LADS 18

1

A is data matrix

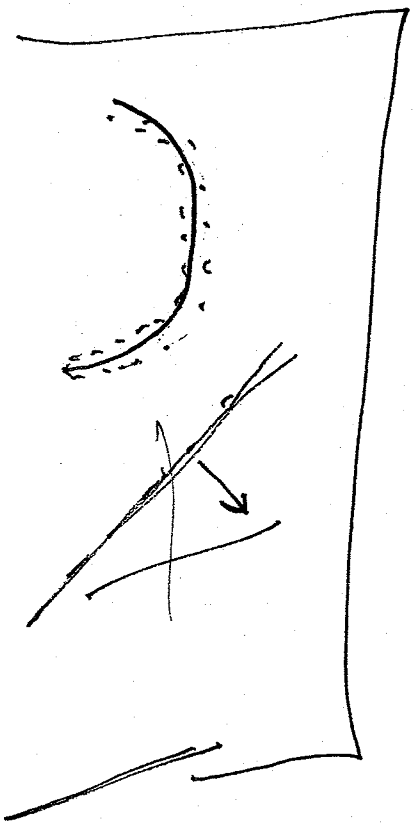
sample customer $\rightarrow$

$$\begin{array}{cc} & n \\ m & \begin{bmatrix} \longleftrightarrow \end{bmatrix} \end{array} = \begin{bmatrix} \vec{r}_1^{\,T} \\ \vdots \\ \vec{r}_n^{\,T} \end{bmatrix}$$

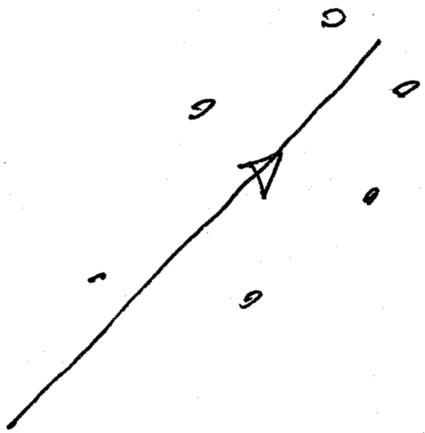
product features

Find a single #
that best describes a customer.

geometric problem



$\vec{r}_j$ are points in $\mathbb{R}^n$



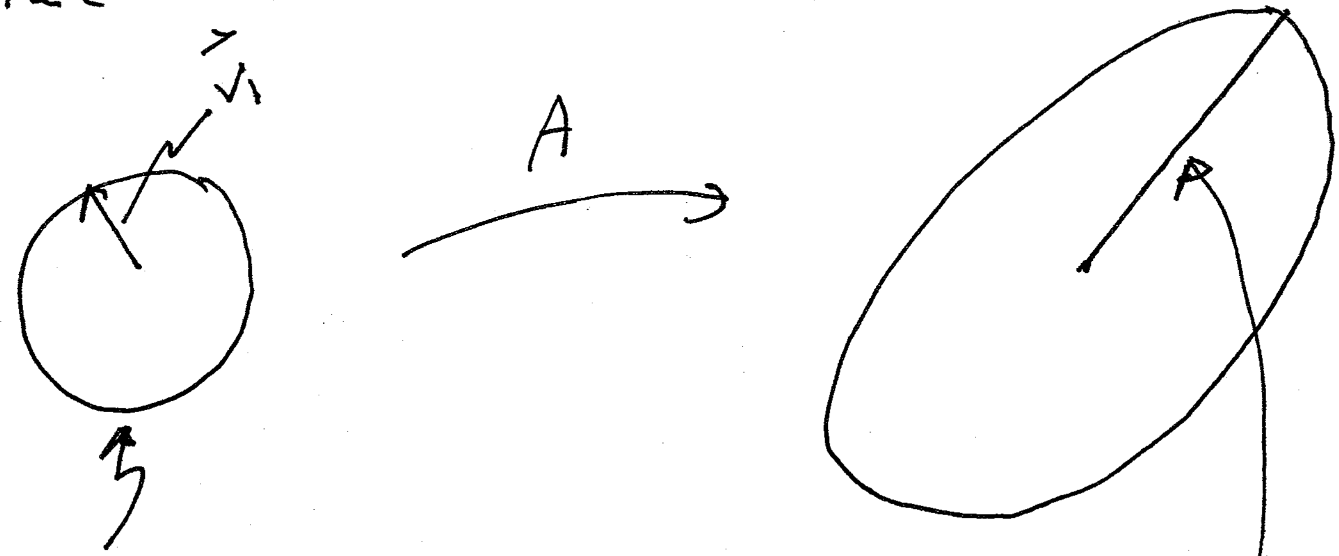
Find $\vec{w}$ (unit vector).
That maximizes the
sum of the the square
of the projection of the
$\vec{r}_j$ onto $\vec{w}$ OR

$$\arg\max \sum_{i=1}^{m} (\vec{r}_i^{\,T} \cdot \vec{w})^2 \doteq \underset{\|\vec{w}\|_2 = 1}{\arg\max} \|A\vec{w}\|_2^2$$

last
time

best number to categorize customer i's $\vec{r}_i \cdot \vec{w}$

Recall



$\|\vec{w}\|_2 = 1$

This vector is $A\vec{v}_1$

σ_2 is $\max\limits_{\|\vec{w}\|=1} \|A\vec{w}\|_2$ so

$\vec{v}_1 = \underset{\|\vec{w}\|=1}{\operatorname{Argmax}} \|A\vec{w}\|_2$

But needs proof

Theorem: If A has full svd

$$A = U \Sigma V^T \Rightarrow$$

1st right sing vect.

$$\vec{v}_1 = \underset{\|\vec{w}\|=1}{\operatorname{argmax}} \|A\vec{w}\|_2^2$$

$$\sigma_1^2 = \max_{\|\vec{w}\|=1} \|A\vec{w}\|_2^2$$

first (largest) sing value

Proof:

Recall $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is o.n. basis for $\mathbb{R}^n$

so we can write

$$\vec{w} = \sum_{i=1}^{n} \alpha_i \vec{v}_i$$ so now we seek $\vec{\alpha}$.

and since $\|\vec{w}\|^2 = 1$, $\boxed{\sum_{i=1}^{n} \alpha_i^2 = 1}$ (pythag)

compute $\|A\vec{w}\|_2^2 = \|U\Sigma V^T\vec{w}\|_2^2$

$$= \|\Sigma V^T \vec{w}\|_2^2 = \left\| \Sigma \begin{bmatrix} \vec{v}_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{bmatrix} \sum_{i=1}^{n} \alpha_i \vec{v}_i \right\|_2^2$$

U is unitary

diag. matrix in SVD

$\vec{v}_i$ are o.n. basis

$$= \left\| \Sigma \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} \right\|_2^2 = \left\| \begin{bmatrix} \sigma_1 \alpha_1 \\ \sigma_2 \alpha_2 \\ \vdots \\ \sigma_n \alpha_n \end{bmatrix} \right\|_2^2$$

matrix not sum

$$= \sum_{i=1}^{n} (\sigma_i \alpha_i)^2$$

Now problem is:

MAXIMIZE $\displaystyle\sum_{i=1}^{n} \sigma_i^2 \alpha_i^2 = \Psi(\vec{\alpha})$

Subject to $\displaystyle\sum_{i=1}^{n} \alpha_i^2 = 1$

(6)

Solve By Lagrange multiplier.

assuming $\sigma_1 > \sigma_2$

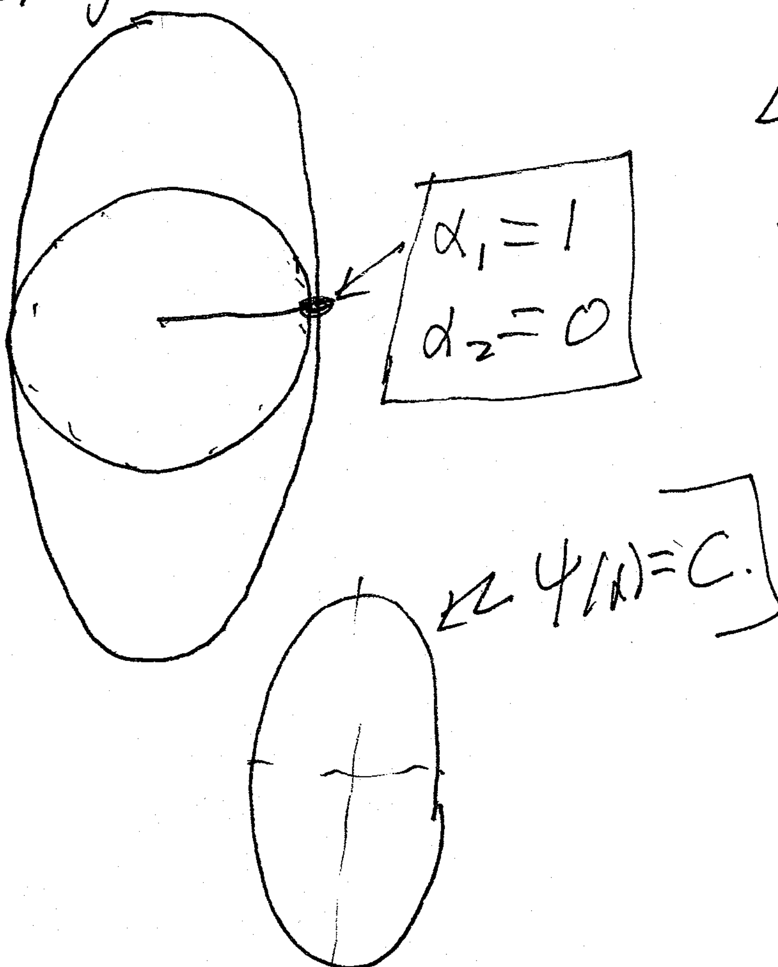
answ: $\alpha_1 = 1 \quad \alpha_i = 0 \quad i > 1$

So $\vec{w} = 1 \cdot \vec{v}_1 = \vec{v}_1$.

$n=2$ geometrically -

$$\psi(\vec{\alpha}) = \sigma_1^2 \alpha_1^2 + \sigma_2^2 \alpha_2^2$$

subject to $\alpha_1^2 + \alpha_2^2 = 1$ i.e. <u>unit circle</u>



Level sets of ψ

$$\sigma_1^2 \alpha_1^2 + \sigma_2^2 \alpha_2^2 = C$$

$$\frac{\alpha_1^2}{\frac{C}{\sigma_1^2}} + \frac{\alpha_2^2}{\frac{C}{\sigma_2^2}} = 1$$

<u>ellipses</u>

$$\frac{C}{\sigma_1^2} < \frac{C}{\sigma_2^2} \text{ since } \sigma_1 > \sigma_2$$

3rd optimization property of SVD

$A = U \Sigma V^T$ with r nonzero sing values

$$= \underbrace{\sigma_1 \vec{u}_1 \vec{v}_1^T}_{\text{rank 1 outer product}} + \cdots + \sigma_r \vec{u}_r \vec{v}_r^T$$

$\sigma_1 > \sigma_2 > \cdots$

$$A_i = \sum_{j=1}^{i} \sigma_j \vec{u}_j \vec{v}_j^T \Rightarrow \text{rank}(A_i) = i$$

but A_i is the best rank i approx to A.

Example

$$A = \begin{bmatrix} 3 & 0 \\ 4 & 5 \end{bmatrix} \quad U = \frac{1}{\sqrt{10}} \begin{bmatrix} 1 & -3 \\ 3 & 1 \end{bmatrix}$$

$$V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \quad \Sigma = \begin{bmatrix} \sqrt{45} & 0 \\ 0 & \sqrt{5} \end{bmatrix}$$

$$A = \sigma_1 \vec{u}_1 \vec{v}_1^{\,T} + \sigma_2 \vec{u}_2 \vec{v}_2^{\,T}$$

$$= \sqrt{45} \begin{bmatrix} 1/\sqrt{10} \\ 3/\sqrt{10} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} + \sqrt{5} \begin{bmatrix} -3/\sqrt{10} \\ 1/\sqrt{10} \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$= \sqrt{45} \begin{bmatrix} \frac{1}{\sqrt{20}} & \frac{1}{\sqrt{20}} \\ \frac{3}{\sqrt{20}} & \frac{3}{\sqrt{20}} \end{bmatrix} + \sqrt{5} \begin{bmatrix} \frac{+3}{\sqrt{20}} & \frac{-3}{\sqrt{20}} \\ -\frac{1}{\sqrt{20}} & \frac{1}{\sqrt{20}} \end{bmatrix}$$

$$= \underbrace{\frac{3}{2}\begin{bmatrix} 1 & 1 \\ 3 & 3 \end{bmatrix}}_{A_1} + \frac{1}{2}\begin{bmatrix} 3 & -3 \\ -1 & 1 \end{bmatrix} = A$$

(11)

<u>Thm</u> If if rank $(B) = i$

$$\Rightarrow \|B - A\|_2 \geq \|A_i - A\|_2 = \sigma_{i+1}$$

$$\|B - A\|_F \geq \|A_i - A\|_F = \sqrt{\sigma_{i+1}^2 + \cdots + \sigma_r^2}$$

<u>OR</u>

$$\min_{\operatorname{rank}(B)=i} \|B - A\|_2 = \|A_i - A\|_2 = \sigma_{i+1}$$

$$\min_{\operatorname{rank}(B)=k} \|B - A\|_F = \|A_i - A\|_F = \sqrt{\sigma_{i+1}^2 + \cdots + \sigma_r^2}$$

SVD and Image compression

8 bit

- A grey scale 400×600 image is a matrix A that is 400×600 and each entry is an integer between 0 and 255. (or scaled 0 to 1)

black white

- Compute the SVD $A = U \Sigma V^T$ and

So $A = \sum_{i=1}^{400} \sigma_i \vec{u}_i \vec{v}_i^{\,T}$

- Often for $N << 400$

$$A_N = \sum_{i=1}^{N} \sigma_i \vec{u}_i \vec{v}_i^{\,T}$$

when rendered is a good version of A and requires much less storage space.

MATLAB Demo