FoDA - L18

Dimensionality Reduction

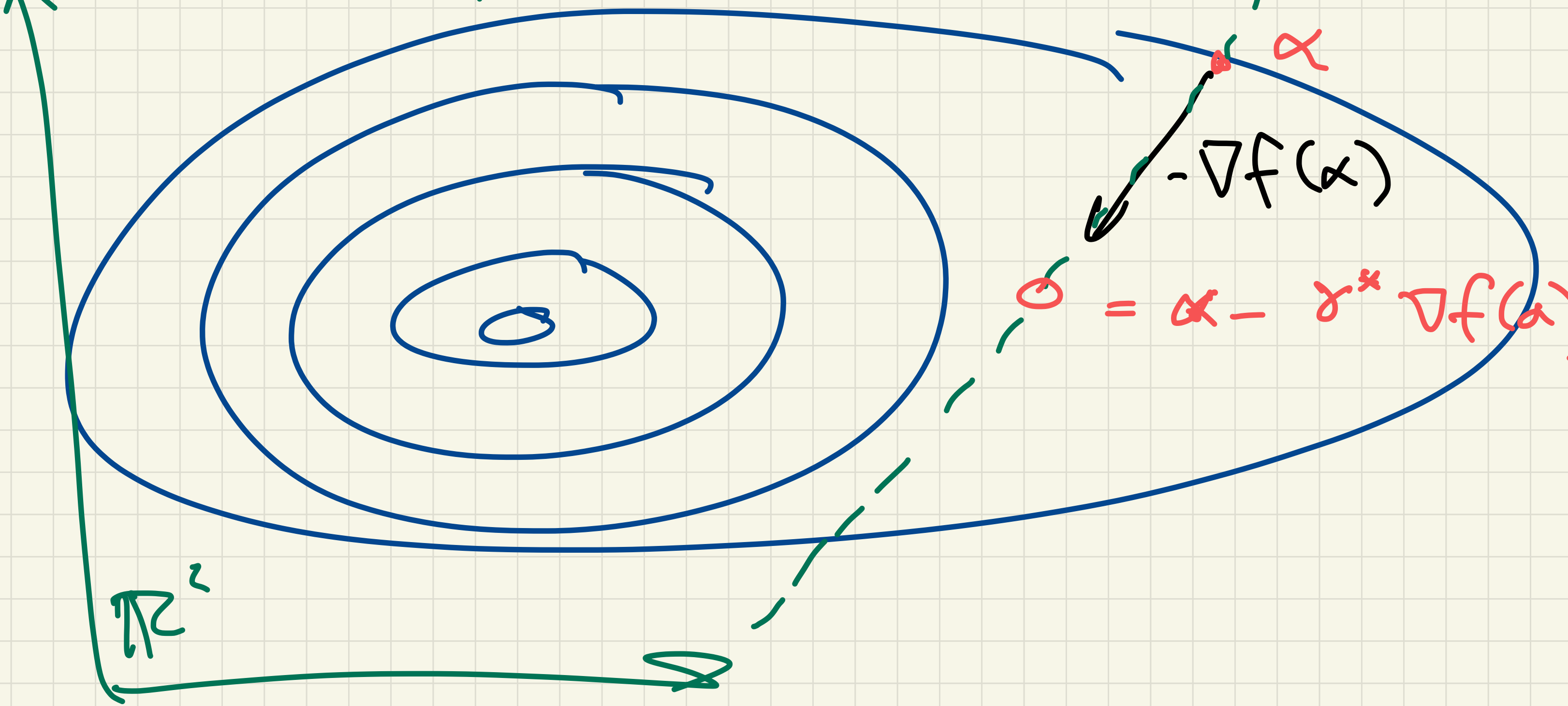
Gradient Descent

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

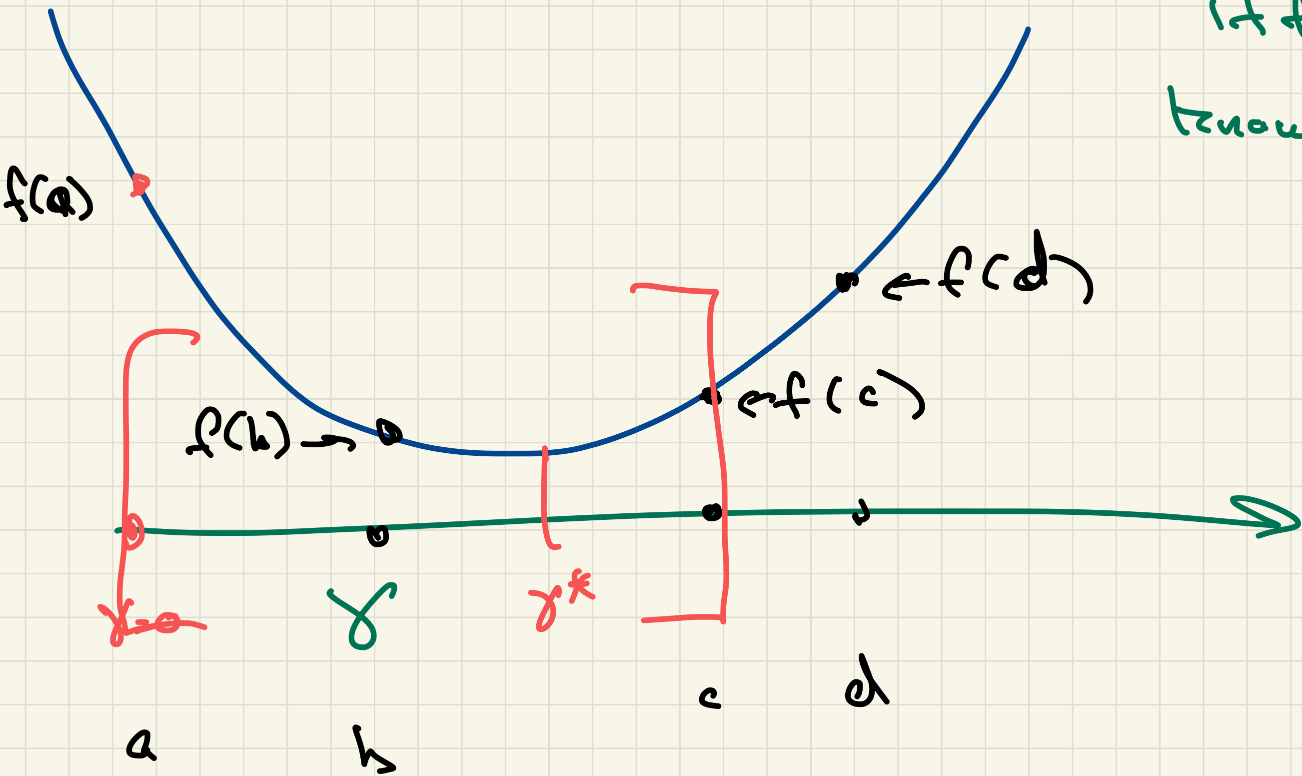
$\underset{\alpha \in \mathbb{R}^d}{\operatorname{argmin}} f(\alpha)$

$\alpha = \alpha - \gamma_k \nabla f(\alpha)$

line search



Golden Section Search



if $f(b) < f(c)$

know $\gamma^* \in [a, c]$

$$\nabla f : \mathbb{R}^d \to \mathbb{R}^d$$

$\alpha \in \mathbb{R}^d$

$\alpha = (\alpha_1, \alpha_2, \dots \alpha_d)$

$$\nabla f(\alpha) = \left(\frac{\partial}{\partial \alpha_1} f(\alpha), \dots, \frac{\partial}{\partial \alpha_d} f(\alpha) \right)$$

$$\frac{\partial}{\partial \alpha_j} f(\alpha) \leftarrow \text{says } \alpha_i \;\; i \neq j \text{ is const.}$$

A function $g: \mathbb{R}^d \to \mathbb{R}^k$ is

L-Lipschitz if

for any $p, q \in \mathbb{R}^d$

$$\|g(p) - g(q)\| \leq L \cdot \|p - q\|$$

$$\frac{\|g(p) - g(q)\|}{\underbrace{\|p - q\|}_{h}} \leq L$$

$g = \nabla f$

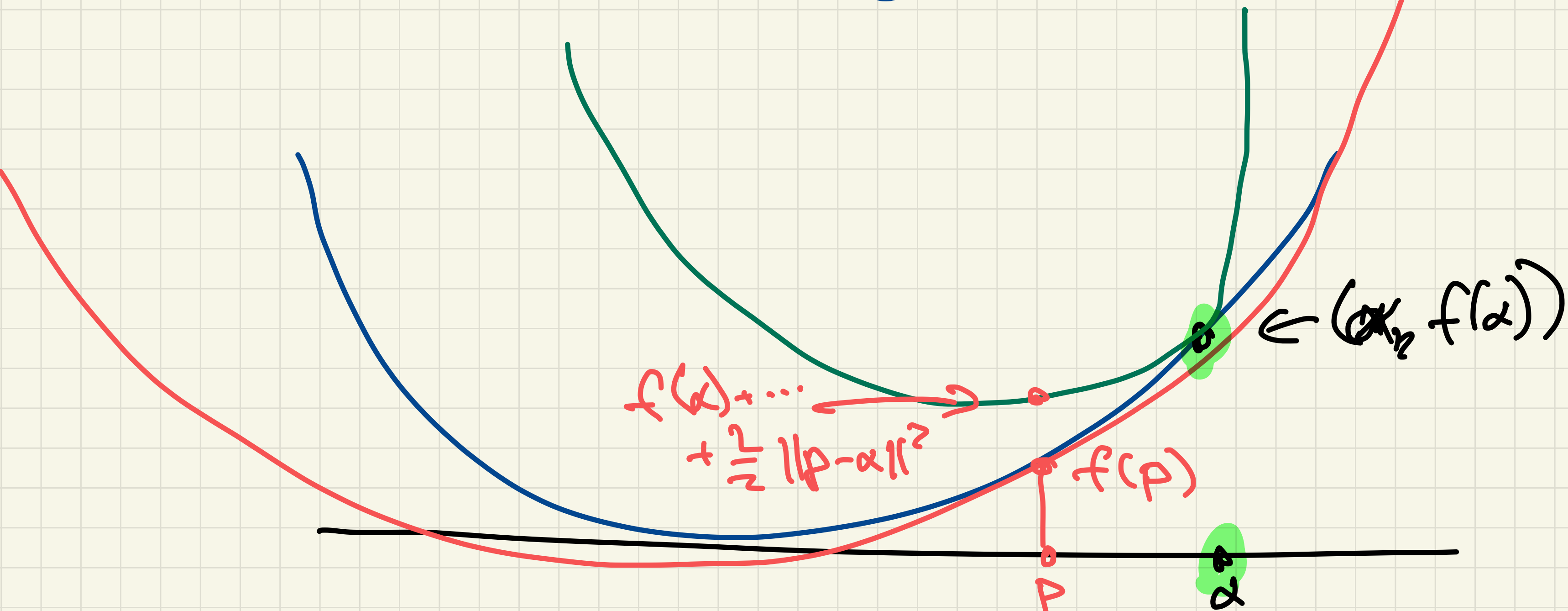
$q = p + h u$

$\|u\| = 1$

$p, \alpha \in \mathbb{R}^d$

∇f is L-Lipschitz if

$$f(p) \leq f(\alpha) + \frac{L}{2}\|p-\alpha\|^2$$

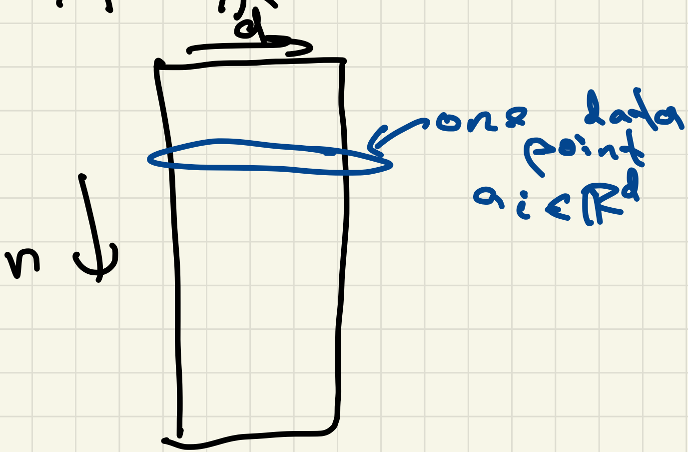


Dimensionality Reduction

n data points in $\mathbb{R}^d$ ← d is too big

$a_1, a_2, \dots a_n \in \mathbb{R}^d$

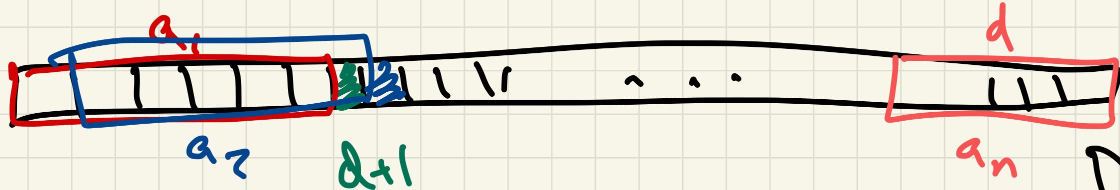
$A \in \mathbb{R}^{n \times d}$



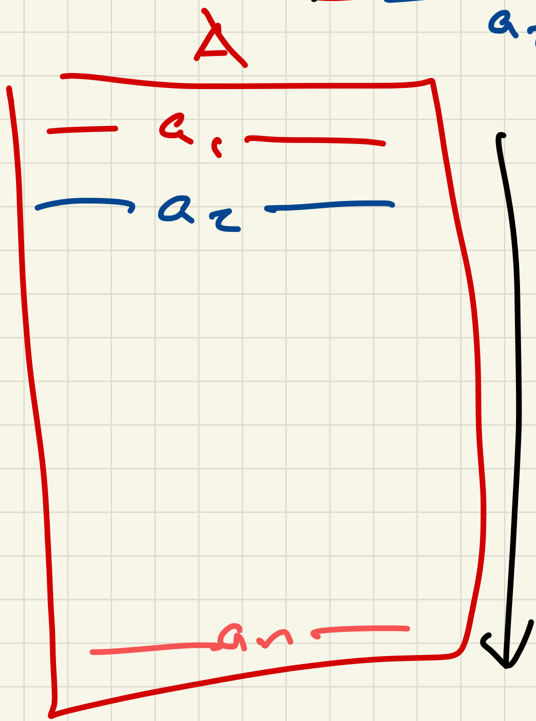
- n weather stations
  d times measure temp.

- n Netflix users) → rating
  d movies catalog

- ~~n~~ stocks track
  D days data closing price



D days



$n = D-(d)$

All columns have the same units!

$a = (a_1, a_2 \dots a_d)$, $a' = (a_1', a_2', \dots a_d')$

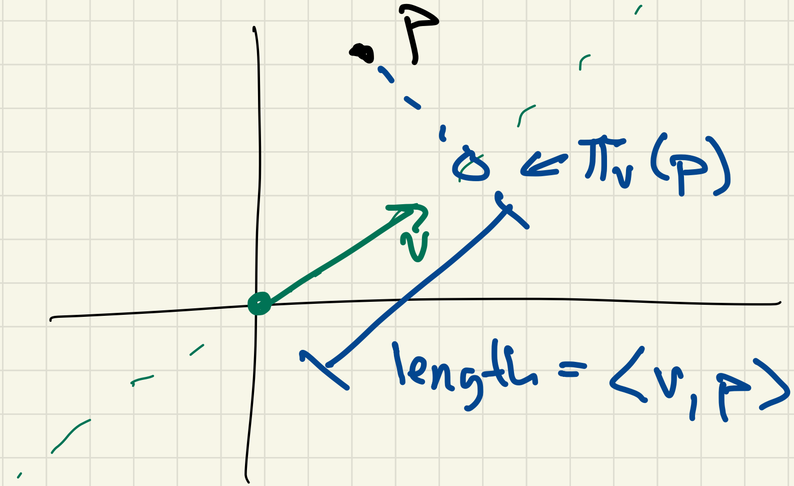
$$\| a - a' \| = \sqrt{\sum_{j=1}^{d} (a_j - a_j')^2}$$

Projections

$p \in \mathbb{R}^d$

unit vector $v \in \mathbb{R}^d$

$\|v\| = 1$

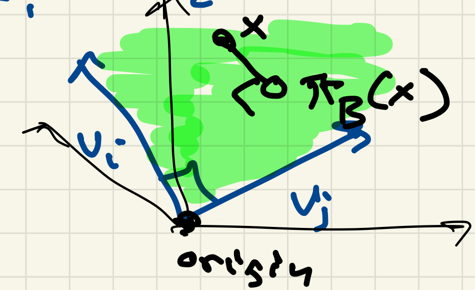


$\langle v, p \rangle$

$\pi_v(p) = \langle v, p \rangle v \in \mathbb{R}^d$

$B \subset k$-dimensional subspace

orthogonal basis $V_B = \{v_1, v_2, \dots v_k\}$



- $\|v_j\| = 1$
- $\langle v_j, v_i \rangle = 0 \quad j \neq i$

For any point $x \in B$

$$x = \sum_{j=1}^{k} \alpha_j v_j \qquad \alpha_j = \langle x, v_j \rangle$$

For $x \notin B$ $x \in \mathbb{R}^d$ project x onto B as

$$\pi_B(x) = \sum_{j=1}^{k} \langle v_j, x \rangle v_j = \sum_{j=1}^{k} \pi_{v_j}(x)$$

Goal: given $a_1 \dots a_n = A$

Find B dim k

minimize $SSE(A, B)$

$$= \sum_{i=1}^{n} \left(a_i - \pi_B(a_i)\right)^2$$