

Singular Value Decomposition

Start w/ any matrix A say $m \times n$ represents a linear trans.

$$T: \mathbb{R}^n \rightarrow \mathbb{R}^m$$

Consider the matrix $A^T A$ ($n \times n$ matrix)

this is symmetric: $(A^T A)^T = A^T (A^T)^T = A^T A$

So we can find (by the principal axes theorem)
an orthonormal basis $v_1, \dots, v_n$ of eigenvectors.

Observe: if v_i has an eigenvalue of λ_i then $\lambda_i \geq 0$

Why? $A^T A v_i = \lambda_i v_i$

$$\begin{aligned} v_i \cdot (A^T A v_i) &= v_i \cdot (\lambda_i v_i) = \lambda_i v_i \cdot v_i = \lambda_i \|v_i\|^2 \\ &= (v_i^T A^T A) v_i \end{aligned}$$

$$= (Av_i)^T (Av_i) = Av_i \cdot Av_i = \|Av_i\|^2$$

$$\Rightarrow \lambda_i \underbrace{\|v_i\|^2}_1 = \|Av_i\|^2 \quad \lambda_i = \|Av_i\|^2 \geq 0$$

Same logic actually shows: the vectors Av_i are mutually orthogonal & $\|Av_i\| = \sqrt{\lambda_i}$

$$\begin{aligned} Av_i \cdot Av_j &= (Av_i)^T Av_j = v_i^T A^T Av_j \\ &= v_i^T (A^T Av_j) \\ &= v_i^T \lambda_j v_j = \lambda_j v_i^T v_j \\ &= \lambda_j v_i \cdot v_j \end{aligned}$$

So if $i \neq j$ then $v_i \cdot v_j = 0$ tells us
 $Av_i \cdot Av_j = 0$

Punchline: Given $m \times n$ matrix A , can find an orthonormal basis $v_1, \dots, v_n$ for $\mathbb{R}^n$ s.t. $Av_1, Av_2, \dots, Av_n$ are orthogonal

let $\sigma_i = \|Av_i\|$

Suppose we have (after reordering them)

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r \geq \underbrace{\sigma_{r+1}}_0 = \sigma_{r+2} = \dots = \sigma_n$$

Side observation: $Av_1, \dots, Av_r$ are nonzero, independent

and span the image of A . (are a basis for the image)
 $\Rightarrow \text{rank } A = \dim \text{image} = r.$

$$\text{let } w_1 = \frac{Av_1}{\|Av_1\|} = \frac{Av_1}{\sigma_1}, w_2 = \frac{Av_2}{\|Av_2\|} = \frac{Av_2}{\sigma_2}, \dots$$

then $w_1, \dots, w_r$ are orthonormal
 $\therefore$ we can extend them to an orthonormal basis of $\mathbb{R}^m$
 (Gram-Schmidt)

$$w_1, \dots, w_r, w_{r+1}, \dots, w_m$$

$$\mathbb{R}^n \xrightarrow{\text{given by } A} \mathbb{R}^m$$

$$v_1 \mapsto Av_1 = \sigma_1 w_1$$

$$v_2 \mapsto Av_2 = \sigma_2 w_2$$

$\vdots$

$$v_r \mapsto Av_r = \sigma_r w_r$$

$$v_{r+1} \mapsto Av_{r+1} = \underbrace{\sigma_{r+1}}_0 w_{r+1} = 0$$

$\vdots$

$$v_m \mapsto 0$$

$\vdots$

So: if $\mathcal{B} = \{v_i\}$ $\mathcal{C} = \{w_j\}$ then

$$e^T_B = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_r \\ 0 & 0 & \dots & 0 \end{bmatrix} = \Sigma_A \quad \text{"the singular matrix for A"}$$

Let Q be the matrix $Q = [w_1 | \dots | w_m]$ $m \times m$ matrix

$P = [v_1 | \dots | v_n]$ $n \times n$ matrix

note: $\{w_i\}$ & $\{v_j\}$ orthogonal bases $\Rightarrow P, Q$ are orthogonal matrices

$$\text{i.e. } P^T = P^{-1} \quad Q^T = Q^{-1}$$

$$\Sigma_A = e^T_B = Q^{-1} A P \quad \text{"change of basis"}$$

$$\text{or: } Q \Sigma_A P^{-1} = A$$

$$Q \Sigma_A P^T$$

Theorem: "Singular value decomposition"

Given an $m \times n$ matrix A can find $n \times n$ & $m \times m$ orthogonal matrices P & Q so that

$$A = Q \Sigma_A P^T \quad \text{where } \Sigma_A \text{ has form } \begin{bmatrix} \sigma_1 & \dots & 0 \\ & \ddots & \\ 0 & \sigma_r & \dots & 0 \end{bmatrix}$$

w/ $\sigma_1, \sigma_2, \dots, \sigma_r$ are the "singular values" and $\text{rk} A = r$.

$$A = \begin{bmatrix} \text{CRAZY} \end{bmatrix} \quad \text{and } \sigma_1 = 35$$

6×10

$$\begin{array}{ccc} \mathbb{R}^{10} & \xrightarrow{\quad} & \mathbb{R}^6 \\ \|v\|=1 & v_1 \xrightarrow{\quad} & 35w_1 \quad \|w_1\|=1 \end{array}$$

$$v = \begin{bmatrix} \vdots \\ \vdots \\ \vdots \end{bmatrix}$$

$$Av = ?$$



$v \mapsto$ project into v_1 , do A

$$\frac{v \cdot v_1}{\|v\|^2} \cdot v_1 = (v \cdot v_1) v_1$$

$$\begin{aligned} A(v \cdot v_1) v_1 \\ = (v \cdot v_1) \cdot Av_1 \end{aligned}$$

$$Av \approx \boxed{(v \cdot v_1) 35w_1}$$

guess other $\sigma_2, \dots, \sigma_r$ are $\ll 35$

approx

$$\begin{array}{ccc} v_1 & \xrightarrow{\quad} & 35w_1 \\ v_2 & \xrightarrow{\quad} & 0 \\ \vdots & & \\ v_n & \xrightarrow{\quad} & 0 \end{array}$$

what if we repeat & get

$$\begin{array}{ccc} v_1 & \xrightarrow{\quad} & 35w_1 \\ v_2 & \xrightarrow{\quad} & 2w_2 \end{array}$$

$$\sigma_1 = 35$$

$$\sigma_2 = 2$$

$$\sigma_3 = 1.9$$

$$\sigma_4 = 1.7$$

$$\sigma_5 = .003$$