# SVD – leftovers.

$$A = \underbrace{\boxed{\phantom{AAAA}}}_{m\times n} = \underbrace{\boxed{U}}_{m\times m}\ \underbrace{\overset{\Sigma}{\boxed{\begin{matrix}\sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \sigma_r & & \\ & & & 0 & \\ & & & & 0\end{matrix}}}}_{m\times n}\ \underbrace{\boxed{V^T}}_{n\times n}$$

$k$ $k$

$$A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \dots + \sigma_r u_r v_r^T$$

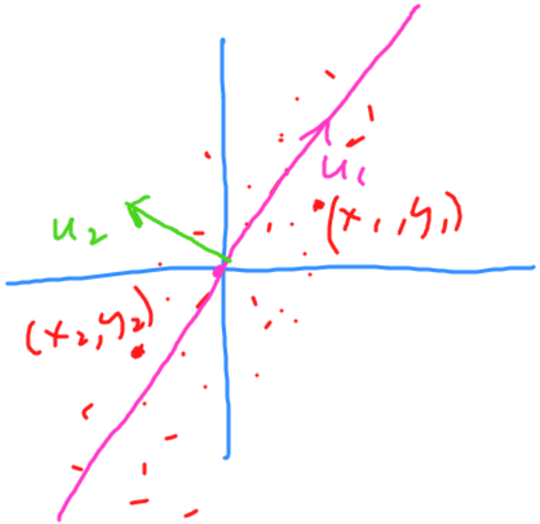
$$U = \begin{bmatrix} | & | & \\ u_1 & u_2 & \cdots \\ | & | & \end{bmatrix}$$

$$A_k = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \dots + \sigma_k u_k v_k^T$$

$$V = \begin{bmatrix} | & | & \\ v_1 & v_2 & \cdots \\ | & | & \end{bmatrix}$$

$\approx$ approximation of $A$ by a matrix of rank $k$

# PCA (principal component analysis)




$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}, \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}, \begin{bmatrix} x_3 \\ y_3 \end{bmatrix}, \ldots$$

$$A = \underbrace{\begin{bmatrix} x_1 & x_2 & \cdots \\ y_1 & y_2 & \cdots \end{bmatrix}}_{2 \times n} = \underbrace{\begin{array}{cc} u_1 & u_2 \\ \multicolumn{2}{c}{\boxed{2 \times 2}} \end{array}}_{U} \; \underbrace{\begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \end{bmatrix}}_{\Sigma,\ 2 \times n} \; \underbrace{\begin{bmatrix} n \times n \end{bmatrix}}_{V^T}$$

How good the approximation of $A$ by $A_k$ is?

Norm of a matrix $A$: $\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$

$A = U \Sigma V^T$

orthogonal

don't change length

$\|\overbrace{V^T x}^{y}\| = \|x\|$ $\quad \|y\| = \|x\|$

$\|U\Sigma(V^T x)\| = \|\Sigma(V^T x)\| = \|\Sigma y\|$

so $\|A\| = \|\Sigma\| = \left\| \begin{bmatrix} \sigma_1 & & & & & \\ & \sigma_2 & & & & \\ & & \ddots & & & \\ & & & \sigma_r & & \\ & & & & 0 & \\ & & & & & \ddots \end{bmatrix} \right\| = \sigma_1$

$\|A - \underbrace{A_k}_{\| \atop U\begin{bmatrix} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_k & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix} V^T}\| = \left\| U \begin{bmatrix} 0 & & & & & \\ & \ddots & & & & \\ & & 0 & & & \\ & & & \sigma_{k+1} & & \\ & & & & \ddots & \\ & & & & & \sigma_r \\ & & & & & & 0 \\ & & & & & & & \ddots \end{bmatrix} V^T \right\| = \sigma_{k+1}$

$A_k$

<u>Thm</u> If $A, B$ are matrices of the same size and rank $B$ is at most $k$ then

$\|A - B\| \geq \|A - A_k\| = \sigma_{k+1}$.

Pseudoinverse.

← as close as you can get to "$A^{-1}$" if $A$ is not invertible.

$$\underset{m\times n}{A} = \underset{m\times m}{U} \underset{m\times n}{\begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \ddots & & \\ & & & \sigma_r & \\ & & & & 0 \\ & & & & & \ddots \end{bmatrix}} \underset{n\times n}{V^T}$$

Pseudoinverse $A^+$ is:

$$A^+ = \underset{n\times n}{V} \underset{n\times m}{\begin{bmatrix} \sigma_1^{-1} & & & & \\ & \sigma_2^{-1} & & & \\ & & \ddots & & \\ & & & \sigma_r^{-1} & \\ & & & & 0 \\ & & & & & \ddots \\ & & & & & & 0 \end{bmatrix}} \underset{m\times m}{U^T}$$

← $\Sigma^+$

$$AA^+ = (U\Sigma V^T)(V\Sigma^+U^T) =$$
$$= U\Sigma(V^TV)\Sigma^+U^T =$$
$$= U\underbrace{\Sigma\Sigma^+}U^T = U\begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix}\underbrace{U^T_r}$$

proj. onto span $A$ $u_1, \dots, u_r$.

$$\Sigma\Sigma^+ = \underset{m\times n}{\begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & 0 \\ & & & & \ddots \end{bmatrix}} \underset{n\times m}{\begin{bmatrix} \sigma_1^{-1} & & & \\ & \sigma_2^{-1} & & \\ & & \ddots & \\ & & & \sigma_r^{-1} \\ & & & & 0 \\ & & & & & \ddots \end{bmatrix}} = \underset{m\times m}{\begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & 0 \\ & & & & & \ddots \\ & & & & & & 0 \end{bmatrix}}$$

$$A^+A = V\begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix}V^T$$
$$\underbrace{\qquad}_{r}$$

← projection matrix onto span $A$ $v_1, \dots, v_r$

# Linear Transformations.

Tranformation $T$ takes input vector $v$ in vector space $V$ and output a vector $w$ in vector space $W$.

Notation: $T(v) = w$, $Tv = w$

$$T(0) = T(2 \cdot 0)$$
$$= 2T(0)$$
so $T(0) = 0$

A transformation $T$ is <u>linear</u> if

$$T(c_1 v_1 + c_2 v_2) = c_1 T(v_1) + c_2 T(v_2).$$

<u>Ex.</u> 1 $V = \mathbb{R}$, $W = \mathbb{R}$

$T(x) = ax$

$$T(c_1 x_1 + c_2 x_2) = a c_1 x_1 + a c_2 x_2$$

$$c_1 T(v_1) + c_2 T(v_2) = c_1 a x_1 + c_2 a x_2$$

so $T(x) = ax$ is linear.

**Ex 2**  $V = \mathbb{R}$, $W = \mathbb{R}$

$$T(x) = x + b$$

$$T(c_1x_1 + c_2x_2) = c_1x_1 + c_2x_2 + b$$

$\neq 0$ unless $b = 0$

$$c_1T(x_1) + c_2T(x_2) = c_1(x_1 + b) + c_2(x_2 + b)$$

$$= c_1x_1 + c_2x_2 + c_1b + c_2b$$

$T(x) = x + b$ is not a linear transformation

Terminology:

linear + shift (translation) = affine transformation

**Ex3** $V = \mathbb{R}^n$ $W = \mathbb{R}^m$

$A$ an $m \times n$ matrix

$T(v) = Av$ is linear:

$$A(c_1 v_1 + c_2 v_2) = Ac_1 v_1 + Ac_2 v_2 =$$
$$= c_1 A v_1 + c_2 A v_2 = c_1 T(v_1) + c_2 T(v_2)$$

$\|$

$T(c_1 v_1 + c_2 v_2)$

**Ex4** If $V, W$ have finite dimension then every linear transformation from $V$ to $W$ can be realized as multiplication by some matrix $A$!

Ex 4') $\frac{d}{dx}(c_1 f + c_2 g) = c_1 \frac{d}{dx} f + c_2 \frac{d}{dx} g$

$T = \frac{d}{dx}$ is a linear transformation.

$V = \{\text{polynomials } ax^2+bx+c\}$ $\qquad ax^2+bx+c \longleftrightarrow \begin{bmatrix} a \\ b \\ c \end{bmatrix}$

$W = \{\text{polynomials } bx+c\}$ $\qquad bx+c \longleftrightarrow \begin{bmatrix} b \\ c \end{bmatrix}$

$$\frac{d}{dx} x^2 = 2x \longleftrightarrow T\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \underline{\begin{bmatrix} 2 \\ 0 \end{bmatrix}}$$

$$\frac{d}{dx} x = 1 \longleftrightarrow T\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \underline{\begin{bmatrix} 0 \\ 1 \end{bmatrix}}$$

$$\frac{d}{dx} 1 = 0 \longleftrightarrow T\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \underline{\begin{bmatrix} 0 \\ 0 \end{bmatrix}}$$

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 2a \\ b \end{bmatrix} \longleftrightarrow \frac{d}{dx}(ax^2+bx+c) = 2ax+b$$

?