Singular Value decomposition.

Diagonalization: $A = X \Lambda X^{-1}$

(theoretical) issues: A has to be square

A has to have a full set of (real) eigenvalues

A has to have a full set of eigenvectors

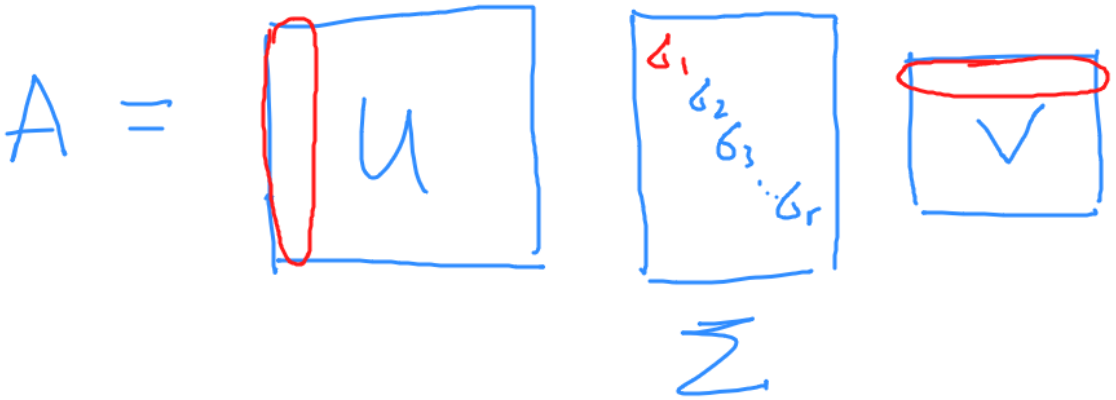
(practical) issue: tiny changes in A may lead to huge changes in X, Λ

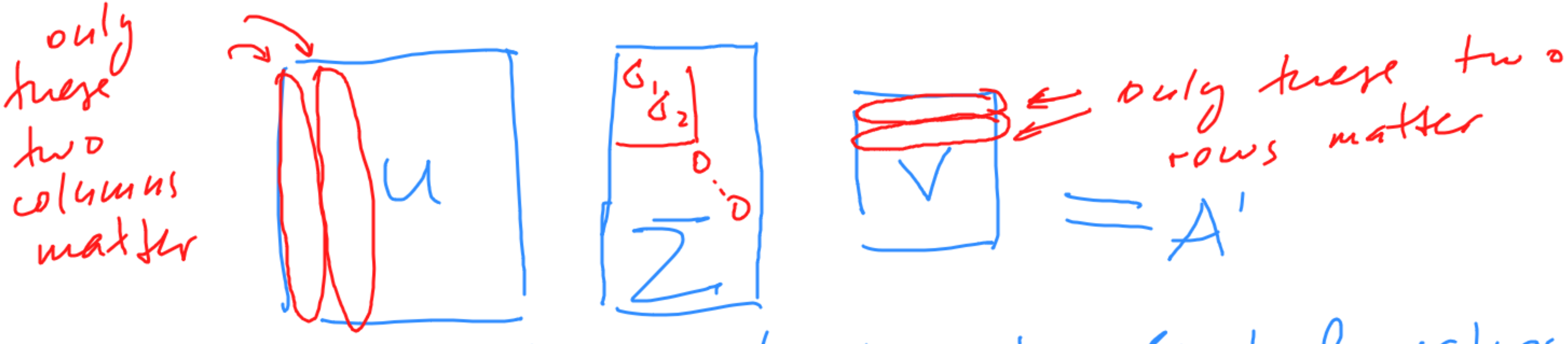
The singular value decomposition resolves all of these issues:

$$\underset{m \times n}{A} = \underset{\substack{m \times m \\ \text{orthogonal}}}{U} \; \underset{\substack{m \times n \\ \text{diagonal}}}{\Sigma} \; \underset{\substack{n \times n \\ \text{orthogonal}}}{V}$$

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

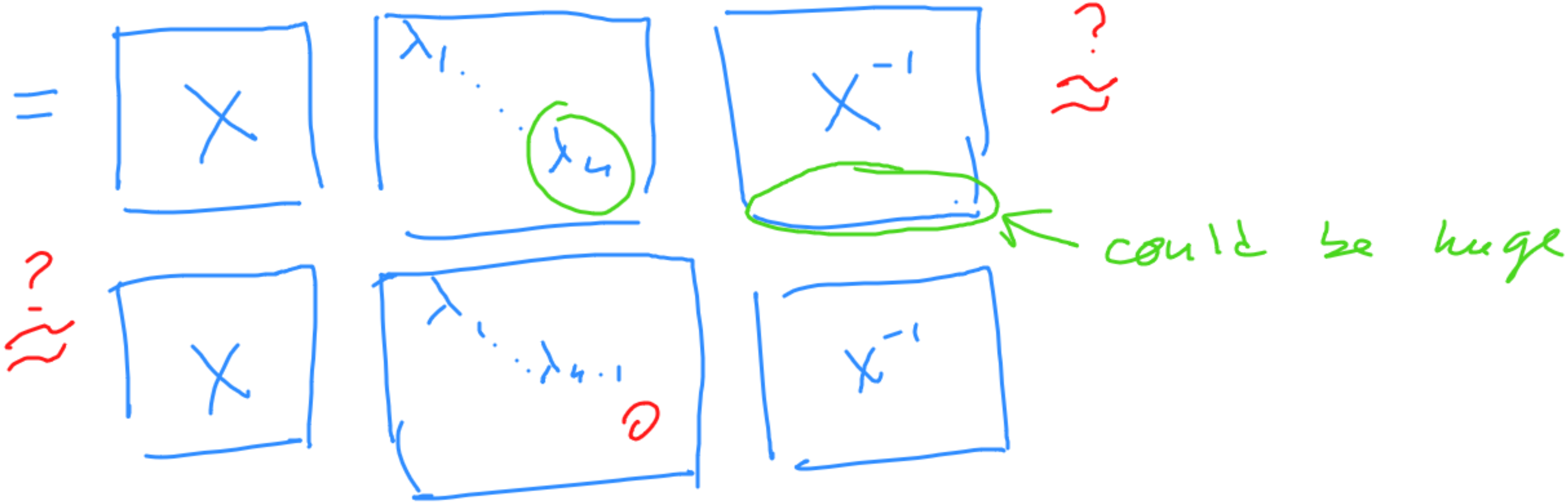
$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$$



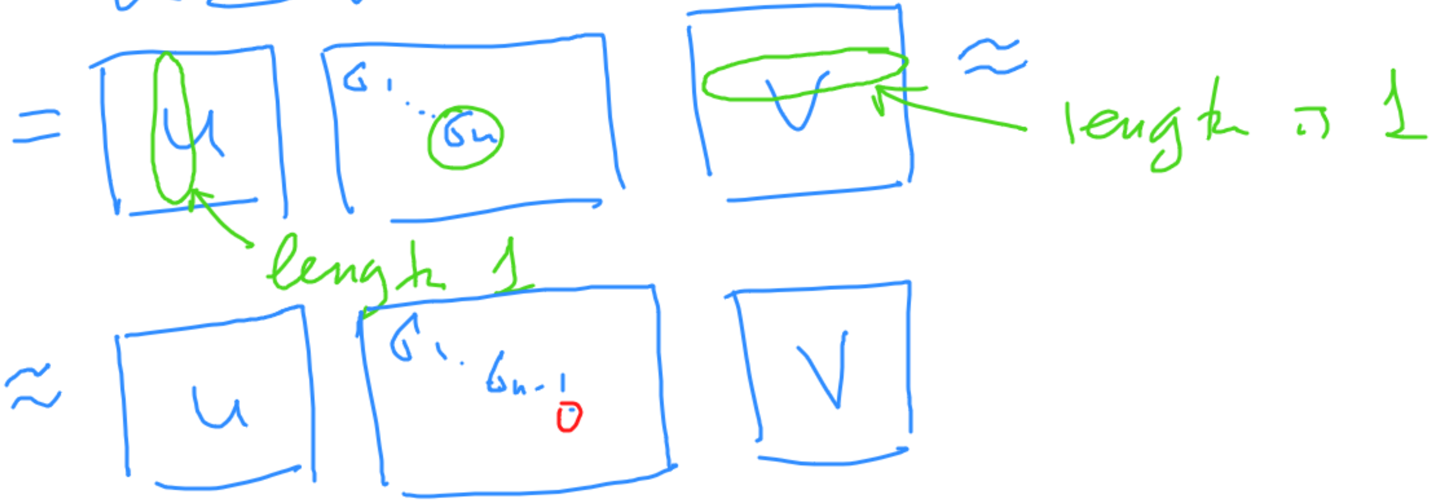


To approximate A by A', keep the first ℓ values of σ_i's, replace the rest σ_i's with 0's. Then the only parts of U, V that matter in the product $A' = U\Sigma' V$ are the first ℓ columns of U and the first ℓ rows of V.

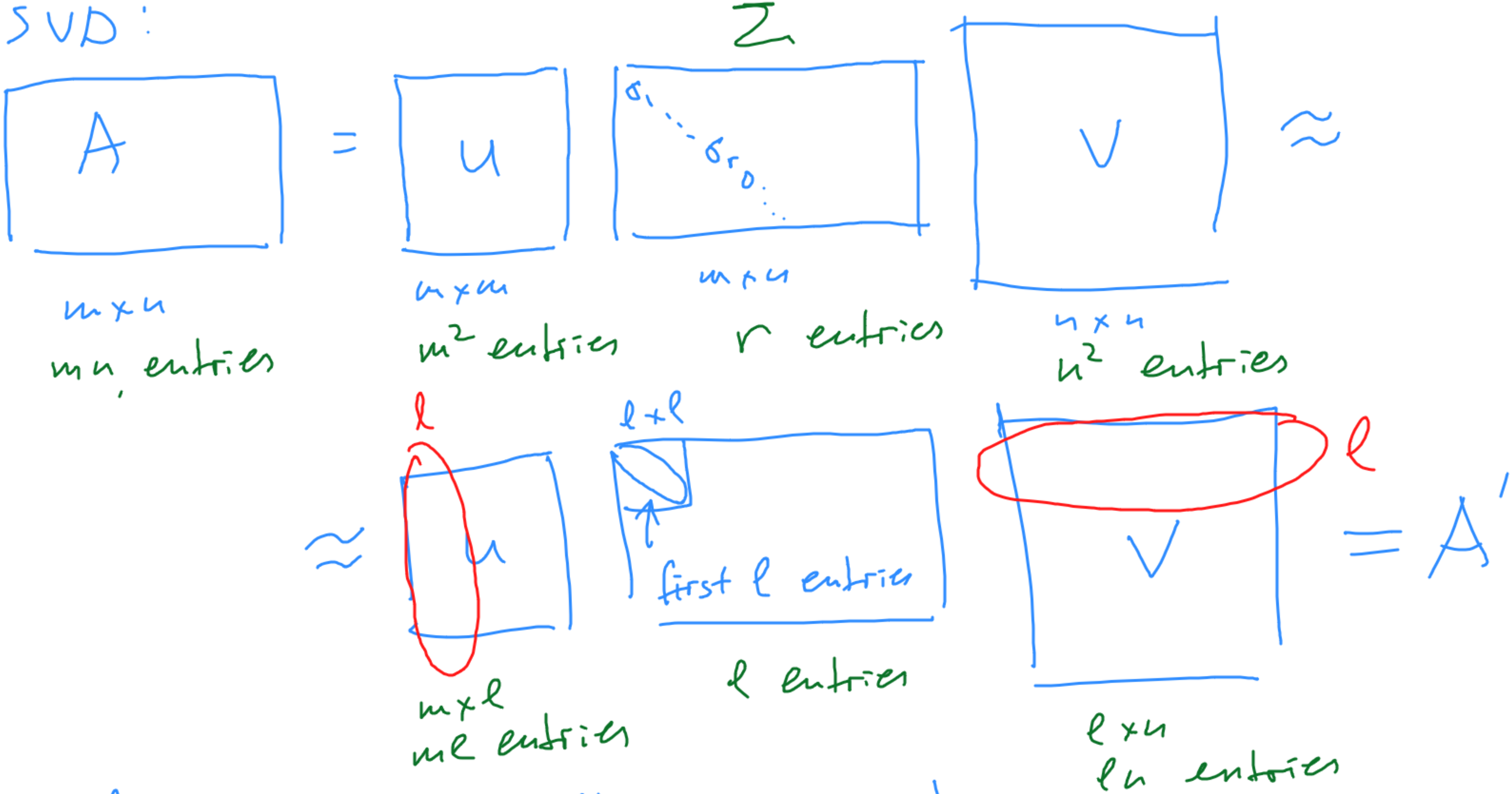
Diagonalization: $A = X \Lambda X^{-1} =$



In SVD: $A = U \Sigma V$



SVD:



information about A' is carried by $m\ell + \ell + \ell n = \ell(m+n+1)$ entries

How to find SVD-decomp.? Why does every matrix have SVD-decomp.?

SVD produces vectors $u_1, u_2, \dots u_m$ (orthonormal system)

$v_1, v_2 \dots v_n$ (orthonormal system)

singular values of A → numbers $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \dots \geq \sigma_r > 0$ s.t.

$$Av_i = \sigma_i u_i \quad \text{for } i = 1 \dots r$$

$$Av_i = 0 \quad \text{for } i > r$$

in matrix form: $AV = U\Sigma$ where

$$U = \begin{bmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_m \\ | & | & & | \end{bmatrix} \quad \Sigma = \begin{bmatrix} \sigma_1 & & & & & \\ & \ddots & & & & \\ & & \sigma_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix} \quad V = \begin{bmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{bmatrix}$$

orthogonal   orthogonal

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

$\sigma_1, \sigma_2, \dots, \sigma_n$ are called the singular values A A.

Finding v_i, σ_i, u_i :

$u_1, u_2, \dots, u_r$ are an orthonormal basis for $C(A)$ (col. space)

$u_{r+1}, \dots, u_m$ are an orthonormal basis for $N(A^T)$ (left nullspace)

$v_1, \dots, v_r$ are an orthonormal basis for $C(A^T)$ (row space)

$v_{r+1}, \dots, v_n$ are an orthonormal basis for $N(A)$ (nullspace)

Suppose $A = U\Sigma V^T$

Compute $\underbrace{A^TA}_{\text{symmetric}} = (U\Sigma V^T)^T(U\Sigma V^T) =$

$$= V\Sigma^T\underbrace{U^TU}_{=I}\Sigma V^T = V\underbrace{\Sigma^T\Sigma}_{\text{diagonal square}}V^T$$

orthogonal

$$\Sigma^T\Sigma = \begin{bmatrix}\sigma_1^2 & & & & \\ & \sigma_2^2 & & & \\ & & \ddots & & \\ & & & \sigma_r^2 & \\ & & & & \ddots\ 0\end{bmatrix}$$

diagonal

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

orthogonal

so this product is the orthogonal diagonalization of symmetric $S = A^TA$ (and eigenvalues ≥ 0).

To find V, diagonalize S!

$$S = A^TA = Q\Lambda Q^T$$

rows of Q^T are v_i

$\sigma_i = \sqrt{\lambda_i}$

$$A^TAv_j = \lambda_j v_j = \sigma_j^2 v_j\ , \quad \|v_j\| = 1$$

Then u_i are found from $(i=1,\dots,r)$

$$Av_i = \sigma_i u_i$$

$u_1,\dots,u_r$ automatically will be an orthonormal system!

$$u_i^T u_j = \left(\frac{Av_i}{\sigma_i}\right)^T \left(\frac{Av_j}{\sigma_j}\right) = \frac{v_i^T (A^T A v_j)}{\sigma_i \sigma_j} = \frac{v_i^T \sigma_j^2 v_j}{\sigma_i \sigma_j} =$$

$$= \frac{\sigma_j^2}{\sigma_i \sigma_j} v_i^T v_j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

1) Find orthonormal system of eigenvectors $v_1,\dots,v_r$ for A^TA with nonzero eigenvalues: $\sigma_1^2 \geq \sigma_2^2 \geq \sigma_3^2 \geq \dots \geq \sigma_r^2 > 0$

2) Compute u_i from $Av_j = \sigma_j u_j$ for $j=1,\dots,r$.

3) Take any orthonormal basis for left nullspace $N(A^T)$ $u_{r+1},\dots,u_m$

4) Take any orthonormal basis for nullspace $N(A)$ $v_{r+1},\dots,v_n$