LADS 16

SVD – various forms

U, V orthogonal

$- A = U \Sigma V^T \Rightarrow \Sigma$ is generalized diagonal

(1)

· FULL S.V.D.

$m = n$: $\quad \begin{bmatrix} \; \end{bmatrix} = \begin{bmatrix} U \end{bmatrix} \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_n \end{bmatrix} \begin{bmatrix} V^T \end{bmatrix}$ all $n \times n$

$m < n$: $\quad \underset{m \times n}{\begin{bmatrix} \; \end{bmatrix}} = \underset{m \times m}{\begin{bmatrix} \; \end{bmatrix}} \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_n & \end{bmatrix} \underset{n \times n}{\begin{bmatrix} \; \end{bmatrix}}$

← rows added to make ~~orthogonal~~ orthonormal

$m > n$: $\quad \begin{bmatrix} \; \end{bmatrix} = \underset{m \times m}{\begin{bmatrix} \; \end{bmatrix}} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ \hline & 0 & \end{bmatrix} \underset{n \times n}{\begin{bmatrix} \; \end{bmatrix}}$

↑ columns added to make orthonormal

Reduced SVD (various versions)

Strang's version

Let A have r nonzero singular values. NOTE: $r \leq \min(m,n)$

Rewrite SVD as $\underline{AV = U\Sigma}$

$$\boxed{A\vec{v}_L = \sigma_L \vec{u}_L}$$

with $V = [\vec{v}_1 \ldots \vec{v}_n]$

$U = [\vec{u}_1 \ldots \vec{u}_m]$

$\Sigma = \text{diag}(\sigma_1, \ldots, \sigma_n)$.

only nonzero σ_L are important.

if $\sigma_{r+1} = 0$

$$A\vec{v}_{r+1} = 0 \cdot \vec{u}_{r+1} = \vec{0}$$

$\vec{v}_{r+1} \in \text{Null}(A)$.

So $A\vec{v}_i = \sigma_i \vec{u}_i$ for $i = 1, \ldots, r$

In matrix form.

$$A \underset{V_r}{\begin{bmatrix} \vec{v}_1 \cdots \vec{v}_r \end{bmatrix}} = \underset{U_r}{\begin{bmatrix} u_1 \cdots \vec{u}_r \end{bmatrix}} \underset{\Sigma_r}{\begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{bmatrix}}$$

but V_r and U_r are not square usually.

but they still have orthonormal col.

so $V_r^T V_r = I_r \quad U_r^T U = I_r$

but usually

$$V_r V_r^T \neq I_r \quad \text{and} \quad U_r U_r^T \neq I_r$$

so V_r^T is a one-sided inverse.

U_r^T " " " " .

But $A = U_r \Sigma_r V_r^T$

(Trefethen + Bau , Golub)

Thin, reduced or economy

$m \geq n$

$$\begin{array}{c} \\ m \end{array} \begin{bmatrix} \quad & \quad \\ & \\ & \end{bmatrix} = \begin{array}{c} \\ m \end{array} \overset{n}{\begin{bmatrix} | & | & | & | \\ | & | & | & | \\ | & | & | & | \end{bmatrix}} \; \begin{array}{c} n \end{array} \overset{n}{\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix}} \; \begin{array}{c} n \end{array} \overset{n}{\begin{bmatrix} & V^T & \end{bmatrix}}$$

$\hat{U}$

allowing 0 sing values

orthogonal.

Be careful with built-in SVD's

eg matlab has $[u\,s\,v] = svd(A)$

$[\cancel{U}\, S\, V] = svd(A, \text{'econ'})$

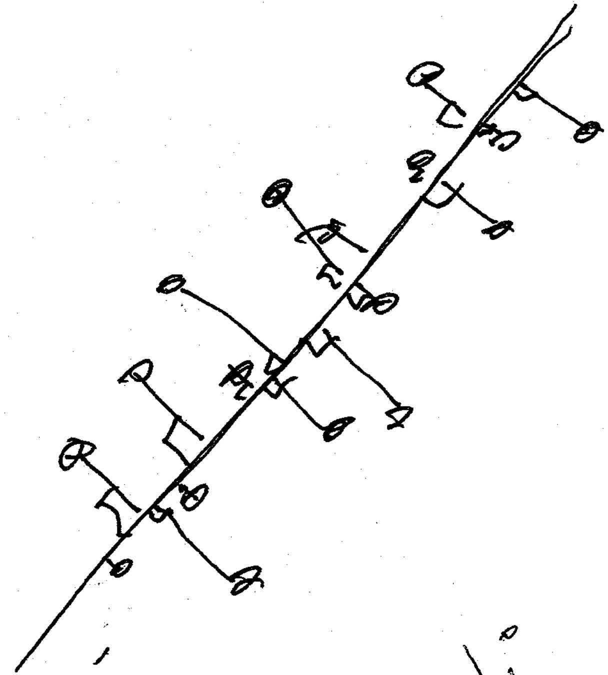
- Most D.S. problems are high dimensional and there are many methods to reduce the dimensions in order to analyze and understand
- The start is the SVD.

Three optimization properties of the SVD.

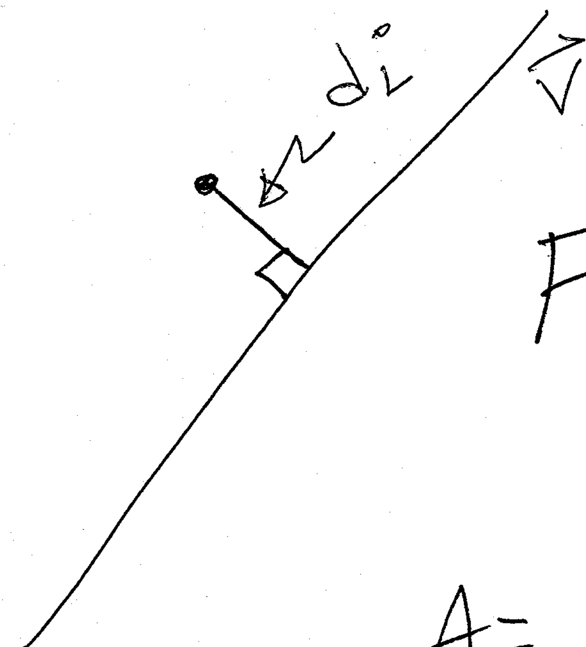
$\boxed{\text{First informally}}$ → stated geometrically.

(2)

(1)



Find best line approx the data



direction of line

Find $\vec{v}$ that minimizes

$\sum d_i^2$ – Ans

A = matrix of data

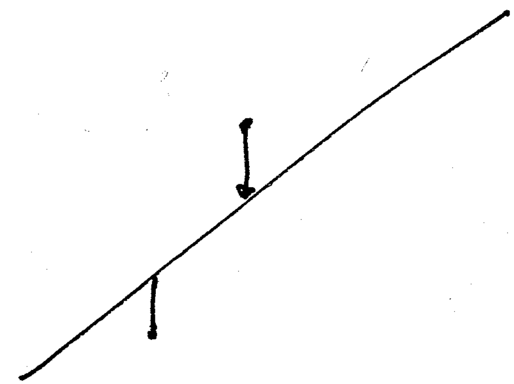
$v = \vec{v}_1$ minimizes

first right singular vector

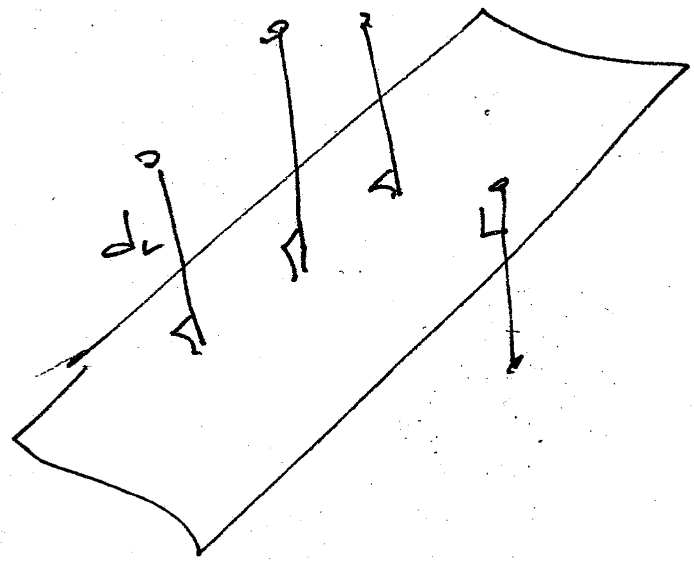
CONTRAST TO Least squres fit

(x_i, y_i)

$$\sum_i (m x_i + b - y_i)^2 \text{ minimized}$$

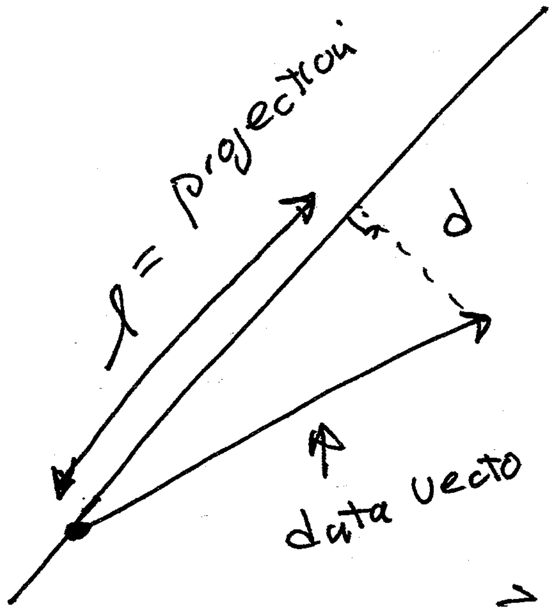


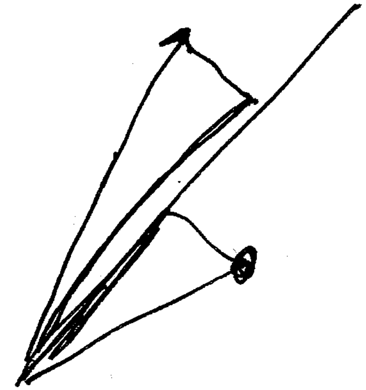
(1) cont Find Subspace V of dim k that minimizes the sum of the squared perp. distances to V – now data in $\mathbb{R}^n$, $n >> k$



Answer = $\mathrm{span}\{\vec{v}_1, \ldots, \vec{v}_k\}$.

(2)





Find direction $\vec{v}$ that maximizes the sum of the squared projections = PCA = Principal Component Analysis

Answer is $\vec{v}_1$, first right sing vect.

Now find a k-dim subspace that maximizes the sum of the squared lengths of projections of the data i.e. Find best lower dimensional approximation of the data (in this sense) - Answer is Span$\{\vec{v}_1, .., \vec{v}_k\}$

(3) Find a rank k matrix that best approximates A:

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

we can use col. x rows to write

so

$$A = \sigma_1 \underbrace{u_1 v_1^T}_{\substack{\text{outer products}\\ \text{rank 1.}}} + \sigma_2 u_2 v_2^T + \dots + \sigma_r u_r v_r^T$$

$$\sigma_1 \ge \sigma_2 \ge \dots \ge \sigma_r > 0$$

$\boxed{k << r}$

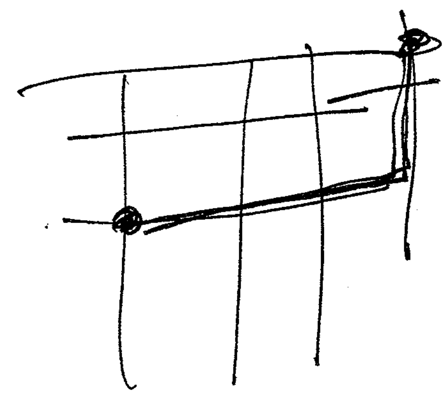
Answ to the approx problem is

$$A_k = \sigma_1 u_1 v_1^T + \dots + \sigma_k u_k \vec{v_k}$$

$\boxed{\text{Compression}}$

What does approximate mean?
how do you measure how close vectors are and matrices are?

VECTOR and matrix norms.



$\ell^2 \quad \|\vec{v}\|_2 = \sqrt{|v_1|^2 + \dots + |v_n|^2}$

$\ell^1 \quad \|\vec{v}\|_1 = |v_1| + \dots + |v_n|$

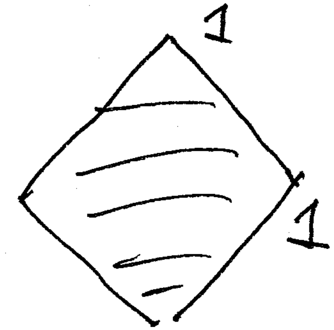
$\ell^\infty \quad \|\vec{v}\|_\infty = \max\{|v_1|, \dots, |v_n|\}$

$\ell^p \quad \|\vec{v}\|_p = \left(\sum |v_i|^p\right)^{1/p}$

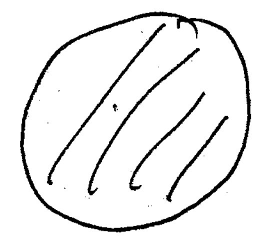
$1 < p < \infty$

Unit balls $= \{\vec{x} : \|\vec{x}\| \leq 1\}$.

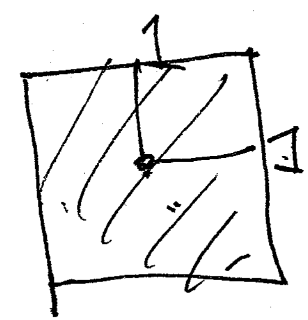
$\ell^1 \quad |x_1| + |x_2| \leq 1$



$\ell^2 \quad (|x_1|^2 + |x_2|^2)^{1/2} \leq 1$



$\ell^\infty \quad \max\{|x_1|, |x_2|\} \leq 1$



how close is $\vec{x}$ to $\vec{y}$?

$\|\vec{x} - \vec{y}\|$ in some <u>norm</u>

14

All norms satisfy

$$\|\alpha\vec{v}\| = |\alpha|\,\|\vec{v}\|$$

$$\|\vec{v}+\vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$$

$\|\vec{v}\| \geq 0$ $\quad \|\vec{v}\| = 0$ only when $\vec{v} = \vec{0}$