) - f(x^*)}{\quad}$$

Luenberger.

$$\|x_{k+1} - x^*\|_Q^2 \le \left(\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}\right) \|x_k - x^*\|_Q^2$$

$\to r^2$

eigs of Q

$$\left(\frac{K-1}{K+1}\right)$$

$\Rightarrow$

① Convergence is linear.

$f(x_k) - f(x^*) < \varepsilon \Rightarrow O(1/\varepsilon)$ iterations.

$\ldots r\, r\, \|x_0 - x^*\|$

② $K(Q) = \frac{\sigma_{max}}{\sigma_{min}} = \frac{\lambda_n}{\lambda_1}$

— x —.

Conjugate Gradient Methods (Ch 5)

(CG)

↳ What is the problem

→ To solve $Ax = b$ iteratively

where A is sym pos def.

Gaussian Elim $O(n^3)$

Equivalent problem:

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

$\nabla \phi(x^*) = Ax^* - b = 0$ at stat pt.

$$\nabla \phi(x) = Ax - b = \underbrace{r(x)}_{\text{residual}}$$

residual.
$\updownarrow$
also gradient.

↳ If A is sym pos definite
$\Rightarrow$ $A = U \Lambda U^T$

$\underbrace{[u_1 \; \dots \; u_n]}_{\text{eig vectors}}$ $\quad$ $\underbrace{\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}}_{\text{eigvalues.}}$

eig vectors $\quad$ eigvalues.

1) $U^{-1} = U^T$
2) $\| Uq \| = \| q \|$
3) $(Uq, Ut) = (q, t)$.

Is true for all types of norms?

↳ <u>Visualizing quadratic forms.</u>

$q(x) = x^T A x.$ $\quad [\quad] A \begin{bmatrix} \\ \\ \end{bmatrix}$

if A is diagonal. $\rightarrow$ $q(x) = \sum_i A_{ii} x_i^2$

say $A_{ii} > 0$ → n-dim ellipsoid.

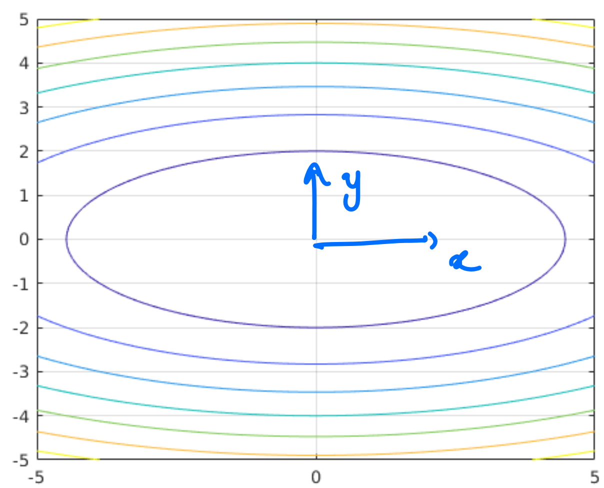
$$q(x) = A_{11} x_1^2 + A_{22} x_2^2$$

If A is not diagonal →

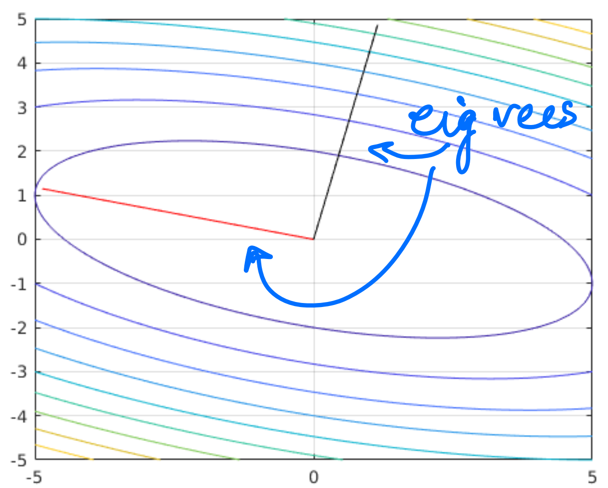
$$q(x) = x^T U \Lambda U^T x$$
$$= y^T \Lambda y.$$

Denote $U^T x = y$

Ellipsoid in the y- cords.



diag



non diag.

↪ Conjugacy property

: orthogonality → $P_i^T P_j = 0$ for $i \neq j$.

We say that $\{p_0, \dots, p_{n-1}\}$ is conjugate w.r.t. a sym pos def matrix A if:

$$\boxed{p_i^T A p_j = 0 \qquad i \neq j.}$$

→ such vectors are linearly independent.

proof: By contradiction. $(p_1, \cdots, p_\ell)$

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

$$p_k^T A p_k = \sum_i \alpha_i \underbrace{p_k^T A p_i}_{\text{by conjugacy}} = 0$$

$$A = U \Lambda U^T$$
$$= \sum_{i=1}^{n} \lambda_i u_i u_i^T$$

$$\left.\begin{array}{l} \sum_i \underbrace{p_k^T u_i u_i^T p_k}_{\|u_i^T p_k\|^2} \underbrace{\lambda_i}_{>0} = 0 \end{array}\right\} \Rightarrow u_i^T p_k = 0 \quad \forall i$$

u's are orthogonal ⇒ form a basis.

$= 0 \iff p_k = 0$ $\qquad\qquad$ $p_k = 0.$

for pos def. matrix by defn.

—— x ——