EECS 16B ## 4.2 LECTURE 20

STRUCTURAL PROPERTIES OF A

$A\vec{x} = U_1 \underbrace{(SV_1^T\vec{x})}_{\vec{y}} = [\vec{u}_1 \dots \vec{u}_r]\begin{bmatrix} y_1 \\ \vdots \\ y_r \end{bmatrix}$ ← columns of U_1 form an orthonormal basis for the column space of A.

Also, the columns of V_2 form an orthonormal basis for the null space of A.
↳ If $\vec{x}$ is a linear combination of cols of V_2, $V_1^T\vec{x} = 0$, so $A\vec{x} = 0$.

SVD: LEAST SQUARES

$\vec{y} = A\vec{x}$ → unknowns $\vec{y} \in \mathbb{R}^m$ $A \in \mathbb{R}^{m\times n}$ overdetermined system
↳ measurements $\vec{x} \in \mathbb{R}^n$ $m > n = r$

$\boxed{A} = \boxed{U}\begin{bmatrix} S \\ 0 \end{bmatrix}_{\Sigma}\boxed{V^T}$ → $\|\vec{e}\| = \|\vec{y} - A\vec{x}\| = \overbrace{\|U(U^T\vec{y} - \begin{bmatrix} S \\ 0 \end{bmatrix}V^T\vec{x})\|}^{*} = \|U^T\vec{y} - \begin{bmatrix} S \\ 0 \end{bmatrix}V^T\vec{x}\|$ → $\left\|\begin{bmatrix} U_1^T \\ U_2^T \end{bmatrix}\vec{y} - \begin{bmatrix} SV^T \\ 0 \end{bmatrix}\vec{x}\right\|$ → $\left\|\begin{bmatrix} U_1^T\vec{y} - SV^T\vec{x} \\ U_2^T\vec{y} \end{bmatrix}\right\|$ →

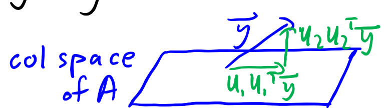
$\|\vec{e}\|^2 = \underbrace{\|U_1^T\vec{y} - SV^T\vec{x}\|^2}_{\text{we can 0 this}} + \underbrace{\|U_2^T\vec{y}\|^2}_{\text{constant}}$

we know $S^{-1} = \begin{bmatrix} 1/\sigma_1 & & \\ & \ddots & \\ & & 1/\sigma_r \end{bmatrix}$ & $V^{T^{-1}} = V$

$\boxed{\vec{x} = VS^{-1}U_1^T\vec{y}}$ → $SV^T\vec{x} = SV^TVS^{-1}U_1^T\vec{y} = U_1^T\vec{y}$

this is another least squares formula, equivalent to $\vec{x} = (A^TA)^{-1}A^T\vec{y}$!

* we multiplied $\vec{y}$ by $UU^T = I$ → $[U_1\ U_2]\begin{bmatrix} U_1^T \\ U_2^T \end{bmatrix}\vec{y} = \underbrace{U_1U_1^T\vec{y}}_{\text{projection } \vec{y} \text{ onto col space of A}} + \underbrace{U_2U_2^T\vec{y}}_{\text{remainder orthogonal to col space}}$



MINIMUM NORM SOLUTION

$\vec{y} = A\vec{x}$ $m < n$ assume that A has independent rows: $r = m$
there are many solutions $\vec{x}$ that satisfy this.

One approach is to find the smallest length $\vec{x}$ (ex. smallest inputs)

$A = \boxed{U}\underbrace{\boxed{S \mid 0}}_{\Sigma}\boxed{V^T}$ → $\vec{y} = U[S\,|\,0]V^T\vec{x} = USV_1^T\vec{x}$ → $U^T\vec{y} = SV_1^T\vec{x}$ → $\boxed{S^{-1}U^T\vec{y} = V_1^T\vec{x}}$

$U^{-1} = U^T$ ↳ $A\vec{x} = \vec{y}$ $\boxed{V_1^T}\boxed{\vec{x}}$

There are $n-r$ degrees of freedom. How do find the smallest-length $\vec{x}$?

$\|\vec{x}\| = \|V^T\vec{x}\| = \left\|\begin{bmatrix} V_1^T\vec{x} \\ V_2^T\vec{x} \end{bmatrix}\right\|$ → $\|\vec{x}\|^2 = \underbrace{\|V_1^T\vec{x}\|^2}_{\|S^{-1}U^T\vec{y}\|^2} + \underbrace{\|V_2^T\vec{x}\|^2}_{0}$

$V^T\vec{x} = \begin{bmatrix} S^{-1}U^T\vec{y} \\ 0 \end{bmatrix}$ → $VV^T\vec{x} = V_1S^{-1}U^T\vec{y}$ → $\boxed{\vec{x} = V_1S^{-1}U^T\vec{y}}$

This is equivalent to $\vec{x} = A^T(AA^T)^{-1}\vec{y}$.

Any $\vec{x} = [V_1\ V_2]\begin{bmatrix} S^{-1}U^T\vec{y} \\ * \end{bmatrix}$ will solve $A\vec{x} = \vec{y}$, though the minimum norm solution occurs for $* = 0$.

Principal Component Analysis

Suppose $A \in \mathbb{R}^{m \times n}$ contains n measurements from m samples, but the measurements are of different characteristics.

EX

Child	height	weight
1		
⋮		
m		

how do we find the correlation?

we start by finding avg. height & weight, and form a new A with averages subtracted (center data around 0)

$A^TA \in \mathbb{R}^{2 \times 2}$ $\quad AA^T \in \mathbb{R}^{m \times m}$

↳ $A^TA\vec{v}_1 = \lambda_1 \vec{v}_1$ $\quad \sigma_1 > \sigma_2$, so most info is in $\vec{v}_1$

$\frac{1}{m-1} A^TA$ is called the "covariance theorem." $\frac{1}{m-1} A^TA = \frac{1}{m-1}\begin{bmatrix} \vec{h}^T\vec{h} & \vec{h}^T\vec{w} \\ \vec{w}^T\vec{h} & \vec{w}^T\vec{w} \end{bmatrix}$

The diagonal here contains the variances of height & weight.