10/27/22

last time: eigen decomposition

$$\mathrm{Var}(X) = \Sigma = Q \Lambda Q^T$$

$$Q = \begin{bmatrix} q_1 & \cdots & q_i & \cdots & q_d \end{bmatrix} \qquad \Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_d \end{bmatrix}$$

diagonal matrix

$$X = Q \Lambda^{\frac{1}{2}} Z$$

$$Z = \Lambda^{-\frac{1}{2}} Q^T X$$

$$\mathrm{Var}(Z) = I$$

Data version:

$$\begin{array}{c|c|} & 1 \cdots i \cdots d \\ \hline 1 & \\ \vdots & \\ i & x_i^T \\ \vdots & \\ n & X_{n \times d} \end{array} \qquad z_i^T \leftarrow d \times 1 \text{ matrix} \qquad Z_{n \times d}$$

also called "design" matrix

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i = 0$$

decomposition

$$\hat{\Sigma} = \frac{1}{n} \sum_{i=1}^{n} x_i x_i^T = Q \Lambda Q^T$$

$d \times 1$ $\quad$ $1 \times d$ $\quad = d \times d$

$$x_i = Q \Lambda^{\frac{1}{2}} z_i$$

$$z_i = \Lambda^{-\frac{1}{2}} Q^T x_i$$

$$\frac{1}{n} \sum_{i=1}^{n} z_i z_i^T = I \qquad \frac{1}{n} Z^T Z = I$$

$$X^T = \begin{bmatrix} x_1 \cdots \boxed{x_i} \cdots x_n \end{bmatrix}_{d \times n} = Q \Lambda^{\frac{1}{2}} Z^T$$

singular value decomposition

multiply by Σ

nom verb

do not confuse w/ x^T

$$\begin{tikzcd} X \arrow[r, "\Sigma"] & X' \\ Y \arrow[u, "Q"] \arrow[r, "\Lambda"] & Y' \arrow[u, "Q"'] \end{tikzcd}$$

$$x' = \Sigma x$$
$$QY' = \Sigma QY$$
$$Y' = Q^T \Sigma Q Y$$

$$\Sigma = Var(X)$$
$$\Lambda = Var(Y) = Var(Q^T X) = Q^T var(X) Q = Q^T \Sigma Q$$

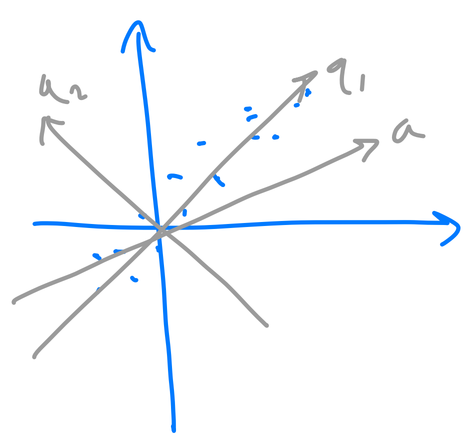
Power method

$$\begin{tikzcd} X \arrow[r, "\Sigma^t"] & X' \\ Y \arrow[u, "Q"] \arrow[r, "\Lambda^t"] & Y' \arrow[u, "Q"'] \end{tikzcd}$$

$$\Lambda^t = \begin{bmatrix} \lambda_1^t & & & \\ & \lambda_2^t & & \\ & & \ddots & \\ & & & \lambda_d^t \end{bmatrix}$$

$$\lambda_1 > \lambda_2 \cdots > \lambda_d > 0$$

$$Y' = \begin{bmatrix} \lambda_1^t y_1 \\ \lambda_2^t y_2 \\ \vdots \\ \lambda_d^t y_d \end{bmatrix} \xrightarrow[\text{normalize}]{\frac{y'}{\|y'\|}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \xrightarrow{Q} q_1$$



$Var(\langle a, x \rangle)$ — projecting x onto a axis

$$\|a\| = 1$$

$$Var(\langle a, x \rangle) = Var(a^T x) = a^T Var(x) a$$
$$= \underbrace{a^T}_{1\times d} \underbrace{\Sigma}_{d\times d} \underbrace{a}_{d\times 1} \geq 0 \quad \forall a$$

$$\Sigma \geq 0, \ \lambda_i \geq 0 \ \forall i$$

first use power method to find solutions, then maximize or minimize

$$\max_{|a|=1} \; a^T \Sigma a$$

$$\begin{tikzcd} q_1 = a \arrow[r, "\Sigma"] & {} \\ b \arrow[u, "Q"] \arrow[r, "\Lambda"] \arrow[d] & {} \arrow[u] \\ \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} & \end{tikzcd}$$

$$a^T \Sigma a = b^T \Lambda b = \sum_{i=1}^{d} \lambda_i b_i^2 \xrightarrow[b_1 = 1]{\max} \lambda_1$$

$$a^T \Sigma a = (Q b)^T (Q \Lambda Q^T)(Q b)$$
$$= b^T \cancel{Q^T Q} \Lambda \cancel{Q^T Q} b$$

$$\min_{|a|=1} \; a^T \Sigma a = \lambda_d$$

$$a = q_d$$

$$\max_{\substack{|a|=1 \\ a \perp q_1}} \; a^T \Sigma a = \lambda_2$$

$a \perp q_1$ – ensures $b_1 = 0$

$$\begin{array}{c|c} \begin{bmatrix} \\ \dfrac{d}{dx_i} \\ \\ \end{bmatrix} & \begin{matrix} 1 \\ 2 \\ \vdots \\ i \\ \vdots \\ d \end{matrix} \end{array} \qquad \begin{array}{c} 1 \cdots j \cdots d \\ \begin{bmatrix} & \dfrac{d}{dx_j} & \end{bmatrix} \end{array}$$

Taylor expansion

$$y = F(x)$$
$1 \times 1 \quad d \times 1$

$\dfrac{dy}{dx}$ ← $d \times 1$

$$\dfrac{d^2 y}{dx\, dx^T} \quad \begin{array}{c} j \\ i \rightarrow \begin{bmatrix} \dfrac{d^2}{dx_i dx_j} \end{bmatrix} \end{array} = \left(\underset{d \times 1}{\dfrac{d}{dx}}\right)\left(\underset{1 \times d}{\dfrac{d}{dx^T}}\right)$$

$$F(x + \Delta x) = F(x) + \langle F'(x), \Delta x \rangle + \frac{1}{2} \Delta x^T F''(x) \Delta x + o(|\Delta x|^2)$$

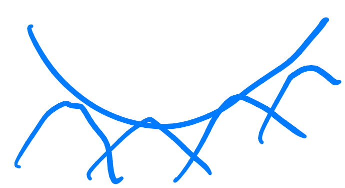
$F'(x) = 0$

second derivative

$F''(x) > 0$ x is local min

$F''(x) < 0$ x is local max

saddle point



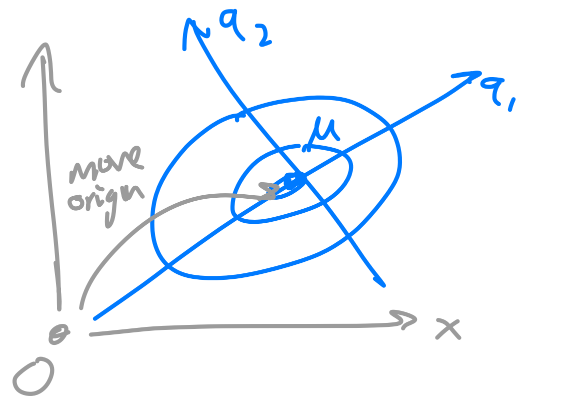
$$\begin{array}{ccc} \Delta x & \xrightarrow{F''(x) \,=\, Q\Lambda Q^T} & \\ Q\uparrow & & \\ \Delta y & \xrightarrow{\ \Lambda\ } & \end{array}$$

$$\Delta x^T F''(x)\,\Delta x = \Delta y^T \Lambda^{-1} \Delta y = \sum_{i=1}^{d} \lambda_i^{-1} \Delta y_i^2$$

Back to Normal example. $\therefore\ y^T \Lambda^{-1} y = \sum_{i=1}^{d} \lambda_i^{-1} y_i^2 = \text{constant}$

$$f(x) = \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \boxed{(x-\mu)^T \Sigma^{-1} (x-\mu)}\right) \qquad \sum_{i=1}^{d} \frac{y_i^2}{\sigma_i^2} = \text{const.}$$

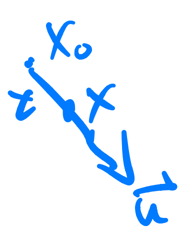
$$\begin{array}{ccc} x-\mu & \xrightarrow{\ \Sigma\ } & \\ Q\uparrow & & \\ y & \xrightarrow{\ \Lambda\ } & \end{array}$$



Taylor expansion

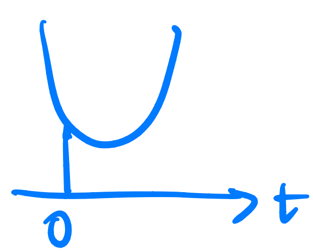
$$F(x) = F(x_0) + \langle F'(x_0),\ x - x_0 \rangle + \frac{1}{2}(x-x_0)^T F''(x_0)(x-x_0)$$

$F(x)$



$$x = x_0 + t\vec{u}$$

$$f(t) = F(x_0 + t\vec{u})$$

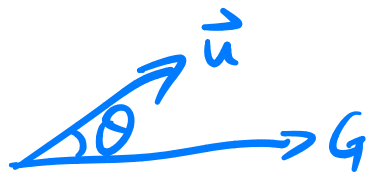


$$f(t) \doteq f(0) + f'(0)t + \frac{1}{2} f''(0)\, t^2$$

$$f'(t) = \frac{\partial F}{\partial x^T}\frac{\partial x}{\partial t} = \langle F'(x), \vec{u}\rangle = \vec{u}^T F'(x)$$

$$f'(0) = \langle F'(x_0), \vec{u}\rangle = \langle G, \vec{u}\rangle$$

$$= |G| \cos\theta$$



max $= |G|$, $\theta = 0$

gradient

$$f''(t) = u^T \frac{\partial F'}{\partial x^T}\frac{\partial x}{\partial t} = u^T \frac{\partial \frac{\partial F}{\partial x}}{\partial x^T}\, u$$

$$= u^T \frac{\partial^2 F}{\partial x \partial x^T}\, u = u^T F''(x)\, u$$

$$f''(0) = u^T F''(x_0)\, u = u^T H u$$

$$= v^T \Lambda v$$

$$= \sum_i \lambda_i v_i^2$$

Curvature

$u \xrightarrow{H}$

$a \uparrow$

$v \xrightarrow{\Lambda}$

Hessian

$\lambda_1 v_1^2 + \lambda_2 v_2^2$