SVD:

②

$$A = U D V^T$$

D → valores singulares

U → vectores singulares izquierdo

V^T → Vectores singulares derechos

$$A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$$

Calculamos los valores singulares σ_3 mediante los autovalores de AA^T

$$AA^T = \begin{pmatrix} 17 & 8 \\ 8 & 17 \end{pmatrix}$$

Cuyo polinomio característico es $|AA^T - \lambda I| = \lambda^2 - 34\lambda + 255$
$= (\lambda - 25)(\lambda - 9)$

$\Rightarrow$ los valores singulares son $\sigma_1 = \sqrt{25} = 5$
$\sigma_2 = \sqrt{9} = 3$

los vectores singulares derechos se encuentran mediante el conjunto ortonormal de autovalores de A^TA. Del mismo modo para los vectores singulares izquierdos. Los autovalores de A^TA son 25, 9 y 0, además A^TA es simétrica, entonces los autovectores serán ortogonales.

Para $\lambda = 25$

$$A^TA - 25I = \begin{pmatrix} -12 & 12 & 2 \\ 12 & -12 & -2 \\ 2 & -2 & -17 \end{pmatrix} \xrightarrow[\text{fila}]{\text{operaciones}} \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Un vector unitario en el núcleo de la matriz anterior es $v_1 = \begin{pmatrix} 1\sqrt{2} \\ 1\sqrt{2} \\ 0 \end{pmatrix}$

De mismo modo, para $\lambda = 9$:

$$A^TA - 9I = \begin{pmatrix} 4 & 12 & 2 \\ 12 & 4 & -2 \\ 2 & -2 & -1 \end{pmatrix} \xrightarrow[\text{filas}]{\text{operaciones}} \begin{pmatrix} 1 & 0 & -1/4 \\ 0 & 1 & 1/4 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{con } v_2 = \begin{pmatrix} 1/\sqrt{18} \\ -1/\sqrt{18} \\ 4/\sqrt{18} \end{pmatrix}$$

Para el último autovector, nosotros computamos el núcleo de $A^T A$ o encontramos un vector unitario perpendicular a v_1 y v_2. Para ser perpendicular a $v_1 = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$, necesitamos $-a = b$. Entonces la condición que $v_2^T v_3 = 0$ se convierte en $2a/\sqrt{18} + 4c/\sqrt{18} = 0$ o $-a = 2c$. Entonces $v_3 = \begin{pmatrix} a \\ -a \\ -a/2 \end{pmatrix}$ y para que v_3 sea un vector unitario, necesitamos $a = \frac{2}{3}$

$$\Rightarrow \quad v_3 = \begin{pmatrix} 2/3 \\ -2/3 \\ -1/3 \end{pmatrix}$$

$$A = U \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{18} & -1/\sqrt{18} & 4/\sqrt{18} \\ 2/3 & -2/3 & -1/3 \end{pmatrix}$$

Calculamos U con $\sigma u_i = A v_i$ o $u_i = \frac{1}{\sigma} A v_i$

$$\Rightarrow \quad A = U D V^T = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{18} & -1/\sqrt{18} & 4/\sqrt{18} \\ 2/3 & -2/3 & -1/3 \end{pmatrix}$$

Del mismo modo para $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1.4 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0.7 & 0.7 & 0 \\ 0 & 0 & 1 \\ 0.7 & -0.7 & 0 \end{pmatrix}$

(4)

$$f(\underbrace{x_1 + x_2}_{Y} \mid \underbrace{x_1 - x_2}_{Z} = 0) = \frac{f_{Y,Z}(y,z)}{f_Z(z)}$$

$$\text{piden} \quad F_{Y,Z}(y \mid z) = \int_{-\infty}^{+\infty} f(y \mid z)\, du$$

$$f_Z(z=0) = \frac{1}{\sqrt{(2\pi)^2 \det(\Sigma)}} \exp\left(-\frac{1}{2}(\bar{0} - (1,1))^T \Sigma^{-1} (\bar{0} - (1,1))\right)$$

$$f_Z(z=0) = \frac{1}{\sqrt{(2\pi)^2\, 5}} \exp\left(-\frac{1}{2}(-1,-1)^T \begin{pmatrix} 2/5 & -1/5 \\ -1/5 & 3/5 \end{pmatrix} (-1,2)\right) = 0.0527$$

$\underset{y,z}{f}(y,z)$ con $y = x_1 + x_2$

$z = x_1 - x_2 = 0 \Rightarrow x_1 = x_2$

$\Rightarrow y = 2x_1$

f

5:

$$f(\theta) = \frac{1}{2} \sum_{i}^{n} w_i (\theta - x_i)^2$$

$$\frac{\partial f(\theta)}{\partial \theta} = \sum_{i}^{n} w_i(\theta - x_i) \cdot 1 = \sum_{i}^{n} w_i (\theta - x_i)$$ Aplicando regla de la cadena

con $w_i > 0 \quad \forall i \in \mathbb{N}$

$$\frac{\partial f(\theta)}{\partial \theta} = \sum_{i}^{n} w_i (\theta - x_i) = 0$$

$\Rightarrow \quad \theta = x_i$ minimiza $f(\theta) \quad \forall w_i > 0$

Si $w_i < 0$

6.

$$f(w) = \sum_{i=1}^{n} \sum_{j=1}^{n} (a_i^T w - b_j^T w)^2 + \lambda \|w\|_2^2$$

$\|W\|_2^2$: Norma de Frobenius.

$$\nabla f(w) = \sum_{i=1}^{n} \sum_{j=1}^{n} 2 (a_i^T w - b_j^T w) \cdot (a_i^T - b_j^T) + 2\lambda \sum_{k=1}^{d} w_k$$

$$\nabla f(w) = 2 \left(\sum_{i=1}^{n} \sum_{j=1}^{n} (a_i^T w - b_j^T w) \cdot (a_i^T - b_j^T) + \lambda \sum_{k=1}^{d} w_k \right)$$

Podemos usar esta gradiente para algún algoritmo de optimización.

10-

$$Loss(x,y,w) = \begin{cases} \frac{1}{2}(w\cdot\phi(x)-y)^2 & para\ |w\cdot\phi(x)-y| \leq 1 \\ |w\cdot\phi(x)-y|-\frac{1}{2} & para\ |w\cdot\phi(x)-y| > 1 \end{cases}$$

$$\nabla loss = \frac{\partial Loss}{\partial w}$$

1era caso

$$loss = \frac{1}{2}(w\cdot\phi(x)-y)^2 \quad para \quad |w\cdot\phi(x)-y| \leq 1$$

$$\Rightarrow (|w\cdot\phi(x)-y|)^2 \leq 1$$

$$(w\cdot\phi(x)-y)^2 \leq 1$$

$$\Rightarrow \frac{\partial Loss}{\partial w} = (w\cdot\phi(x)-y)\cdot\phi(x)$$

$$\frac{\partial Loss}{\partial w} = ((1,2)\cdot(4,10)-16)\cdot(4,10) = (-8,-20)$$

2do caso

$$loss = |w\cdot\phi(x)-y|-\frac{1}{2} \quad para \quad |w\cdot\phi(x)-y|>1 : \quad w\phi(x)-y>1 \ \cup \ w\cdot\phi(x)-y<-1$$

$$loss = -w\cdot\phi(x)+y-\frac{1}{2}$$

$$\frac{\partial Loss}{\partial w} = -\phi(x) = (-4,-10)$$