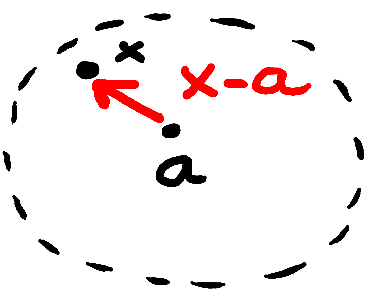
Session 15 MLE and regression

Recap: Quadratic approximation
Geometric interpretation of critical points

$f:\mathbb{R}^n \to \mathbb{R}$

Jacobian $\left[\frac{\partial f}{\partial x_i}\right]_i \in \mathbb{R}^{1\times n}$ $\left[\frac{\partial^2 f}{\partial x_i \partial x_j}\right] \in \mathbb{R}^{n\times n}$

$$f(x) \approx f(a) + Jf(a)(x-a) + \frac{1}{2}(x-a)^t Hf(a)(x-a)$$



$Hf(a)$ symmetric $\leftrightarrow$ $\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a)$

orthonormal eigenbasis $\{v_1, \dots, v_n\}$
eigenvalues $\lambda_1, \dots, \lambda_n$

Let $Q = \left[v_1 | \cdots | v_n \right]$ the orthogonal matrix with v_i as columns

decoder

$$\begin{tikzcd} \text{eigenbasis} \arrow[r, "D=\mathrm{diag}(\lambda_1,\dots,\lambda_n)"] \arrow[d, "Q"'] & \text{eigenbasis} \\ \text{canonical} \arrow[r, "Hf(a)"'] & \text{canonical} \arrow[u, "Q^t"'] \end{tikzcd}$$

decoder encoder

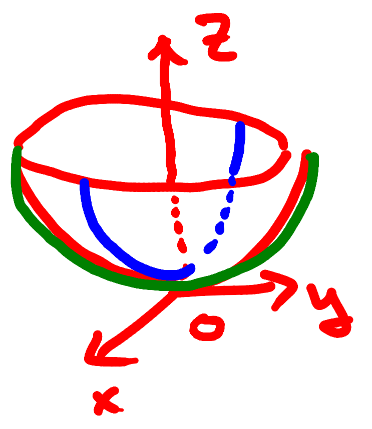
$$D = Q^t Hf(a) Q$$

let $\hat{a}$ be a critical point
$Jf(a) = [0 \dots 0]$ then

$$f(x) \approx f(a) + (x-a)^t Hf(a)(x-a)$$

$QQ^t = Id$

graphs are paraboloids



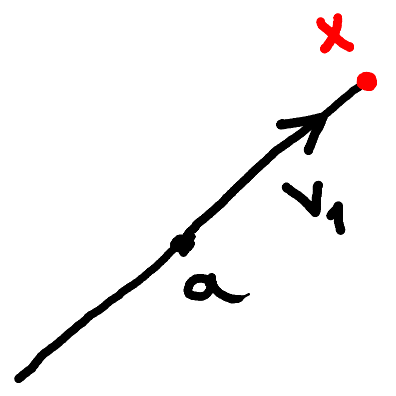
Check theory of quadratic forms, maps of the form
$x \longmapsto x^t A x$ $x \in \mathbb{R}^n$ A: symmetric $\mathbb{R}^{n\times n}$

$$f(x) \approx f(a) + \underbrace{(x-a)^t Q}_{w^t} \underbrace{Q^t Hf(a) Q}_{D} \underbrace{Q^t (x-a)}_{w}$$

$D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$

w: vector $x-a$ written in coords of the eigenbasis

For example, if x is along the line
$L = a + \mathrm{span}(v_1)$ then $x = a + \mu v_1$ for some $\mu \in \mathbb{R}$
$\Rightarrow x - a = \mu v_1 \Rightarrow Q^t(x-a) = Q^t(\mu v_1) = \mu Q^t v_1 = \begin{bmatrix} \mu \\ 0 \\ \vdots \\ 0 \end{bmatrix}$



Then the vertical cross-section along v_1 at a of the graph of f has equation

$$z = f(a) + [\mu\ 0 \dots 0] \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \begin{bmatrix} \mu \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$z = f(a) + \boxed{\lambda_1} \mu^2$$ $\to$ equation of a parabola

vertical cross-section at a along v

- $\begin{cases} \lambda_1 > 0 & \cup \\ \lambda_1 < 0 & \cap \end{cases}$
- $|\lambda_1|$ how open the branches are. the lower the more open

$\Rightarrow$ Clasification critical points of $f:\mathbb{R}^n \to \mathbb{R}$ at a using the quadratic approximation.