

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:

$$A = U \Sigma V$$

$m \times n$ $m \times m$ $m \times n$ $n \times n$
orthogonal diagonal orthogonal

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

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

$$A = \begin{bmatrix} | & \\ & U \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots \\ & & & \sigma_r \end{bmatrix} \begin{bmatrix} \hline \\ & V \end{bmatrix}$$

only these two columns matter

$$\begin{bmatrix} \sigma_1 & & \\ \hline \sigma_2 & & \\ & 0 & \\ & & \ddots \\ & & & 0 \end{bmatrix}$$

only these two rows matter

$$= A'$$

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

$$= \begin{bmatrix} X \\ \sim? \\ X \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \\ & & & 0 \end{bmatrix} \begin{bmatrix} X^{-1} \\ \sim? \\ X^{-1} \end{bmatrix}$$

could be huge

In SVD: $A = U \Sigma V$

$$= \begin{bmatrix} U \\ \text{length } 1 \\ U \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ & & & 0 \end{bmatrix} \begin{bmatrix} V \\ \approx \text{length } 1 \\ V \end{bmatrix}$$

\approx length ≈ 1

SVD:

$$\begin{array}{ccccc}
 \boxed{A} & = & \boxed{U} & \boxed{\Sigma} & \boxed{V} \approx \\
 m \times n & & m \times m & m \times n & n \times n \\
 mn \text{ entries} & & m^2 \text{ entries} & r \text{ entries} & n^2 \text{ entries}
 \end{array}$$

Diagram of Σ : A rectangle with a dashed diagonal line from the top-left to the bottom-right, labeled $\sigma_1, \sigma_2, \dots$.

$$\begin{array}{ccccc}
 \approx & \boxed{U} & \boxed{\Sigma} & \boxed{V} & = A' \\
 m \times l & l \times l & l \times n & l \times n & \\
 ml \text{ entries} & l \text{ entries} & & &
 \end{array}$$

Diagram of U : A rectangle with a red oval around the first l rows, labeled l at the top.

Diagram of Σ : A rectangle with a red oval around the top-left corner, labeled $l+l$ at the top-left and "first l entries" below it.

Diagram of V : A rectangle with a red oval around the first l columns, labeled l at the top-right.

information about A' is carried
by $ml + l + ln = l(m+n+1)$ entries

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

SVD produces vector u_1, u_2, \dots, u_m (orthonormal system)
 singular values of A v_1, v_2, \dots, v_n (orthonormal system)
 \rightarrow number $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \dots \geq \sigma_r > 0$ s.t.

$$\begin{aligned} A v_i &= \delta_i u_i & \text{for } i = 1 \dots r \\ A v_i &= 0 & \text{for } i > r \end{aligned}$$

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

$U = \begin{bmatrix} u_1 & u_2 & \dots & u_m \\ 1 & 1 & & 1 \end{bmatrix}$ $\Sigma = \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_m & \\ & & & 0 \dots 0 \end{bmatrix}$ $V = \begin{bmatrix} 1 & 1 & \dots & 1 \\ v_1 & v_2 & \dots & v_n \\ 1 & 1 & & 1 \end{bmatrix}$

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 of A .

Finding v_i, G_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 $A^T A = (U \Sigma V^T)^T (U \Sigma V^T) =$

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

$\underbrace{A^T A}_{\text{symmetric}}$
 $\underbrace{V}_{\text{orthogonal}}$
 $\underbrace{\Sigma^T \Sigma}_{\text{diagonal "square"}}$
 $\underbrace{V^T}_{=V^{-1} \text{ orthogonal}}$

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

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

To find V , diagonalize S :

$$S = A^T A = Q \Lambda Q^T \leftarrow \text{rows of } Q^T \text{ are } \frac{v_i}{\sigma_i = \sqrt{\lambda_i}}$$

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

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

$$A v_i = \sigma_i u_i$$

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

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

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

- 1) Find orthonormal system of eigenvectors v_1, \dots, v_r for $A^T A$ 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 $A v_j = \sigma_j u_j$ 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