Lesson 19

Read ch 7, 8

Polynomial spaces

Definition of SVD, First examples

Recall from last time

$$\mathbb{R}[x]_{\le d} \approx \mathbb{R}^{d+1}$$

$$a_0 + a_1 x + \cdots + a_d x^d \to \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_d \end{bmatrix}$$

$\mathbb{R}[x] \approx ?$ No finite basis

So we will focus on $\mathbb{R}[x]_{\le d}$

$$T \quad \mathbb{R}[x]_{\le 1} \to \mathbb{R}[x]_{\le 2}$$

$$T(p(x)) = \int_0^x p(t)\,dt$$

Question: what is the matrix of T?

What is the metrix of T ?

Answer 1

$$\begin{array}{ccc} T \quad \mathbb{R}[x]_{\leq 1} & \longrightarrow & \mathbb{R}[x]_{\leq 2} \\ p = ax+b & \longrightarrow & q = cx^2+dx+e \\ \downarrow & & \downarrow \\ S \quad \mathbb{R}^2 & \longrightarrow & \mathbb{R}^3 \\ \begin{bmatrix} a \\ b \end{bmatrix} & & \begin{bmatrix} c \\ d \\ e \end{bmatrix} \end{array}$$

$$M \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} e \\ d \\ c \end{bmatrix}$$

$$M = \begin{bmatrix} S\begin{bmatrix} 1 \\ 0 \end{bmatrix} & S\begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1/2 \end{bmatrix}$$

$$1 \longrightarrow \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad T(1) = x \longrightarrow \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$x \longrightarrow \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad T(x) = \frac{x^2}{2} \longrightarrow \begin{bmatrix} 0 \\ 0 \\ 1/2 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ b \\ 1/2\, a \end{bmatrix},$$

tells us $\displaystyle\int_0^x b + at\, dt = bx + \frac{1}{2}ax^2$

Answer 2: Fix bases $B_1 = 1, x$ in $R[x]_{\le 1}$ and $B_2 = 1, x, x^2$ in $R[x]_{\le 2}$ then we want M s.t.

$$T: R[x]_{\le 1} \to R[x]_{\le 2}$$

$$M[p]_{B_1} = [T(p)]_{B_2}$$

$ax + b = b \cdot 1 + a \cdot x$ so $[ax+b]_{B_1} = \begin{bmatrix} b \\ a \end{bmatrix}$

$cx^2 + dx + e = e \cdot 1 + d \cdot x + c \cdot x^2$ $[cx^2 + dx + e]_{B_2} = \begin{bmatrix} e \\ d \\ c \end{bmatrix}$

so $M = \begin{bmatrix} [T(1)]_{B_2} & [T(x)]_{B_2} \end{bmatrix}$

$T(1) = x$ $T(x) = \frac{1}{2}x^2$

$[x]_{B_2} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ $[\frac{1}{2}x^2]_2 = \begin{bmatrix} 0 \\ 0 \\ 1/2 \end{bmatrix}$

$$M = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1/2 \end{bmatrix}$$

In generel Given $T \quad V \longrightarrow W$

$B_1 = b_1 \cdots b_m$ a basis for V

$B_2 = c_1 \cdots c_n$ a basis for W

The matrix of T is $\begin{bmatrix} [T(b_1)]_{B_2} \cdots [T(b_m)]_{B_2} \end{bmatrix}$

and $M [v]_{B_1} = [T(v)]_{B_2}$

Wronskian: $V = C^{\infty}[a, b]$ not finite dimensional

$f_1 \ldots f_n$ are linearly dependent if there are constants $c_1 \ldots c_n$ not all equal to 0 s.t $c_1 f_1 + c_2 f_2 + \cdots + c_n f_n = \vec{0}$

$\vec{0}$ is the 0 function i.e $\vec{0}(x) = 0$ for all $x \in [a, b]$. So when $f_1 \ldots f_n$ are dependent we have

$\forall x \in [a\ b] \quad c_1 f_1(x) + \cdots + c_n f_n(x) = 0$

Differentiation is a linear transformation therefore if

$$c_1 f_1 + \cdots + c_n f_n = \vec{0} \quad \text{then}$$
$$c_1 f_1' + \cdots + c_n f_n' = \vec{0}$$
$$c_1 f_1^{(2)} + \cdots + c_n f_n^{(2)} = \vec{0} \quad \text{and so on}$$

so
$$\begin{bmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & & f_n'(x) \\ \vdots & \vdots & & \vdots \\ f_1^{(n-1)}(x) & f_2^{n-1}(x) & & f_n^{n-1}(x) \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

for all x in $[a, b]$; this tells us that $w(x) = \det \begin{bmatrix} f_1(x) \cdots f_n(x) \\ f_1^{(n-1)}(x) \cdots f_n^{(n-1)}(x) \end{bmatrix} = 0$ for all $x \in [a, b]$

Chapter 8

Chapter plan

$A = U \Sigma V^T$ singular values decomposition of A

1) Understand U, Σ, V^T
2) Reduced SVD and rank 1 decomposition of A
3) Geometry of SVD
4) Rank k approximation / compression
5) Best fit k-planes (PCA)

Th: any matrix $A \in \mathbb{R}^{m \times n}$ can be factored as

$$\underset{m\times n}{A} = \underset{m\times m}{U} \; \underset{m\times n}{\Sigma} \; \underset{n\times n}{V^T}$$ singular value decomposition

$U = \begin{bmatrix} u_1 \cdots u_m \\ \end{bmatrix}$ $\quad V = \begin{bmatrix} v_1 \cdots v_n \\ \end{bmatrix}$ Orthogonal i.e columns form an orthonormal set

Σ "diagonal" i.e

$\Sigma = (s_{ij})$ $\quad s_{ij} = 0$ if $i \neq j$ $\quad s_{ii} = \sigma_i$ for $i = 1 \cdots r$

$r = \text{rank } A$, $\quad \sigma_{ii} = 0$ for $i > r$ $\quad \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$

σ singular values

Us are left singular vectors $\quad V$ right singular vectors

---

Ex $\quad A$ 2×3 of rank 2 $\quad \Sigma = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \end{bmatrix}$

$\quad A$ 3×2 of rank 2 $\quad \Sigma \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{bmatrix}$

Ex :

$$A=\begin{pmatrix}1&2\\1&2\\1&2\end{pmatrix}=\begin{array}{c}\begin{array}{ccc}u_1&u_2&u_3\end{array}\\\begin{pmatrix}1/\sqrt{3}&1/\sqrt{2}&1/\sqrt{6}\\1/\sqrt{3}&0&-2/\sqrt{6}\\1/\sqrt{3}&-1/\sqrt{2}&1/\sqrt{6}\end{pmatrix}\end{array}\begin{array}{c}\sigma_1\quad\quad\\\begin{pmatrix}\sqrt{15}&0\\0&0\\0&0\end{pmatrix}\end{array}\begin{array}{c}\begin{bmatrix}1/\sqrt{5}&2/\sqrt{5}\\-2/\sqrt{5}&1/\sqrt{5}\end{bmatrix}\end{array}\begin{array}{c}v_1\\v_2\end{array}$$

3×2 $\qquad$ 3×3 $\qquad$ 3×2 $\qquad$ 2×2

rank 1

$r=1$

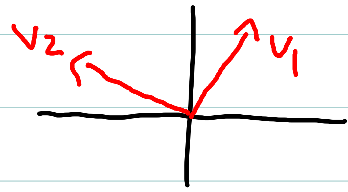
orthonormal basis for col(A)

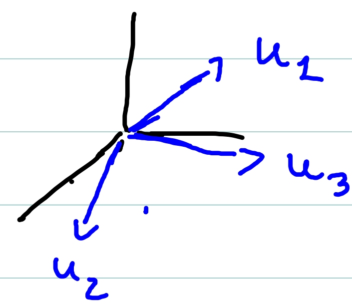
orthonormal basis for Null (A^T)

orthonormal basis for Null(A)

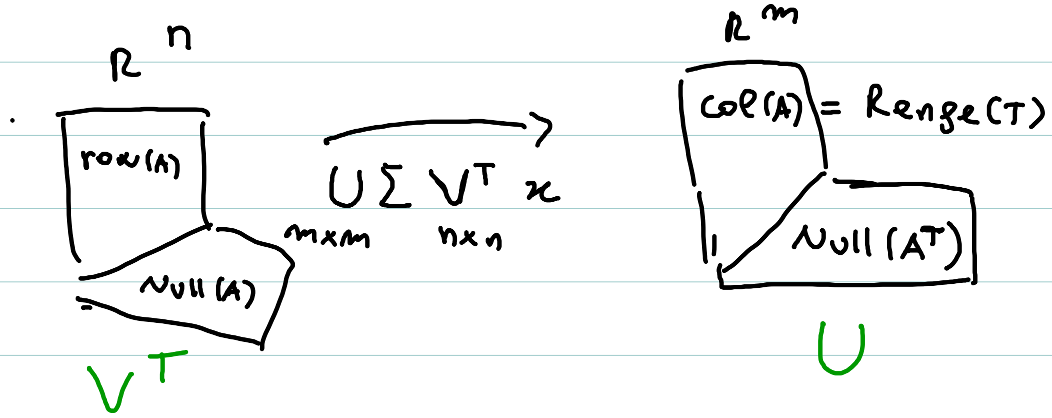
orthonormal basis for row A

$$T: \mathbb{R}^2 \longrightarrow \mathbb{R}^3 \qquad T(v) = Av$$





$A \quad m \times n \qquad T: \mathbb{R}^n \to \mathbb{R}^m \qquad T(v) = Av$



$$\underset{m\times m}{U} = \Big[\underbrace{u_1 \cdots\cdots u_r}_{\text{basis for } Col(A)} \; \overbrace{u_{r+1} \cdots u_m}^{\text{basis for } Null(A^T)}\Big]$$

left singular vectors

$$\underset{n\times n}{V} = \Big[\underbrace{v_1 \cdots\cdots v_r}_{\text{basis for } row(A)} \; \overbrace{v_{r+1} \cdots v_n}^{\text{Basis for } Null(A)}\Big]$$

right singular vectors

Note:

$$\Sigma^T \Sigma = \underset{n\times m}{\begin{bmatrix} \sigma_1 & & & 0 \\ & \ddots & & \\ & & \sigma_r & \\ 0 & & & \ddots & 0 \end{bmatrix}} \underset{m\times n}{\begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ & & \sigma_r \\ 0 & & & \ddots & 0 \end{bmatrix}} = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ 0 & \cdots & \sigma_r^2 & \cdots 0 \\ 0 & \cdots & \cdots & 0 \end{bmatrix}$$

Proof:

Let $\Sigma = (a_{ij})$, $a_{ij} = \begin{cases} \sigma_i & \text{if } i=j\leq r \\ 0 & \text{ow} \end{cases}$

$\Sigma^T = (c_{ij})$ $\quad c_{ij} = a_{ji}$

$\Sigma^T \Sigma = (b_{ij})$

$$b_{ii} = \sum_{k=1}^{n} c_{ik}\, a_{ki} = \sum_{k=1}^{n} a_{ki}\, a_{ki} = a_{ii}^2.$$

$$i\neq j \quad b_{ij} = \sum_{k=1}^{n} c_{ik} a_{kj} = \sum_{k=1}^{n} a_{ki} a_{kj} = 0 \quad \text{because}$$

$a_{ki} = 0$ if $k\neq i$ $\quad a_{kj}=0$ if $k\neq j$

$$\text{so } \Sigma^T\Sigma = \begin{bmatrix} \sigma_1^2 & & & 0 \\ & \ddots & & \\ & & \sigma_r^2 & \\ 0 & & & \ddots & 0 \end{bmatrix}$$

How do we find the σ, U, V ?

Idea:

Suppose $\overset{m\times n}{A} = \overset{m\times n}{U\Sigma V^T}$ then

$$\underset{n\times n}{A^T A} = \underset{n\times n}{V} \underset{n\times m}{\Sigma^T} U^T U \underset{m\times n}{\Sigma} \underset{n\times n}{V^T} = V \underset{n\times n}{\Sigma^T \Sigma} V^T$$

$$\underbrace{= V \begin{bmatrix} \sigma_1^2 & & & \\ & \ddots & & \\ & & \sigma_r^2 & \\ & & & \ddots \ 0 \end{bmatrix} V^T}_{\text{diagonalization of } A^T A} \quad (*)$$

σ_i^2 eigenvalues of $A^T A$

right singular vectors v_i = eigenvectors of $A^T A$

To find left singular values use

$$A V = \begin{bmatrix} u_1 & \cdots & u_m \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix}$$

$\frac{1}{\sigma_i} A v_i = u_i$ for $i = 1 \cdots r$

Complete $u_1 \cdots u_r$ with basis of $\mathrm{Null}(A^T)$

Note that we have assumed here, to start with, what we were trying to prove i.e that A has an SVD, and this is not ok in a proof. The argument in the previous page just gives us an idea of how a careful proof of the existence of SVD could go:

1) Consider $A^T \cdot A$ show it is orthogonally diagonalizable and has non negative eigenvalues $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_r > 0 \ldots 0$. Let $\sigma_i = \sqrt{\lambda_i}$
2) Let $v_1 \ldots v_n$ be the eigenvectors for $\lambda_1 \ldots \lambda_n$ then I can choose them to be orthonormal, $v_1 \ldots v_r$ are a basis for row(A) $v_{r+1} \ldots v_n$ are a basis for Null(A)
3) Let $u_i = \frac{1}{\sigma_i} A v_i$ for $i = 1 \cdots r$ then

   $u_1 \cdots u_r$ are an orthonormal basis for col(A) which can be completed with a basis for Null(A^T) to form an orthonormal basis for R^m
4) with these choices $A = U \Sigma V^T$

Th: If $A \in R^{m \times n}$ has rank r

$\underset{n \times n}{A^T A} \in R^{n \times n}$, is PSD, of rank r

Proof 1) $(A^T A)^T = A^T A$ so $A^T A$ is symmetric

2) it is PSD because is the product of the transpose of a matrix and the matrix

3) rank $(A^T A) = r$:

We will show $\text{Null}(A) = \text{Null}(A^T A)$ $(\subseteq R^n)$

so rank $(A^T A) = n - \text{nullity}(A^T A) = n - \text{nullity}(A)$

$= n - (n - r) = r$ by rank-nullity theorem applied to $A^T A$ and A.

1) suppose $y \in \text{Null}(A)$ then $A^T A y = 0$ so $y \in \text{Null}(A^T A)$

2) suppose $y \in \text{Null}(A^T A)$ then $y^T A^T A y = 0$ so

$\|Ay\|^2 = 0$ so $Ay = 0$ so $y \in \text{Null}(A)$

This tells us that if we look at the eigenvalues of $A^T A$ in descending order they are

$\lambda_1 \dots \lambda_r \; 0 \dots 0$

Given $A \in R^{m \times n}$ of rank r to compute SVD:

1) Diagonalize $A^T A = V \begin{pmatrix} \lambda_1 & & & & \\ & \ddots & & & \\ & & \lambda_r & & \\ & & & 0 & \\ & & & & \ddots \; 0 \end{pmatrix} V^T \qquad \lambda_1 \geq \lambda_2 \cdots \geq \lambda_r$

2) $\sigma_i = \sqrt{\lambda_i}$ are the singular vectors $i = 1 \cdots r$

3) The eigenvectors of $A^T A$ (columns of V) are the right singular vectors

4) $u_i = \frac{1}{\sigma_i} A v_i$ for $i = 1 \cdots r$

$u_{r+1} \cdots u_m$ are an orthonormal basis for $Null(A^T)$

Ex What is the SVD of $\underset{3\times 2}{\begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 1 & 2 \end{bmatrix}}$? $r=1$

1) $A^T A = \begin{bmatrix} 3 & 6 \\ 6 & 12 \end{bmatrix}$ has eigenvalues 15, 0

$E_{15} = \text{span}\left(\underbrace{\begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}}_{= v_1} \right)$ $E_0 = \text{span}\left(\underbrace{\begin{bmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}}_{= v_2} \right)$

$\sigma_1 = \sqrt{15}$

$u_1 = \frac{1}{\sqrt{15}} A v_1 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$

Now we need u_2, u_3 : orthonormal basis for $Null(A^T)$

$A^T = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \end{bmatrix}$ Null (A^T) has dimension 2

$\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \in Null(A^T)$ $\begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix} = u_2$ has length 1

I want to find another vector $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$

s.t $x+y+z=0$ $(\in Null(A^T))$

$\frac{x}{\sqrt{2}} - \frac{z}{\sqrt{2}} = 0$ $\left(\perp \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix} \right)$ so $2x+y=0$

$\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ would work $u_3 = \begin{bmatrix} 1/\sqrt{6} \\ -2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$ has length 1

$$\begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \end{bmatrix} \begin{bmatrix} \sqrt{15} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}$$

$$B = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 2 \\ 0 & 1 & 3 \end{bmatrix}$$ Julia output

svd (B)

```
SVD{Float64,Float64,Array{Float64,2}}
U factor:
33 Array{Float64,2}:
 -0.30519    0.757315  -0.57735
 -0.503259  -0.64296   -0.57735
 -0.808449   0.114355   0.57735
singular values:
3-element Array{Float64,1}:
  3.904484344750072
  1.6598198702273692
 -0.0
Vt factor:
33 Array{Float64,2}:
 0.0507284  -0.285221  -0.957119
 0.84363     0.525159  -0.111784
 0.534522   -0.801784   0.267261
```

$= U$

σ_1, σ_2, σ_3

B has rank 2

$= V^T$

Truncating the above numbers at 5 decimal places, the singular value decompostion of B is

$$\begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 2 \\ 0 & 1 & 3 \end{bmatrix} = \begin{bmatrix} -0.30519 & 0.75731 & -0.57735 \\ -0.50325 & -0.64296 & -0.57735 \\ -0.80844 & 0.11435 & 0.57735 \end{bmatrix} \begin{bmatrix} 3.90448 & 0 & \\ 0 & 1.65981 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0.05072 & -0.28522 & -0.95711 \\ 0.84363 & 0.52515 & -0.11178 \\ 0.53452 & -0.80178 & 0.26726 \end{bmatrix}$$

Notice that B has only two singular values and $\text{rank}(B) = 2$.

Complexity of computing SVD of $A \in \mathbb{R}^{m \times n}$ $\quad \Theta(m \cdot n \cdot \min(m,n))$