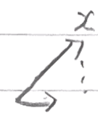
Lec 19



projection matrix

$\pi_b(x) = P \cdot x$

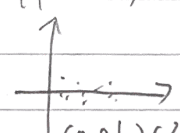
$P^2 = P$

$\pi_b(x) = P\,\pi_b(x)$

$\{b_1, \ldots, b_i, \ldots, b_m\}$ $b_i \in \mathbb{R}^n$
$(m \le n)$

U = span of $\{b_1, \ldots, b_m\}$
if $U = \mathbb{R}^n$, U is basis

Suppose you collect data

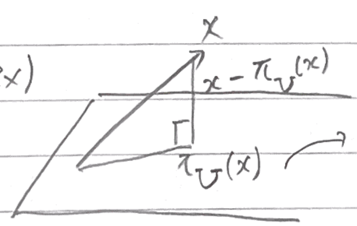


$\begin{pmatrix} 0.01 \\ 0 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \end{pmatrix} \begin{pmatrix} 1.5 \\ 0.002 \end{pmatrix} \begin{pmatrix} 3 \\ -0.001 \end{pmatrix}$

↳ ys are close to 0
↳ can make it one-dimensional when you lose $\pi_b(x)$ for each

SVD → can detect structure
↳ understands subspace that data lies

ex)



→ combination from b_1 to b_m

$= b_1 \cdot \lambda_1 + \ldots + b_m \lambda_m$

$= \begin{bmatrix} | & & | \\ b_1 & \cdots & b_m \\ | & & | \end{bmatrix}^{n \times m} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{bmatrix}^{m \times 1}$

Find $\lambda_1 \ldots \lambda_m$
↳ we express the fact that $x - \pi_U(x) \perp b_i$, $i = 1 \ldots m$

$$\left.\begin{aligned} b_1^T(x - \pi_U(x)) &= 0 \\ &\vdots \\ b_n^T(x - \pi_U(x)) &= 0 \end{aligned}\right\} \Rightarrow \begin{bmatrix} -b_1^T- \\ \vdots \\ -b_m^T- \end{bmatrix} (a - B\lambda) = 0 \text{ or } B^T(x - B\lambda) = 0^{m \times 1}$$

$B^T x = (B^T B)\lambda$

$\lambda = (B^T B)^{-1} B^T x$

$\pi_U(x) = B \cdot \lambda = B(B^T B)^{-1} B^T x$

ex) If $b_1 \dots b_m$ is orthonormal $(B^T B = I)$ then $P = B \cdot B^T$

In this case $b_i^T b_j \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$, this is equivalent to

$$\pi_U(x) = BB^T x = \begin{bmatrix} | & & | \\ b_1 & \dots & b_m \\ | & & | \end{bmatrix} \begin{bmatrix} - b_1^T - \\ \vdots \\ - b_m^T - \end{bmatrix} = \begin{bmatrix} | & & | \\ b_1 & \dots & b_m \\ | & & | \end{bmatrix} \begin{bmatrix} b_1^T x \\ \vdots \\ b_m^T x \end{bmatrix}$$

$$= (b_1^T x) b_1 + \dots + (b_m^T x) b_m$$

Eigenvalue decomposition for symmetric $(A = A^T)$ real matrices

Theorem

Let A be a real symmetric matrix. Then,

1) The eigenvalues $\lambda_1, \dots, \lambda_n$ are real, as are the components of the corresponding eigenvectors $v_1, \dots, v_n$

2) A is orthogonally diagonalizable,

$$A = \sum_{i=1}^{n} \lambda_i v_i v_i^T \quad \left(v_i^T v_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \right)$$

Equivalently, $A = VDV^T$ where $D = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$

and $V = \begin{bmatrix} | & & | \\ v_1 & \dots & v_n \\ | & & | \end{bmatrix}$ or $V^T = V^{-1}$

Theorem:

A real matrix A is orthogonally diagonizable if and only if A is symmetric

Singular Value Decomposition (SVD)

for any matrix $A \in \mathbb{R}^{m \times n}$, there exist orthonormal matrices $v \in \mathbb{R}^{m \times m}$, $V \in \mathbb{R}^{n \times n}$ and a diagonal matrix singular values

with diagonal entries $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r \geq \sigma_{r+1} = \dots \geq \sigma_{\min(m,n)} = 0$

σ_r → rank(A)

such that

$$A = \underbrace{U^{m\times m}}_{\text{orthonormal}} \Sigma^{m\times n} \underbrace{V^T}_{\text{orthonormal}}, \quad V: n\times n$$

① if $m<n$, $\Sigma = \begin{pmatrix} \ddots & 0 \cdots \\ & 0 \cdots 0 \end{pmatrix}$

② if $m=n$, $\Sigma = \begin{pmatrix} \ddots \end{pmatrix}$

Rank(A) $\leq$ min(m,n)

full rank rank(a) = min(m,n)

③ if $m>n$ $\begin{pmatrix} \ddots \\ 0 \cdots 0 \end{pmatrix}$

SVD

→ eigenvalues are always greater than 0

$$A = U \Sigma V^T$$
$$= \sum_{i=1}^{n} \sigma_i (v_i v_i^T) \leq \min(m,n)$$

$A^{n\times n} = \sum \lambda_i u_i u_i^T \rightarrow y^{m\times 1} = A^{m\times n} \cdot x^{n\times 1}$

$y^{n\times 1} = A^{T\, n\times m} \cdot x^{m\times 1}$

left singular vectors

$$A = \underbrace{U}_{\text{left singular vectors}} \underbrace{\Sigma}_{\text{singular values}} \underbrace{V^T}_{\text{right singular vectors}} = \sum_{i=1}^{r} \sigma_i u_i v_i^T$$

Lemma

$Av_i = \sigma_i u_i$ ① and $A^T u_i = \sigma_i v_i$ ②

Proof

Since v is orthogonal, $A \cdot V = U \Sigma V^T V = U\Sigma$ which is equation ①

$$A^T u_i = \left(\sum_{k=1}^{r} \sigma_k u_k \cdot v_k^T\right) u_i = \sigma_i v_i \underbrace{(u_i \cdot u_i^T)}_{=1} = \sigma_i \cdot v_i$$

Best k-rank Approximation

Define $A_k := \sum_{i=1}^{k} \sigma_i u_i v_i^T \quad (k = 1, 2, \ldots, r)$

Lemma: The rows of A_k are the projections of the rows of A onto the subspace

$$A^{m \times n} = \left[\begin{array}{ccc|ccc} | & & | & | & & | \\ u_1 & \cdots & u_r & u_{r+1} & \cdots & u_m \\ | & & | & | & & | \end{array}\right] \cdot \Sigma \cdot \left[\begin{array}{c} - v_1^T - \\ \vdots \\ - v_r^T - \\ \hline - v_{r+1}^T - \\ \\ - v_m^T - \end{array}\right]$$

$\text{rank}(A) = r \leq \min(m \cdot n)$

$\hookrightarrow$ when equality holds, A is full rank

$R(A) = R([u_1, \ldots, u_r]) \quad N(A^T) = R([u_{r+1}, \ldots u_m])$

$R(A^T) = R([v_1, \ldots, v_r]) \quad N(A) = R([v_{r+1}, \ldots, v_m])$