# L17: Projection onto Basis

Mar 16, 2026

FoDA

Jeff M. Phillips

# Dimensionality Reduction

Data $x \in \mathbb{R}^d$ $\quad d =$ very big $= 1000$ $\quad \Longrightarrow \quad x' \in \mathbb{R}^k$

$k =$ small

2, 3, 10

$(x_1, x_2, \dots, x_d)$ $\quad x \quad \Longrightarrow \quad x'$

$x' = (x_1', x_2', x_3')$

## Why?

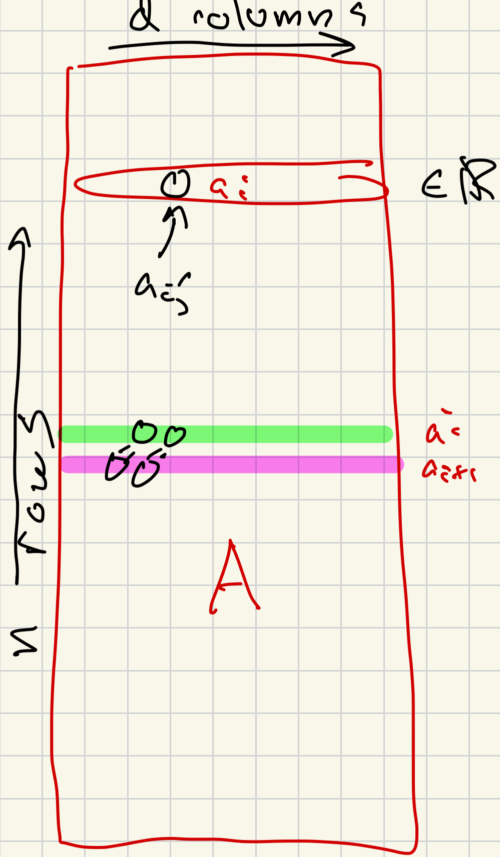
- easier to visualize? **Plot it** $k = 2$
- unneeded info in $\mathbb{R}^d$ "overfit"
- computationally: read 1 data point $O(d)$ time

curse-of-dimensionality. some algo $O(n^d)$ runtime

Input $A = \{a_1, a_2, \dots a_n\} \subset \mathbb{R}^d$

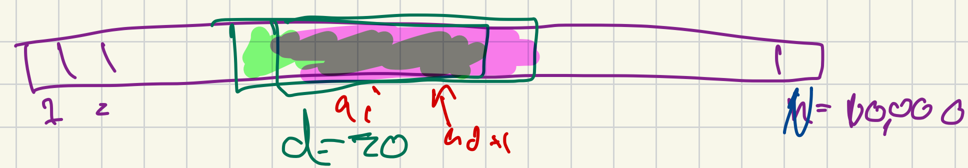
d = too big

still typically $n > d$




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

- word vectors: word2vec $a_i \in \mathbb{R}^{300}$
  $n = 100{,}000$ words in English
- $n = 40{,}000$ weather stations
  $d$ days of max temperature.
  $= 100$
- $N$ days of 1 stock price
  window $d$ days



In all examples each coordinate has the same units (meaning)!

---

if this ↑ not true

then do not apply Dimensionality Reduction.

# Projections

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

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

true

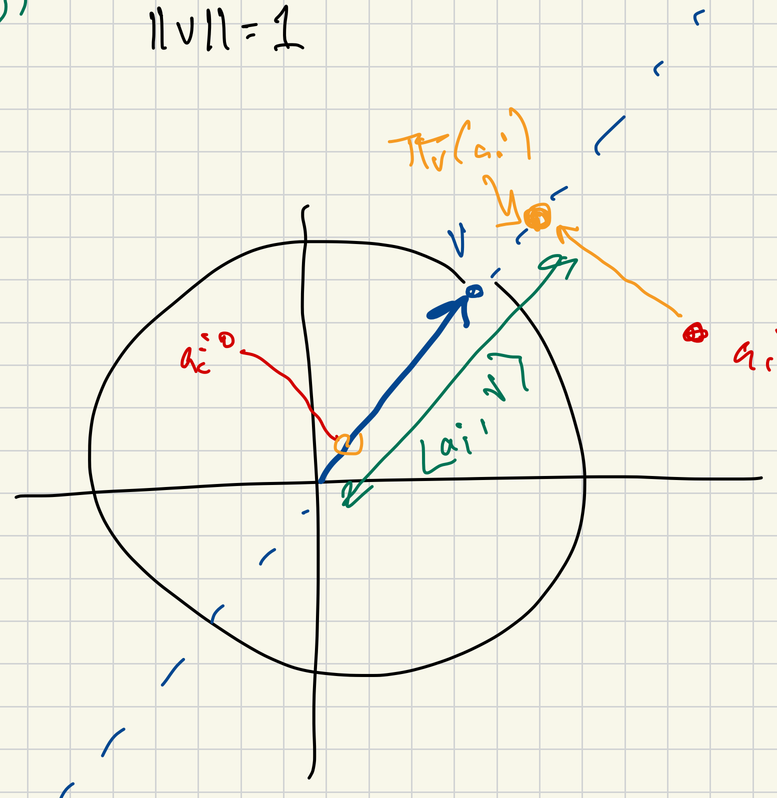
$\|v\| = 1$

dot product

$$\langle a_i, v \rangle = \sum_{j=1}^{d} (a_{ij} \cdot v_j)$$

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

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



# Subspace $B$

(assume contains the origin $0 \in \mathbb{R}^d$)

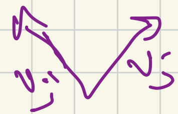
low-dimension

$k$

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

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

- $\|v_j\| = 1$ unit vectors
- $\langle v_j, v_{j'} \rangle = 0$ if $j \neq j'$




orthogonal

- For any $x \in B$ can write

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

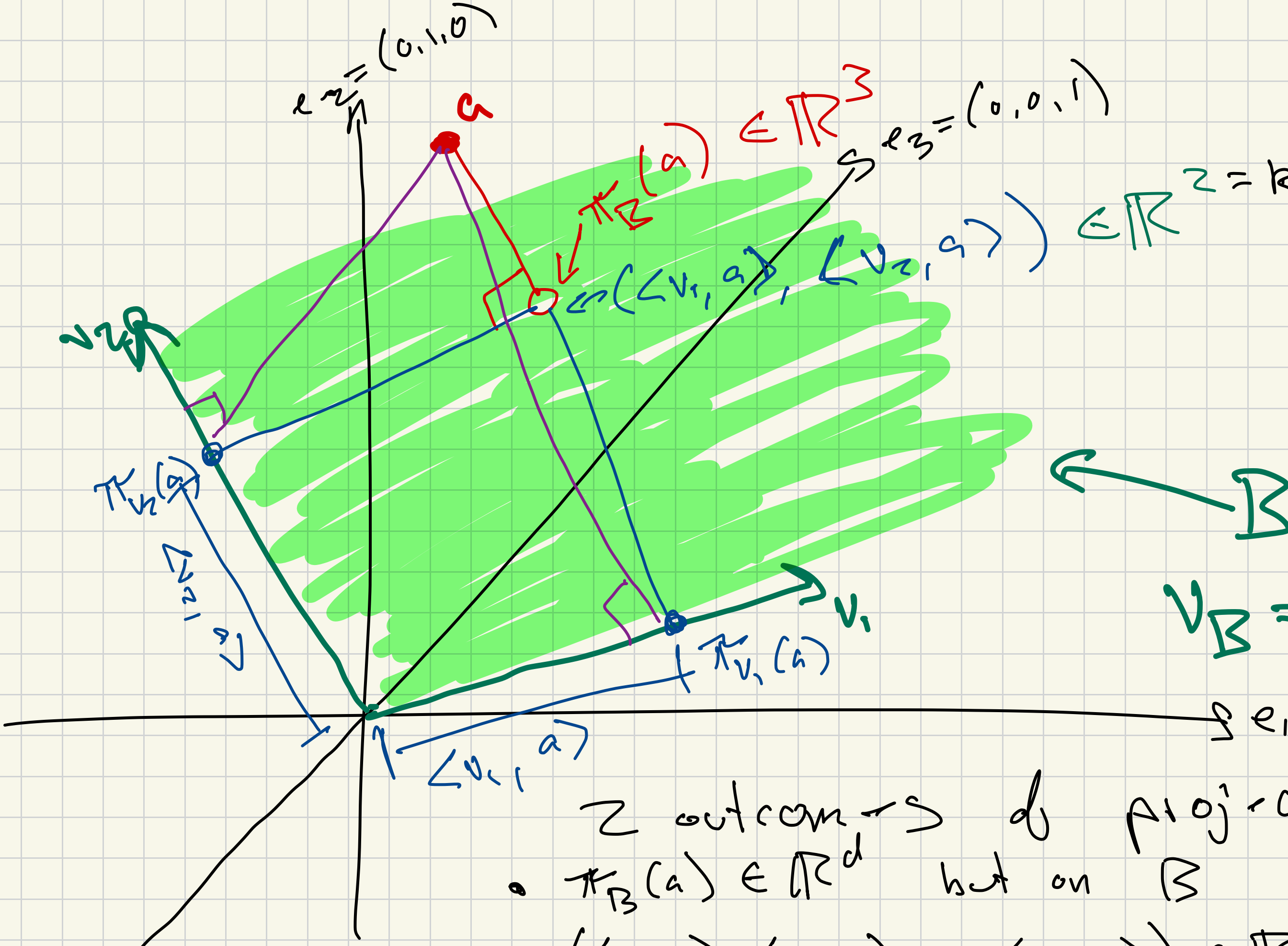
scalar

Projection onto $B$

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

closest point on $B$ to input $a$.

$\mathbb{R}^3 = d$




$V_B = \{v_1, v_2\}$

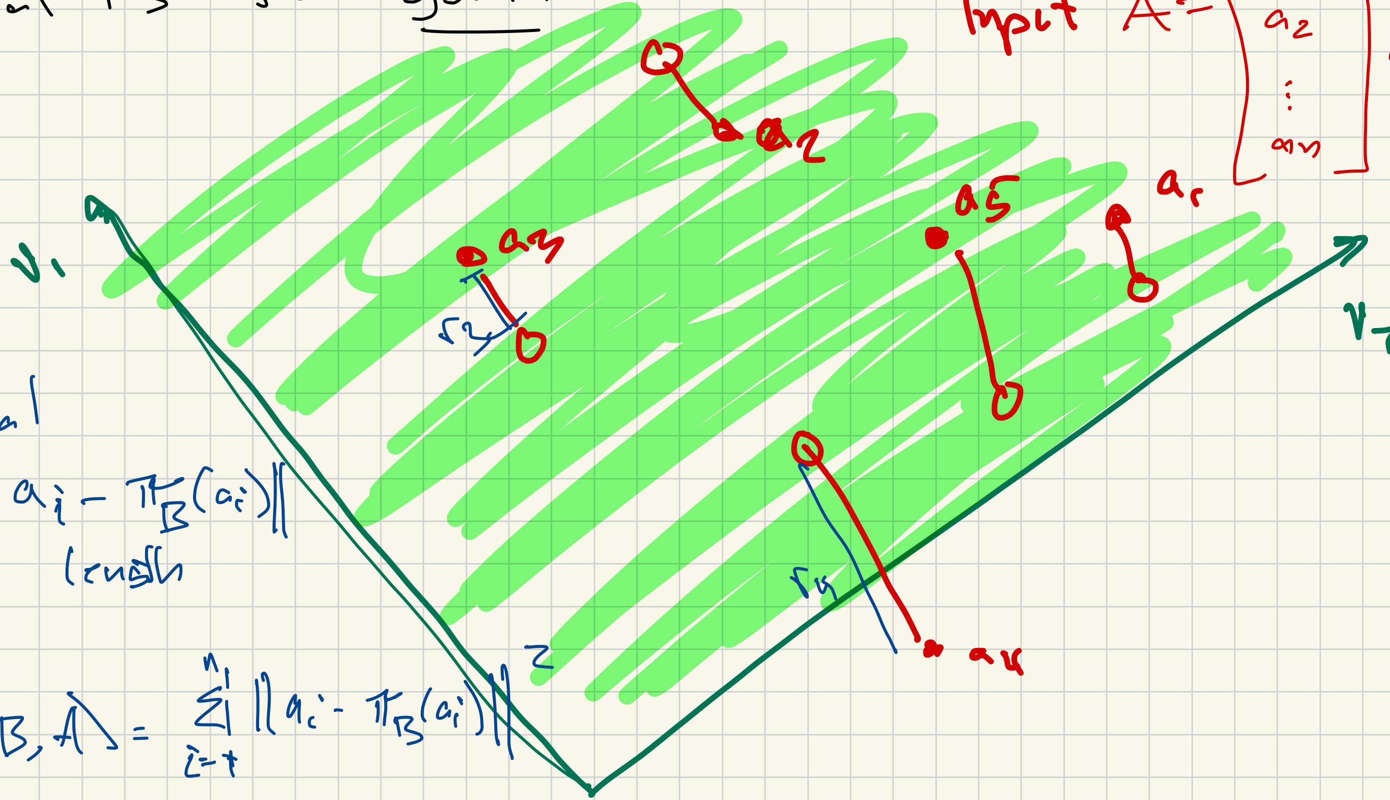
2 outcomes of projection

- $\pi_B(a) \in \mathbb{R}^d$ but on $B$
- $(\langle v_1, a\rangle, \langle v_2, a\rangle \dots \langle v_k, a\rangle) \in \mathbb{R}^k$

# How to choose basis $B$, $V_B = \{v_1, \dots, v_k\}$

What is the goal?

Input $A = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{R}^{n \times d}$




residual

$$r_i = \|a_i - \pi_B(a_i)\|$$

length

SSE

$$\text{cost}(B, A) = \sum_{i=1}^{n} \|a_i - \pi_B(a_i)\|^2$$