25.02.2020 Decomposition and all that

- Shift from prob. to vector spaces, decomposition of vector spaces.
- The best supspaces are observed
- Spectrum tells you all about these best-fit subspaces.
- Matrices are linear operators of vectors space.
- Vector spaces can be functions. - no norm.

$A v = \lambda v$.
└ problem arise when A is not square-matrix.

$As = \lambda s$. — right singular vectors
\ left singular "

$Ab = \delta u$ singular vectors are unit vectors.

$u^T A u = \delta u^T u$ — >

* If v is a singular vector of $A \Leftrightarrow v$ is an eigenvector of $A^T A$.

$$A^T A v = \sigma A^T v = \sigma (u^T A)^T = \sigma^2 (v^T)^T = \sigma^2 v$$

$$(A^T A) v = \lambda v$$

$$v^T (A^T A) v = \|Av\|^2 = \lambda \|v\|^2 \quad \lambda > 0$$

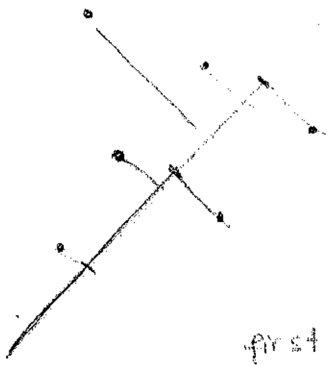
$\{u_i\}$ $u_i = \dfrac{A v_i}{\sqrt{\lambda_i}}$, $\sqrt{\lambda_i} = \sigma_i$ $A v_i = \sigma_i u_i$, $u_i^T A = \sigma_i v_i^T$

$A^T A$ — symmetric hence normal

they will commute

Normal matrices $\rightarrow$ 1. Eigen values are real

2. Set of eigen vectors are orthonormal



A: defn of these points

distance $\rightarrow$ minimized

Projections $\rightarrow$ maximized

$$v^* = \underset{v;\|v\|=1}{\arg\max} \|Av\|_2$$

first singular eigenvector of A

$$u_i^T u_j = \delta_{ij}$$

$v = \sum_i \alpha_i v_i \qquad \{\sigma_i\}\ \{v_i\}\ \{u_i\}$

$\|v\| = 1 \Rightarrow \sum_i \alpha_i^2 = 1$

$\|Av\| = ? \qquad Av = A\left(\sum_i \alpha_i v_i\right) = \sum_i \alpha_i A v_i = \sum_i \alpha_i \sigma_i u_i$

$$\|Av\|^2 = \underbrace{(Av)^T(Av) = \left(\sum_i \alpha_i \sigma_i u_i^T\right)\left(\sum_j \alpha_j \sigma_j u_j\right) = \sum_i \alpha_i^2 \sigma_i^2}_{\text{convex combination of } \sigma_i^2}$$

convex combination $\sum \lambda > 0$ which implies it is maximized

$\sum \lambda = 1$ when you take $\max_{\|v\|=1} \|v^T A\| = v_1$

If I take 1D space, it's the largest singular value

Original k-clustering problem: $\boxed{\text{it is max as the summation of } k \text{ eigenvalues.}}$?

Why max $\|Av\|$? Same with min(error)

v_1 (spanning) V_1. $\underline{V_2}$: (v_1, v_2).

best fit: min error.
max the projection

$$v_1 \equiv \underset{v: \|v\|=1}{\operatorname{argmax}} \|Av\|$$

$$v_2 \equiv \underset{v: \|v\|=1,\ v \perp v_1}{\operatorname{argmax}} \|Av\|$$

Suppose V_2 is not the best fit.

$$v_k \equiv \underset{v \perp v_1, \dots v_{k-1} : \|v\|=1}{\operatorname{argmax}} \|Av\|$$

V_2' is the best fit. $V_2' = (v_1', v_2')$

$\|Av_1'\|^2 \leq \|Av_1\|^2$ Choose v_2' which is perpendicular to v_1.
$(v_2' \perp v_1)$

$+ \qquad +$

$\|Av_2'\|^2 \leq \|Av_2\|^2$

So, V_2 is at least as good as V_2'. However, we know

V_2 is not the best fit.

(Use induction)

We can extend this argument by choosing $v_k \equiv \underset{v \perp v_1 \dots v_k : \|v\|=1}{\operatorname{argmax}} \|Av\|$

(Use induction)

why do we choose $v_2' \perp v_1$? Because we want to find 2D space.

<u>Claim</u>: v_1 and v_2 are indeed singular vectors.

$$v_k \equiv \sum_{i : v_i \perp (v_1 \dots v_{k-1})} \alpha_i^2 \sigma_i^2$$

$$v_1 = \underset{v: \|v\|=1}{\operatorname{argmax}} \underbrace{\|Av\|}_{|\sigma_1|}$$

$$v_2 = \underset{v: \|v\|=1,\ v \perp v_1}{\operatorname{argmax}} \overbrace{\|Av\|}^{|\sigma_2|}$$

σ_1 is the max. of larger set compared to σ_2.
So,

$$\sigma_1 \gg \sigma_2 \gg \sigma_3 \dots \gg \sigma_k \gg \sigma_{k+1}$$

- Square of Singular values of A which are the ~~squa~~ eigenvalues of $A^T A$.

Generalized (for Spaces) Eigenspace Decomposition Theorem

$T: V \longrightarrow W$ $\quad \lambda_1^{n_1}, \lambda_2^{n_2} \dots \lambda_k^{n_k}$; $n_1 + n_2 + \dots + n_k = \dim(V)$

$(\mathbb{R}^+) \to (\mathbb{R}^+)$

consider: $m=5,\ d=10$

$(A^T A)_{10 \times 10}$ $\quad \underline{\underline{\sigma_1^2}} \dots \underline{\underline{\sigma_{10}^2}}$

10 eigen values.

$E_1 \oplus E_2 \oplus E_3 + \dots \oplus E_k = \mathbb{R}^n$

$E_i = \mathrm{Kern}(T - \lambda I)^{n_i}$

A – generic.

A (in previous): number of points

$Av = \lambda v \Rightarrow (A - \lambda I) v_1 = 0$

$\Rightarrow v < \mathrm{Kern}(A - \lambda I)$

$\lambda: v_1, v_2 \in \mathrm{Kern}(A - \lambda I)^2$

$\vdots$

$v_1, \dots v_k \in \mathrm{Kern}(A - \lambda I)^n$

$(\sigma_1 \gg \dots \sigma_k \gg \sigma_{k+1})$ does not create any problems.

(SVD) Singular (for matrices) Decomposition Theorem: (Eigen – Decomposition Thm).

$$A = \sum_i \underbrace{u_i}_{m \times 1} \sigma_i \underbrace{v_i^T}_{1 \times d}$$

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

the columns: left singular eigen-vectors

diag matrix.

the rose of V: transpose of the right eigen vectors

Prove $\longrightarrow$ HW!

$$Av = \sum_i u_i \sigma_i v_i^T v, \quad \forall v.$$

(any v has a component in singular subspace and has a component orthogonal to subspace)

Fröbenius Norm

$$\|A\|_F = \sqrt{\sum_{\substack{1\leq i\leq m \\ 1\leq j\leq n}} a_{ij}^2} \qquad \|A+B\|_F \leq \|A\|_F + \|B\|_F$$

Ex: . $A v_i = \lambda v_i$, $A = V \Lambda V^T$

$v \xrightarrow{\pi} \{v_1, \ldots, v_n\}$ using A ?

Remind: Properties of Projection

1. $\pi = \pi^k$, $k \geq 1$

So, $A^2 = V \Lambda V^T \cdot V \Lambda V^T$

$$V \Lambda^2 V^T = A^2$$

$$A v = A \sum_i \alpha_i v_i = \sum_i \alpha_i \lambda_i v_i$$