Dimensionality Reduction

L16: SVD and Relatives

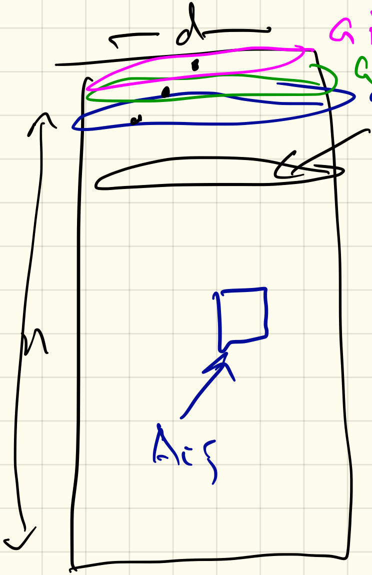
→ PCA

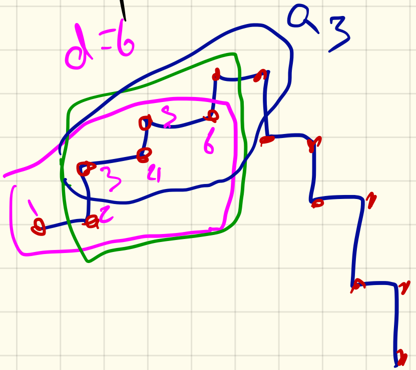
Jeff M. Phillips

March 23, 2020

Data Matrix

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





Using norms $\|a_i\|$

Euclidean distance

Every column

↳ has the same units

↳ representing same object

- n weather station

  d points in time

  ↳ temp

- n users A_{ij} rating movie

  d movies

- 1 stock (stock price)

  N days closing value

  $n = N-d$ a_i shingles

  d consecutive days

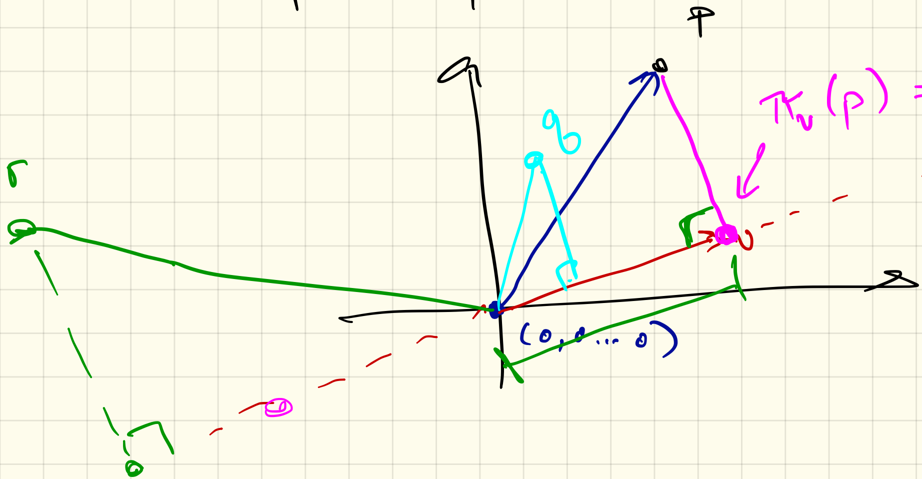
What to do if d is very large

Reduce $d \to k$ $\quad k \ll d$

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

$\|v\| = 1$

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



$\pi_v(p) = \langle v, p \rangle v$

$=$ closest point to p on 1-dim subspace through v (and 0)

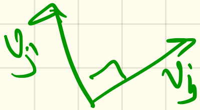
Basis B

represented as $V_B = \{v_1, v_2, .. v_k\}$

$v_j \in \mathbb{R}^d$, $\|v_j\| = 1$, $\langle v_j, v_{j'} \rangle = 0$

either

- $\pi_B(p) \in \mathbb{R}^d$



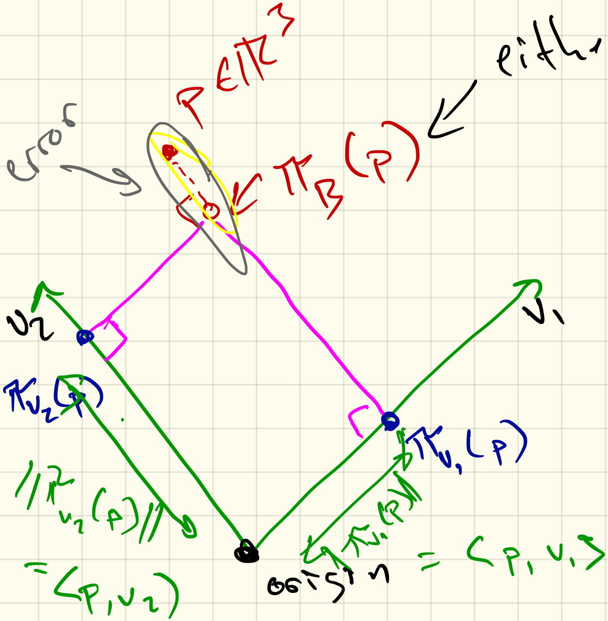
$d = 3$

$k = 2$

or

- $(\|\pi_{v_1}(p)\|, \|\pi_{v_2}(p)\|) \in \mathbb{R}^k$

$V_B = \{v_1, v_2\}$

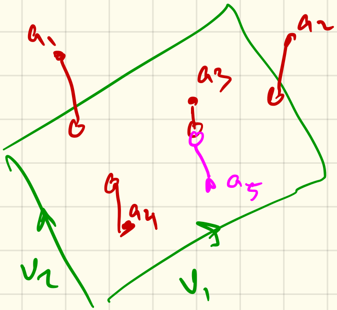


Sum of Squared Errors $SSE(A, B)$

$$SSE(A,B) = \sum_{a_i \in A} \| a_i - \pi_B(a_i) \|^2$$

Goal k-dimensional subspace B

$$B^* = \arg\min_B SSE(A,B)$$



B^* from SVD

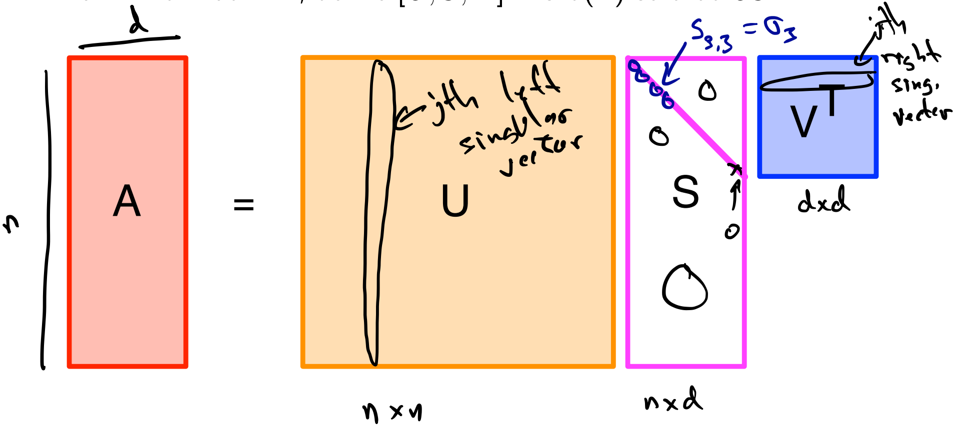
if enforce

B^* contains

0

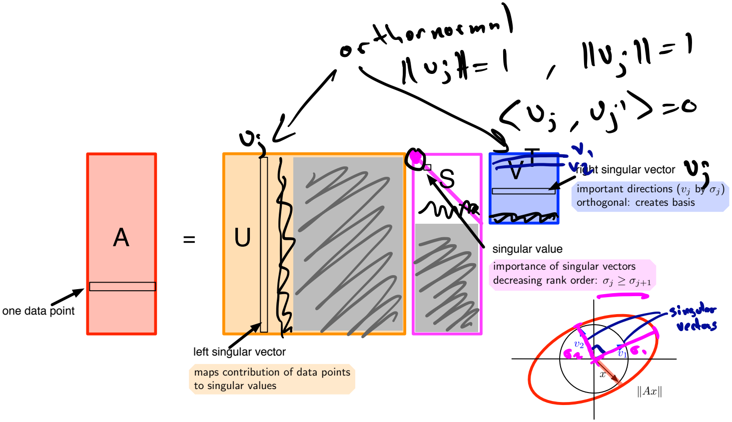
Singular Value Decomposition

For $n \times d$ matrix A, define $[U, S, V] = \text{svd}(A)$ so that $USV^T = A$.



Singular Value Decomposition

$\langle u_j , u_{j'} \rangle = 0$



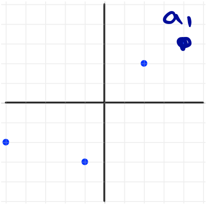
$\sigma_j \geq \sigma_{j+1} \; ... \; \sigma_d \geq 0$

Tracing a Point through SVD

$n = 4$
$d = 2$

Consider a matrix

$$A = \begin{pmatrix} 4 & 3 \\ 2 & 2 \\ -1 & -3 \\ -5 & -2 \end{pmatrix}, \quad a_1$$



and its SVD $[U, S, V] = \text{svd}(A)$:

σ_1 σ_2

$$U = \underbrace{\begin{pmatrix} -0.6122 & 0.0523 & 0.0642 & 0.7864 \\ -0.3415 & 0.2026 & 0.8489 & -0.3487 \\ 0.3130 & -0.8070 & 0.4264 & 0.2625 \\ 0.6408 & 0.5522 & 0.3057 & 0.4371 \end{pmatrix}}_{n \times n} S = \underbrace{\begin{pmatrix} 8.1655 & 0 \\ 0 & 2.3074 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}}_{n \times d} V = \underbrace{\begin{pmatrix} -0.8142 & -.5805 \\ -0.5805 & 0.8142 \end{pmatrix}}_{d \times d}$$

Tracing a Point through SVD

$\|x\| \approx 1$

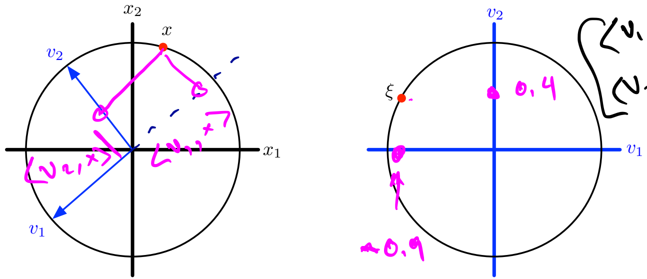
$Ax = USV^T x$

$V^T x$

$SV^T x$

$x = (0.243, 0.97)$, then what is $\xi = V^T x$?

$$(V =) V^T = \begin{pmatrix} -0.8142 & -.5805 \\ -0.5805 & 0.8142 \end{pmatrix}$$



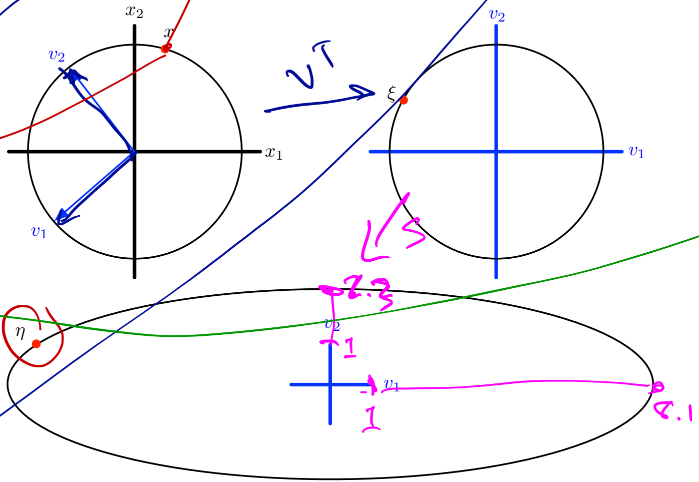
$$\begin{bmatrix} \langle v_1, x \rangle \\ \langle v_2, x \rangle \end{bmatrix} = \boxed{V^T} \boxed{x}$$

Tracing a Point through SVD

$x = (0.243, 0.97)$, then what is $SV^Tx = S\xi$?

$$S = \begin{bmatrix} 8.1 & 0 \\ 0 & 2.3\ldots \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$(V =)V^T = \begin{pmatrix} -0.8142 & -.5805 \\ -0.5805 & 0.8142 \end{pmatrix}$$

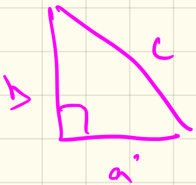


$$\sum_i \|a_i - \pi_B(a_i)\|^2$$

B^*: $V_B = \{v_1, \dots, v_k\}$

top k right sing. vectors

$a^2+b^2=c^2$



$$= \sum_i \Big\| \sum_{j=1}^{d} v_j \langle a_i, v_j\rangle - \sum_{j=1}^{k} v_j \langle a_i, v_j\rangle \Big\|^2$$

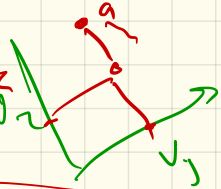
pythagoras thm

$$= \sum_i \Big\| \sum_{j=k+1}^{d} v_j \langle a_i, v_j\rangle \Big\|^2 = \sum_{i} \sum_{j=k+1}^{d} \| v_j \langle a_i, v_j\rangle \|^2$$

unit

scalar

$$= \sum_{i=1} \sum_{j=k+1} \langle a_i, v_j\rangle^2 = \sum_{j=k+1}^{d} \sigma_j^2$$

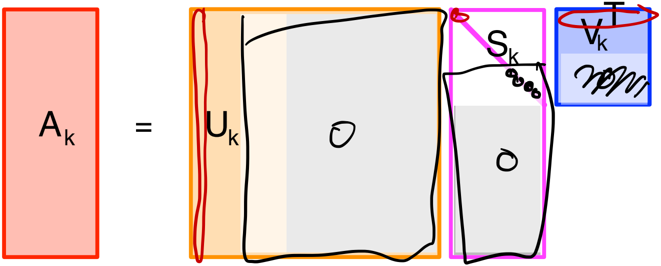


$$= \sum_{j=k+1}^{d} \Big(\sum_{i=1}^{n} \langle a_i, v_j\rangle^2 \Big) \Leftarrow \sum_{j=k+1}^{d} \underbrace{\|A v_j\|^2}_{\sigma_j^2} \Leftarrow \sum_{j=k+1}^{d} \sigma_j^2$$

Best Rank k-Approximation

Find A_k so $\text{rank}(A_k) \le k$

minimize $\|A - A_k\|_F^2$ or $\|A - A_k\|_2^2$



$$A_k = \sum_{j=1}^{k} \underbrace{u_j \sigma_j v_j^T}_{\in \mathbb{R}^{n \times d}}$$

v_1 chosen so $\|v_1\| = 1$

maximizes $\|A v_1\|^2$

then

v_2 chosen so $\|v_2\| = 1$, $\langle v_1, v_2 \rangle = 0$

maximizes $\|A v_2\|^2$