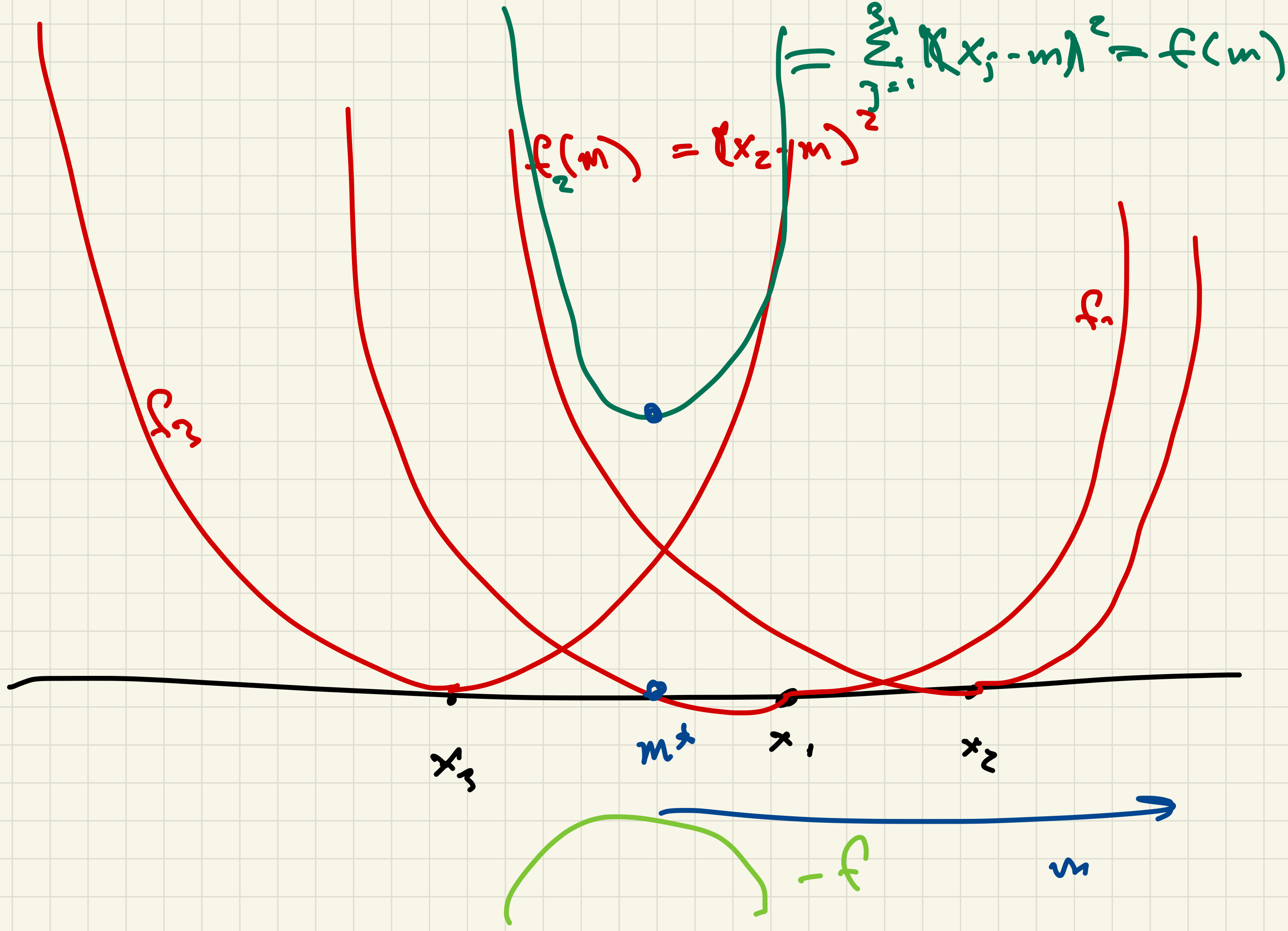
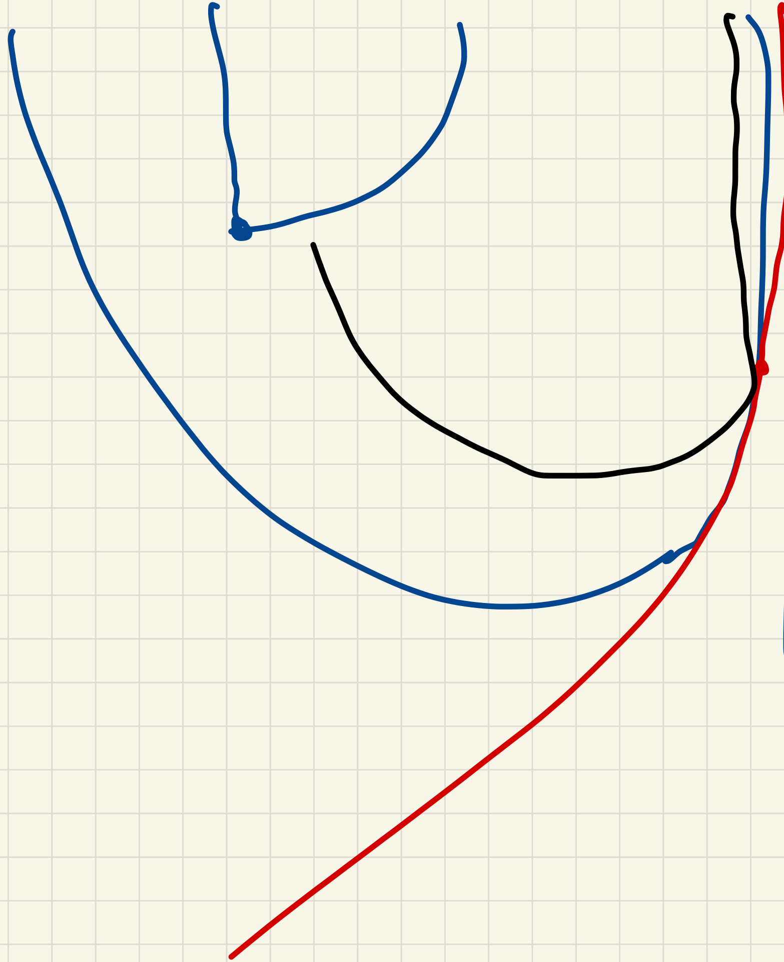
FoDA L18

Dimensionality Reduction

Data Matrices & Projections

Oct 27, 2022


$= \sum_{j=1}^{3} (x_j - m)^2 = f(m)$
$f_2(m) = (x_2 - m)^2$
f_3
f_1
x_3
m^*
x_1
x_2
$-f$
m

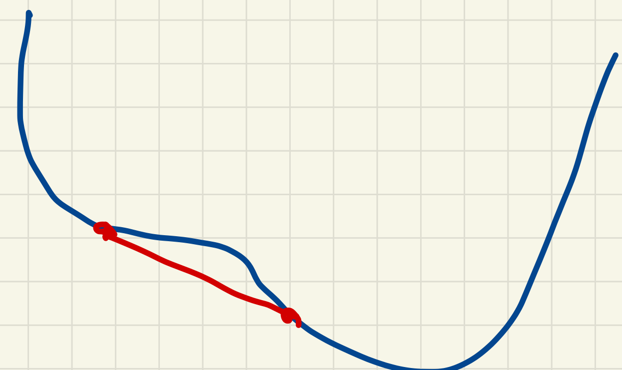


Strongly convex

- no kinks

↳ 2nd deriv continuous

- convex



Dimensional Reduction

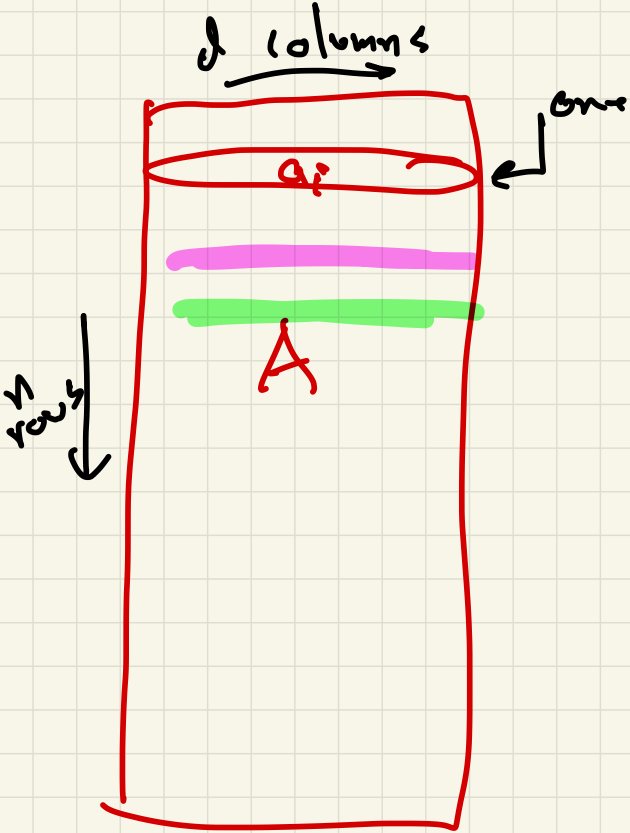
Why high dimensions bad?

- hard to understand
- computationally expensive.
  - read 1 point $O(d)$
  - "curse of dimensionality"
  - algo. runtime $O(n^d)$
- "overfit"

Input

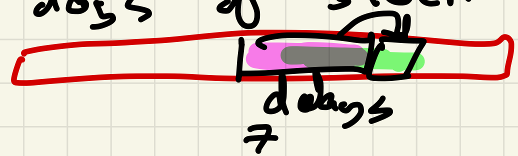
$A = \{a_1, a_2, \ldots, a_n\} \subset \mathbb{R}^d$ $\quad d = \text{big}$

typically $n > d$



$a_i = (a_{i1}, a_{i2}, \ldots a_{id})$

- n weather stations
  d days of max temp.
- n user, d movies
  $\to a_{ij}$ i rates movie j
- N days of stock prices



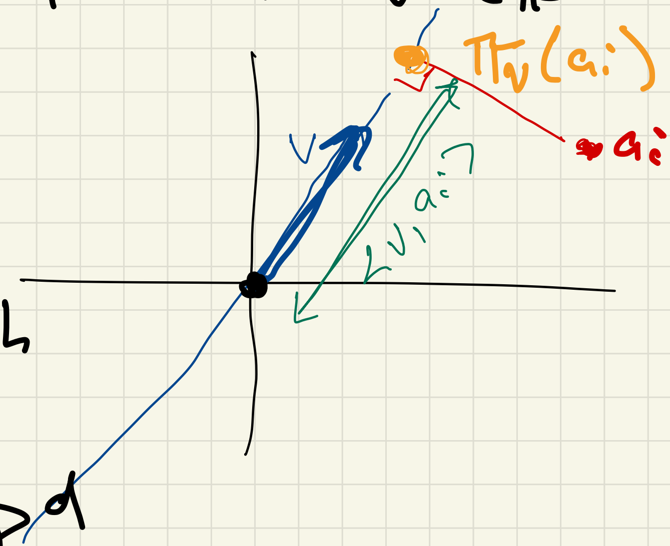
$n = N - d$

Every a_{ij} same units

Projections

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

$\langle a_i, v \rangle$



$\pi_v(a_i) \equiv$ closest point on line through v, to a_i

$= \langle v, a_i \rangle v \in \mathbb{R}^d$

Subspace B (assume contains origin)

set of orthogonal basis vectors $V_B = \{v_1, v_2, \dots, v_k\} \subset \mathbb{R}^d$

k-dimen.

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

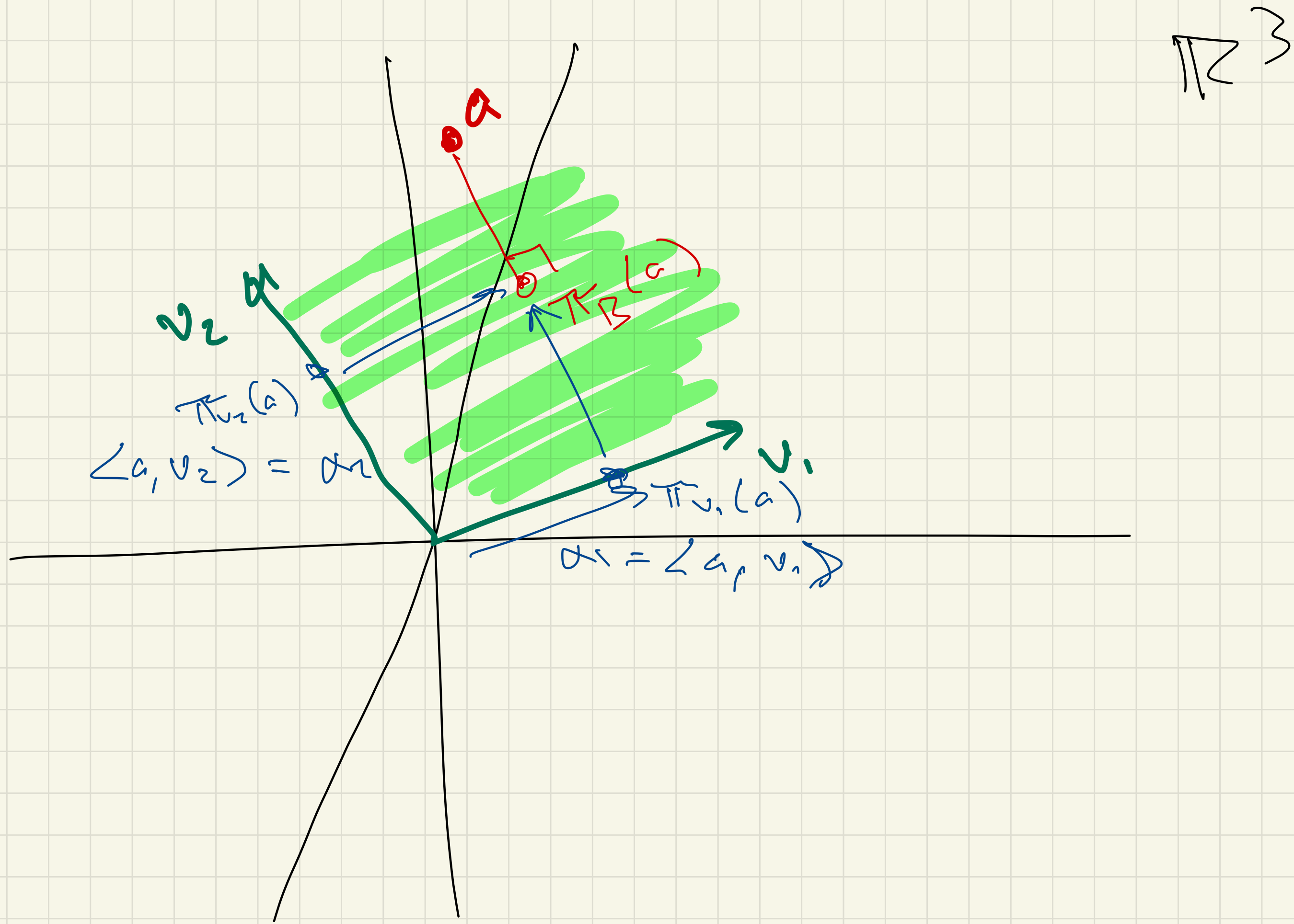
- $\|v_i\| = 1$
- $\langle v_i, v_j \rangle = 0 \quad j \neq i$
- For any $x \in B$ can write

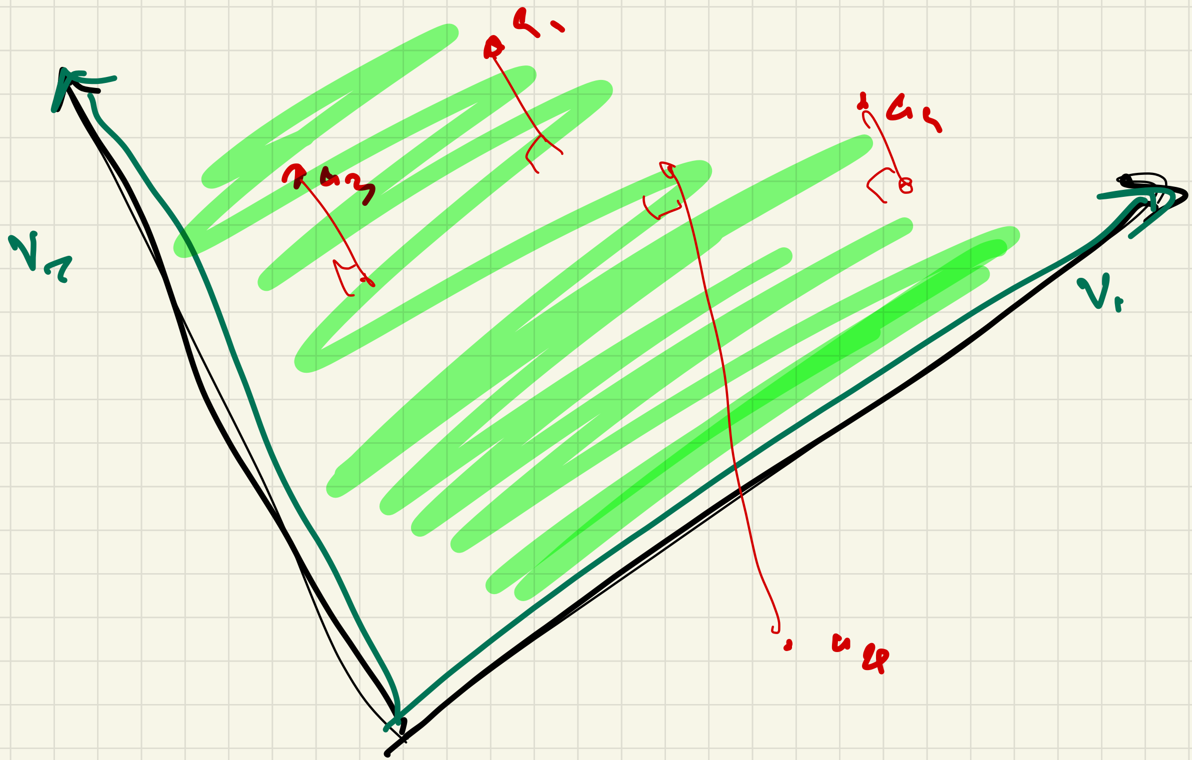
$$x = \sum_{j=1}^{k} \alpha_j v_j$$

α_j ← scalar

Projection onto B

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


$\mathbb{R}^3$
a
$\pi_B(a)$
v_2
α_2
$\pi_{v_2}(a)$
$\langle a, v_2 \rangle = \alpha_2$
v_1
$\pi_{v_1}(a)$
$\alpha_1 = \langle a, v_1 \rangle$


v2
v1