Mathematics Basics

BASIC DEFINITION

$$Ax = b$$

Range $\longrightarrow$ has solution ?

Kernel $\longrightarrow$ how many ?

NORMS

$\langle u, v \rangle = \sum u_i v_i = u^T v$

$\|u\|_2 = \sqrt{\langle u, v \rangle}$

p-norms, 1-norm, ...

$\|u\|_1 = \sum |u_i|$

SVD

$(\sigma_i \geq 0)$

SVD of A is:

$$A = U \Sigma V^T \qquad \Sigma = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix}$$

$$Ax = U \Sigma V^T x$$

SPECTRAL / EIGEN-DECOMPOSITION

$$A = X \Lambda X^{-1}$$

if A symmetric: $A = X \Lambda X^T$

$$Ax = A\left(\sum \langle x, v_i \rangle v_i\right)$$
$$= \sum \langle x, v_i \rangle A v_i$$
$$= \sum \lambda_i \langle x, v_i \rangle v_i$$

$$x^T A x = \sum \lambda_i \langle x, v_i \rangle^2 \leq \sum \lambda_{max} \langle x, v_i \rangle^2 \leq \lambda_{max} \sum \langle x, v_i \rangle^2 = \lambda_{max} \|x\|^2$$

$$\sum \langle x, v_i \rangle^2 = \|x\|^2$$

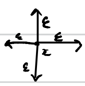
APPROXIMATION

Taylor Approximation:

$$f(x+\varepsilon) \cong f(x) + \langle \nabla f(x), \varepsilon \rangle + \frac{1}{2} \varepsilon^T \underbrace{D^2 f(x)}_{\left[\frac{\partial^2 f}{\partial x_i \partial x_j}\right]_{i,j}} \varepsilon + o(\|\varepsilon\|^2)$$

$$= f(x) + \langle \nabla f(x), \varepsilon \rangle + \frac{1}{2} \varepsilon^T D^2 f(\underbrace{\tilde{x}}_{\text{difference}}) \varepsilon$$

ε random

① 

Q $\mathbb{E}[f(x+\varepsilon)] = ?$ Expected value

$$\mathbb{E}[f(x+\varepsilon)] = \mathbb{E}[f(x) + \langle \nabla f(x), \varepsilon \rangle + \frac{1}{2} \varepsilon^T D^2 f(\tilde{x}) \varepsilon]$$
$$= f(x) + \mathbb{E}\langle \nabla f(x), \varepsilon \rangle + \frac{1}{2} \mathbb{E}[\varepsilon^T D^2 f(x) \varepsilon]$$
$$= f(x) + \langle \nabla f(x), \mathbb{E}\varepsilon \rangle + \frac{1}{2} \mathbb{E}[\varepsilon^T D^2 f(x) \varepsilon]$$
$$= f(x) + 0 + \frac{1}{2} \mathbb{E}[\varepsilon^T D^2 f(x) \varepsilon]$$

Assume that for all $y \in \mathbb{R}^n$, $D^2 f(y)$ has the property that $\lambda_{max} \leq \bar{\lambda}$, $\underline{\lambda} \leq \lambda_{min}(y)$

$$\underline{\lambda} \|\varepsilon\|^2 \leq \varepsilon^T D^2 f(x) \varepsilon \leq \bar{\lambda} \|\varepsilon\|^2$$

$$\underbrace{f(x)}_{\underline{\lambda} \geq 0} \leq f(x) + \underline{\lambda} \underbrace{\mathbb{E}[\|\varepsilon\|^2]}_{=1} \leq \mathbb{E}[f(x+\varepsilon)] \leq f(x) + \bar{\lambda} \underbrace{\mathbb{E}[\|\varepsilon\|^2]}_{=1} \longrightarrow 2^{nd} \text{ Moment } (= \text{variance})$$