The Amazing

$\boxed{\text{Singular Value Decomposition}}$

aka SVD

<u>Setup</u>

$$\underset{n \times p}{X}$$

- $n > p$ (for now)
- $\text{colMeans}(X) = 0$
- $X = [x_1, \dots, x_p]$

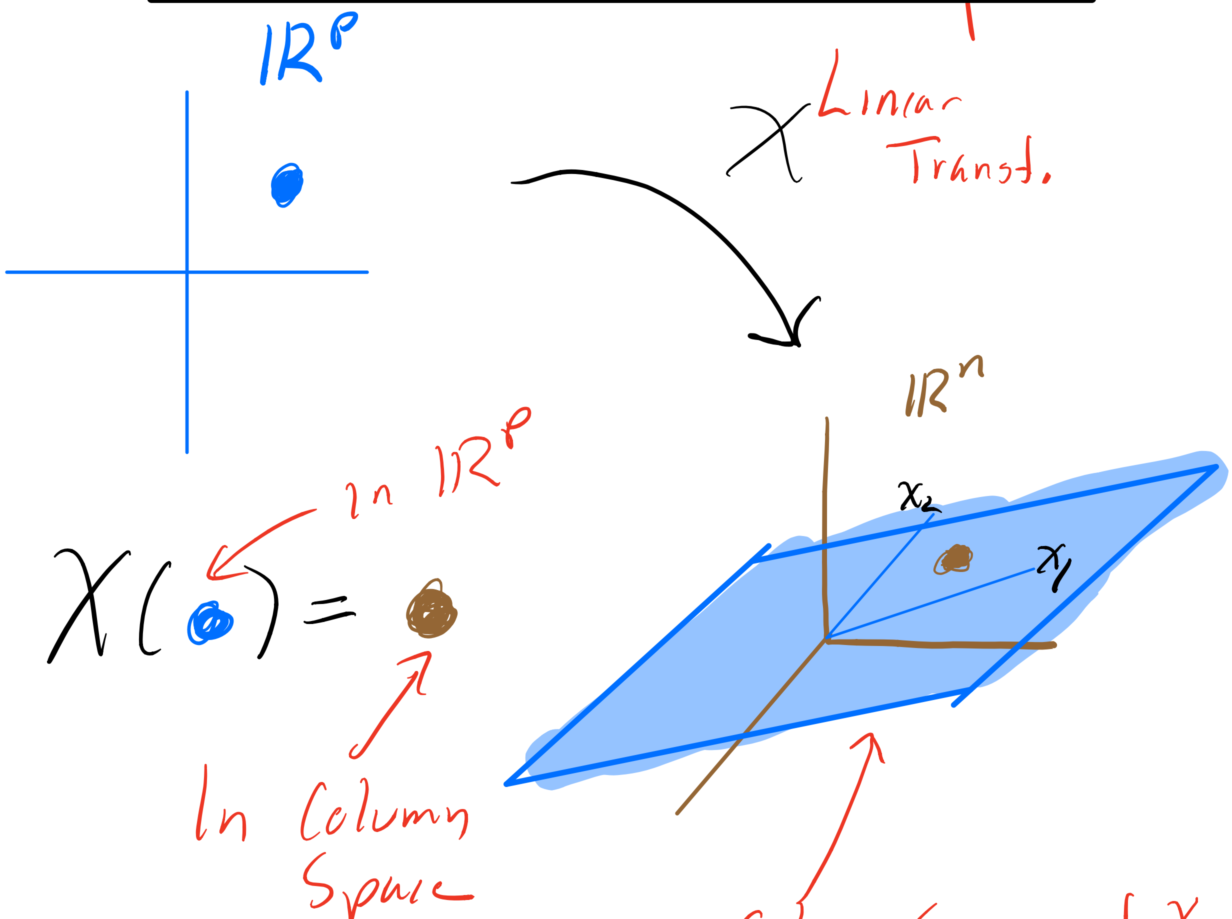
x_1: $n \times 1$ column vector

Think of X as linear transformation

$$X : \mathbb{R}^p \longrightarrow \mathbb{R}^n$$

In the standard basis $\mathcal{E}$

Geometry of X



Column Space of X Spanned by $x_1, \ldots x_p$

Singular Value Decomposition of $\underset{n \times p}{X}$

$$\boxed{\underset{n\times p}{X} = \underset{n\times p}{U}\ \underset{p\times p}{\Sigma}\ \underset{p\times p}{V^T}}$$

$U^{-1} = U^T$!

$U = [u_1 \dots u_p]$ $\quad UU^T =$ Identity

$V = [v_1, \dots, v_p]$ $\quad VV^T =$ Identity

$V^{-1} = V^T$!

i.e. U, V Orthonormal

$$\Sigma = \begin{bmatrix} \sigma_1 & & & 0 \\ & \sigma_2 & & \\ & & \ddots & \\ 0 & & & \sigma_p \end{bmatrix}$$

Diagonal

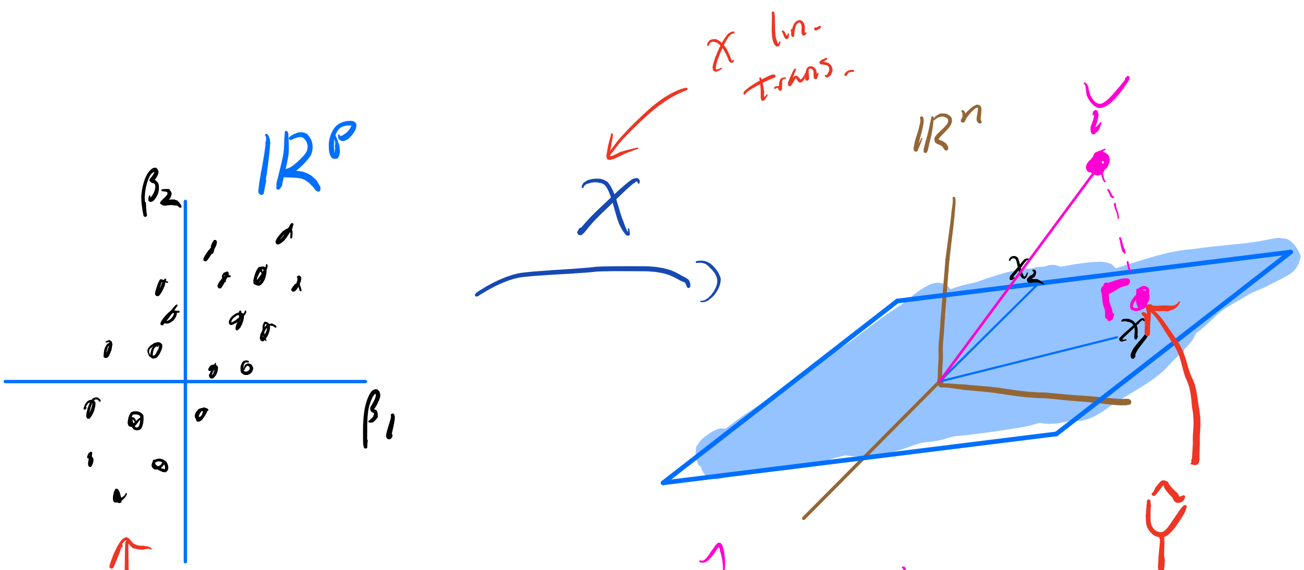
$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p \geq 0$$

Real

SVD & Linear Regression

$\underset{n \times p}{X}$ $\qquad$ $\underset{n \times 1}{Y}$ $\qquad$ $Y = X\beta + \varepsilon$

$$\beta = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} \qquad \beta_0 = 0 \text{ since } colMeans(X) = 0$$



X (data)
n pts in $\mathbb{R}^p$

$\hat{Y}$ = Orthogonal Projection of Y onto Column Space

$\hat{\beta}$ solves $\hat{Y} = X\hat{\beta}$

Minimize $(Y - X\hat{\beta})^T(Y - X\hat{\beta})$

$$\Rightarrow \quad X^T(Y - X\hat{\beta}) = 0$$

$$\Rightarrow \quad X^T X \hat{\beta} = X^T Y$$

SVD of $X = U \Sigma V^T$

$$\Rightarrow \quad X^T X = (U\Sigma V^T)^T (U \Sigma V^T)$$
$$= V \Sigma U^T U \Sigma V^T$$
$$= V \Sigma^2 V^T$$

Invertible if $\sigma_p > 0$ "Full Rank"

$$(X^T X)^{-1} = (V \Sigma^2 V^T)^{-1}$$
$$= V \Sigma^{-2} V^T \quad \text{(nice!)}$$

$V \Sigma U^T$

So, $\hat{\beta} = (X^T X)^{-1} X^T Y$

$$= (V \Sigma^{-2} V^T)(V \Sigma U^T) Y$$
$$= (\underbrace{V \Sigma^{-1} U^T}_{\text{Sorta looks like inverse "Generalized Inverse"}}) Y$$

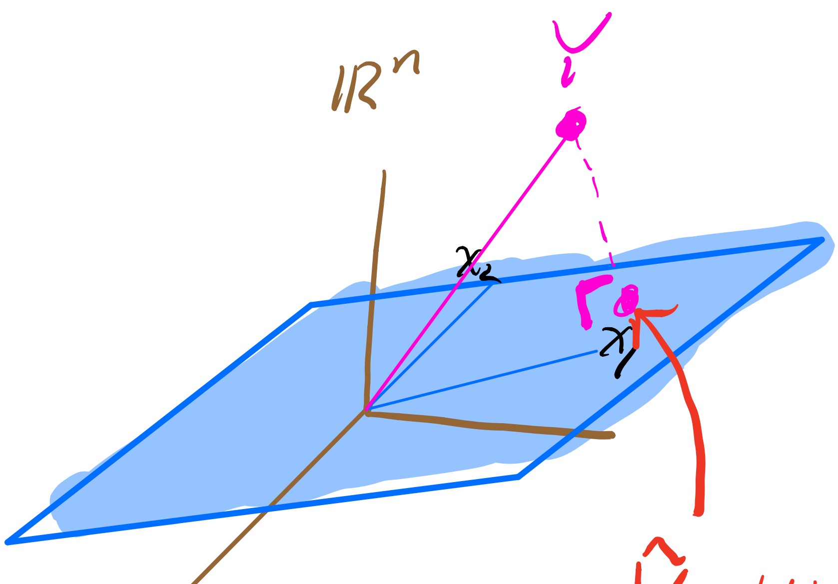
Note: $X = U \Sigma V^T$ not invertible

If $\hat{\beta} = V\Sigma^{-1}U^TY$,

$$\hat{Y} = X\hat{\beta} = (U\Sigma V^T)(V\Sigma^{-1}U^T)Y$$
$$= UU^TY$$

$$\boxed{\hat{Y} = UU^TY = HY}$$

$H = UU^T$ "Hat Matrix"



$\hat{Y} = UU^TY$

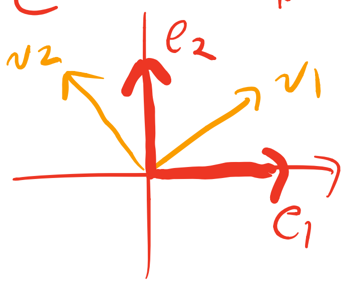
Projection!

$\hat{Y}$ is the closest pt in column space to Y

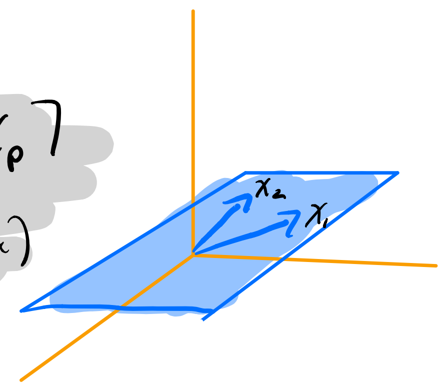
$\|Y - \hat{Y}\| = $ Squared Error $= \sqrt{\sum (y_i - \hat{y}_i)^2}$

Decomposition of Linear Transformation

$\mathcal{E} = e_1 \,..\, e_p$ basis



$$X = [x_1 \,..\, x_p] \xrightarrow[\text{1-1 (full rank)} \\ \text{not onto}]{}$$

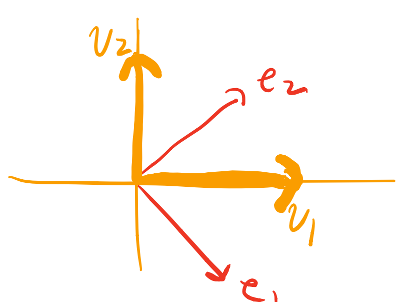


V^T 1-1 onto
rotation

Change of basis!

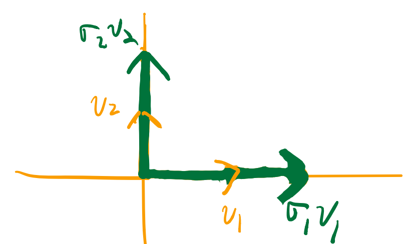
U 1-1
not onto
rotation

V



$V = [V_1 \,..\, V_p]$ basis

$$\xrightarrow[\text{scale}]{\Sigma}$$



i.e. $X = U \Sigma V^T$