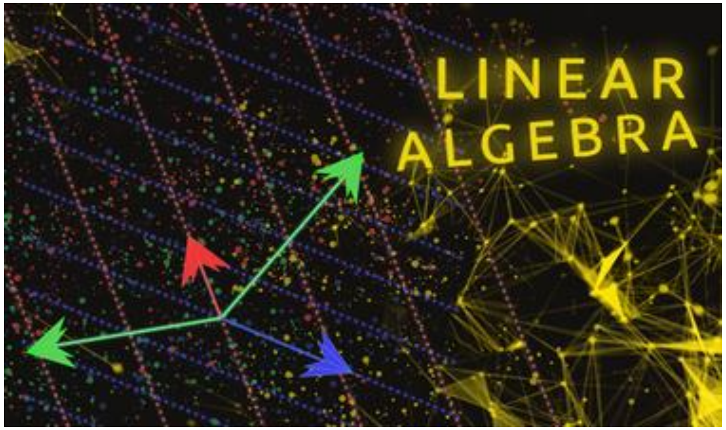
دانشگاه صنعتی نوشیروانی بابل



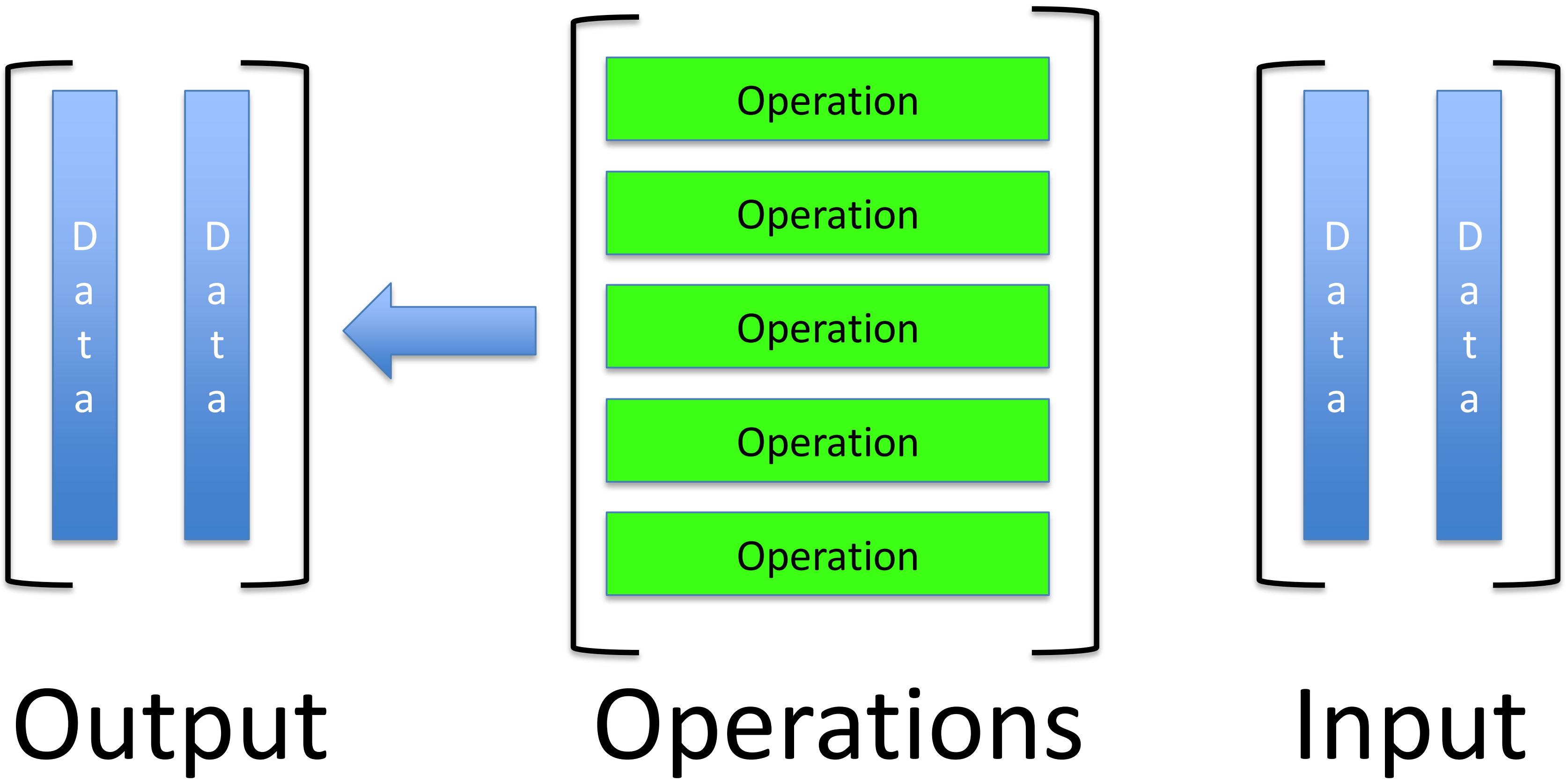
Linear Algebra

Singular Value Decomposition

Nooshin Maghsoodi

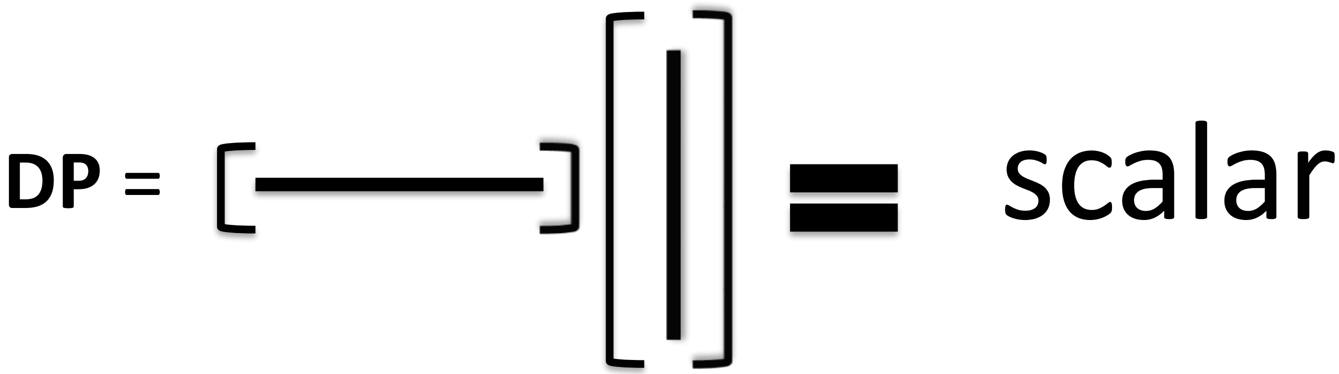
Noshirvani University

A representational view of matrices



Inner Product = Dot product

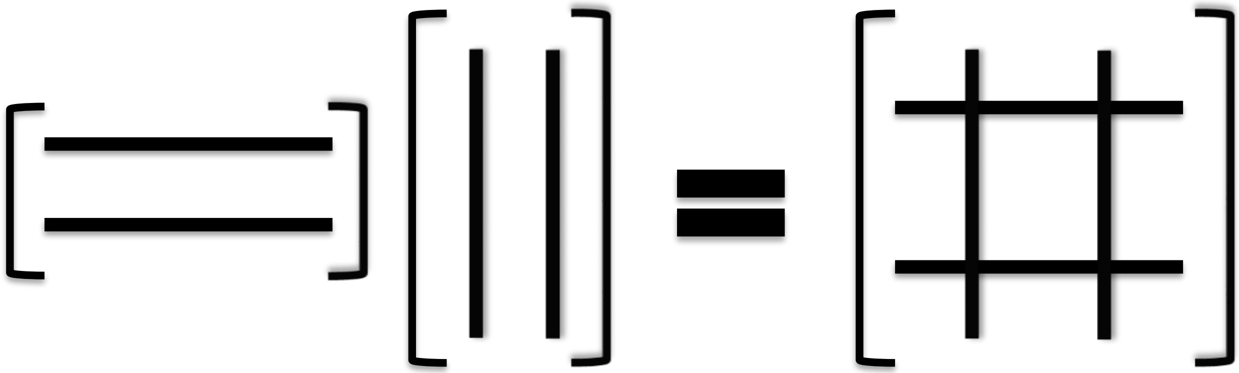
Dot product = Row Vector * Column Vector = Scalar



[1xn] [mx1] [1x1] (m=n)

$$\begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 5 \\ 6 \\ 7 \\ 8 \end{bmatrix} = 1*5 + 2*6 + 3*7 + 4*8 = 70$$

Matrix multiplication

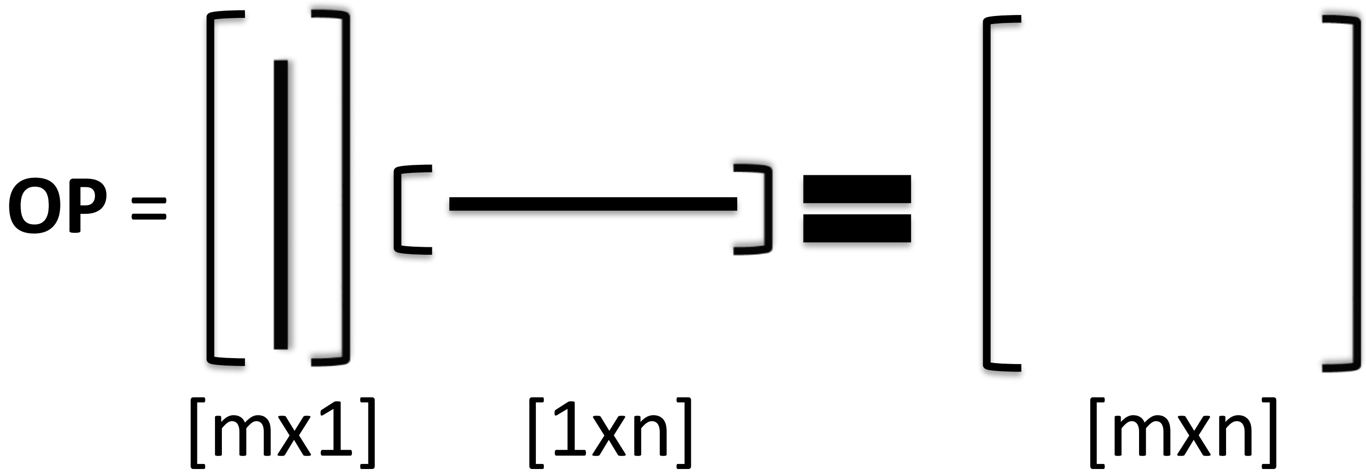


$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} (1*5 + 2*7) & (1*6 + 2*8) \\ (3*5 + 4*7) & (3*6 + 4*8) \end{bmatrix}$$

The outer product

- What if we do matrix multiplication, but when the two matrices are a single column and row vector?



- Output is a *matrix*, not a scalar.

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 3 & 4 \end{bmatrix} = \begin{bmatrix} (1*3) & (1*4) \\ (2*3) & (2*4) \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix}$$

Remember this?

Orthogonal matrices

- square shape (dimensionality-preserving)
- rows are orthogonal unit vectors
- columns are orthogonal unit vectors
- perform a rotation (with possible axis inversion)
- preserve vector lengths and angles
- inverse is transpose (also orthogonal)

Identity matrix

Diagonal matrices

- arbitrary rectangular shape
- off-diagonal entries are zero
- squeeze/stretch along standard axes
- can create/discard standard axes
- (pseudo) inverse: diagonal, with inverse of non-zero diagonal entries of original

Singular Value Decomposition (SVD)

$$M = U \Sigma V^T$$

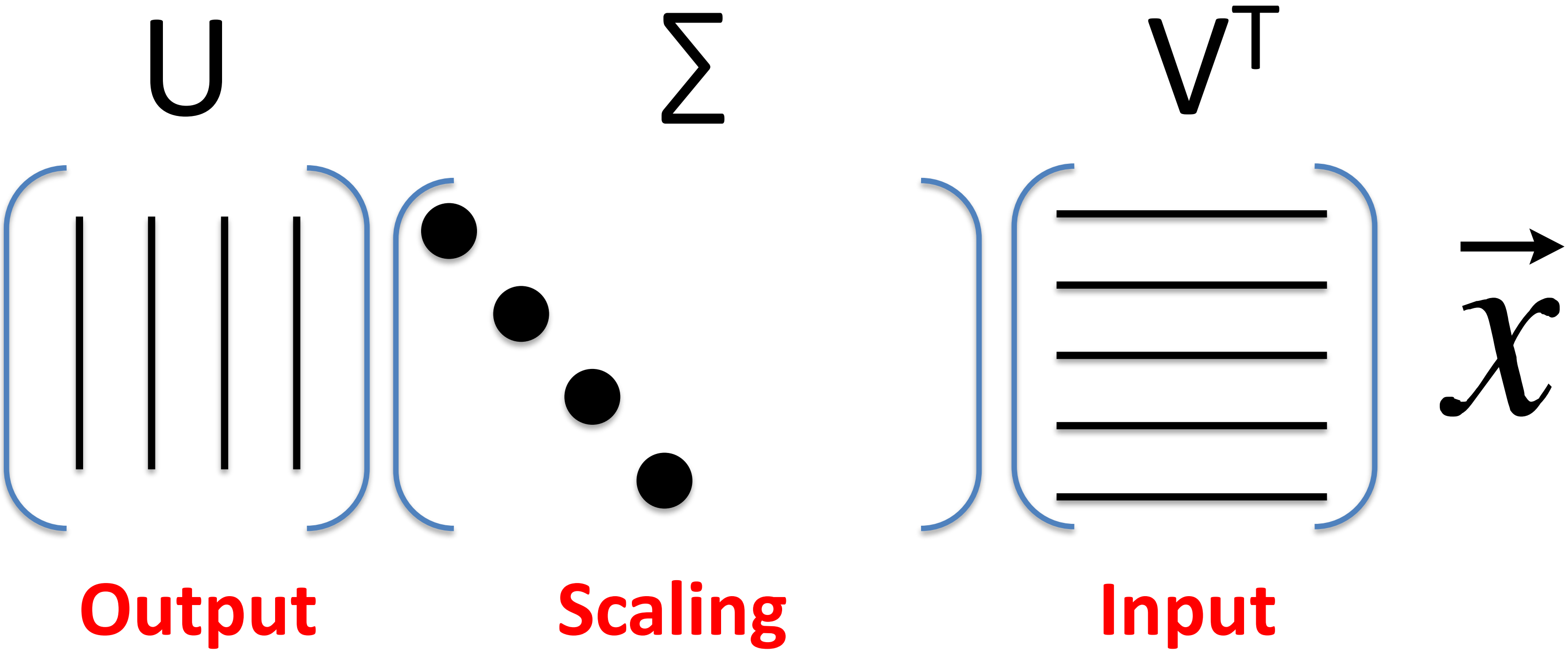
U: Orthogonal
("Rotate")

Σ: Diagonal
("Stretch")

V^T: Orthogonal
("Rotate")

SVD

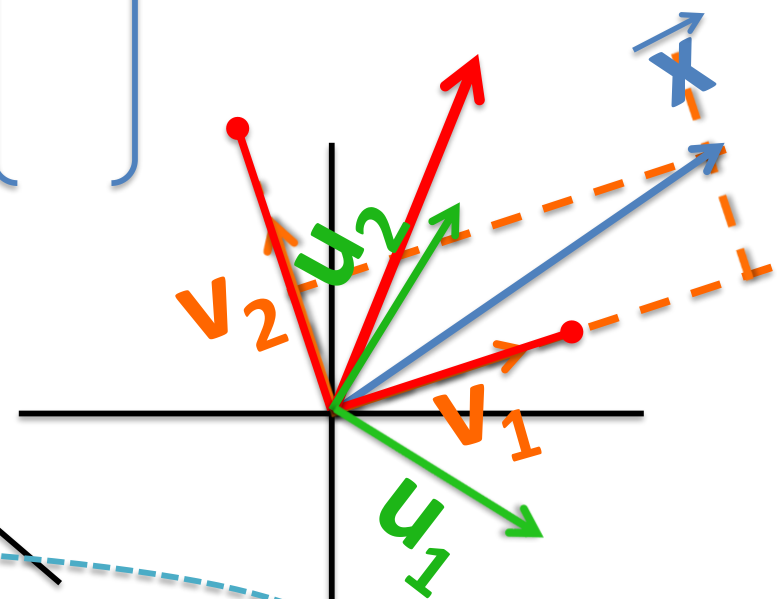
$$M\vec{x}$$



A simple case

$$V = \begin{pmatrix} | & | \\ \hat{v}_1 & \hat{v}_2 \\ | & | \end{pmatrix}$$

$$\underset{U}{\begin{pmatrix} | & | \\ \hat{u}_1 & \hat{u}_2 \\ | & | \end{pmatrix}} \underset{\Sigma}{\begin{pmatrix} s_1 & \\ & s_2 \end{pmatrix}} \underset{V^T}{\begin{pmatrix} - \hat{v}_1^T - \\ - \hat{v}_2^T - \end{pmatrix}} \underset{\vec{x}}{\begin{pmatrix} \\ \\ \end{pmatrix}}$$



Outer Product!

$$\hat{u}_1 s_1 (\hat{v}_1^T)\vec{x} \qquad \begin{bmatrix} s_1(\hat{v}_1^T \bullet \vec{x}) \\ s_2(\hat{v}_2^T \bullet \vec{x}) \end{bmatrix} \begin{bmatrix} \hat{v}_1^T \bullet \vec{x} \\ \hat{v}_2^T \bullet \vec{x} \end{bmatrix}$$

$$+$$

$$\hat{u}_2 s_2 (\hat{v}_2^T)\vec{x}$$

SVD can be interpreted as

- A sum of outer products!
- Decomposing the matrix into a sum of scaled outer products.
- Key insight: The operations on respective dimensions stay separate from each other, all the way – through v, $\sum$ and u.
- They are grouped, each operating on another piece of the input.

Why does this perspective matter? $\vec{x}$

$$U \quad \Sigma \quad V^T \quad \vec{x}$$

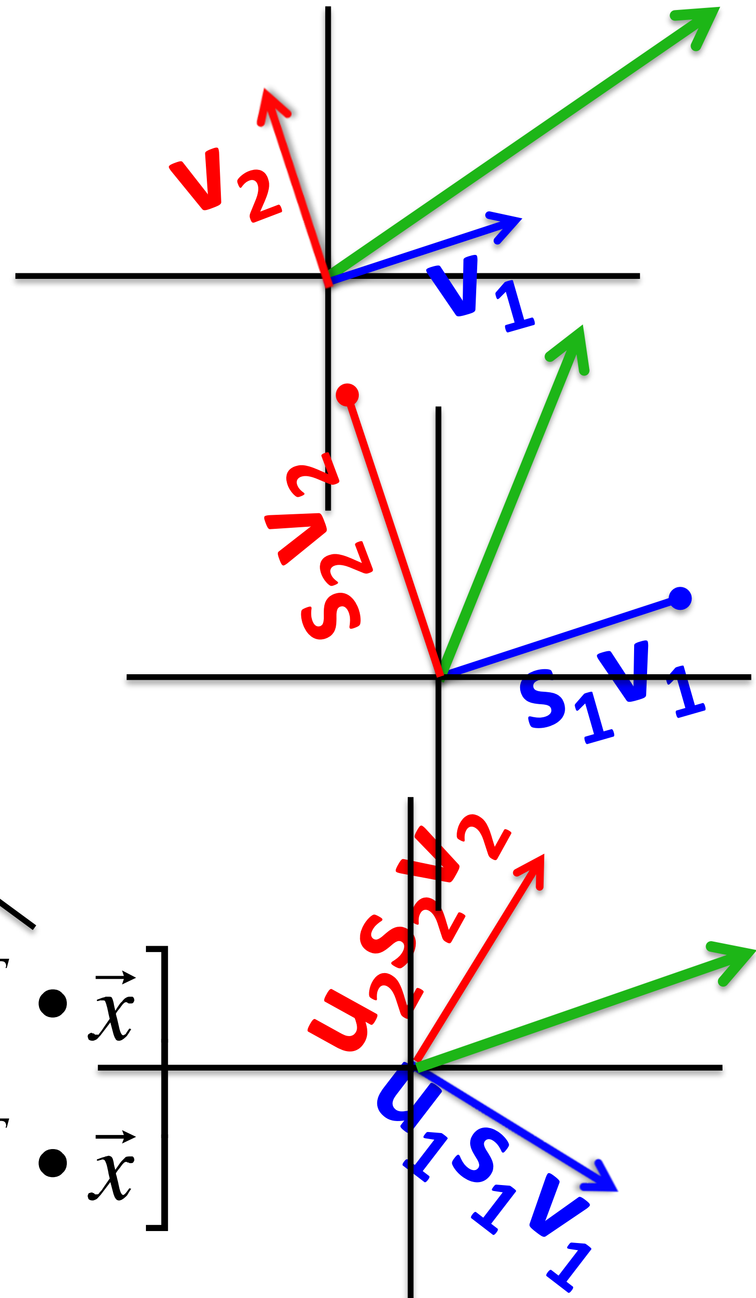
$$\begin{pmatrix} | & | \\ \hat{u}_1 & \hat{u}_2 \\ | & | \end{pmatrix} \begin{pmatrix} s_1 & \\ & s_2 \end{pmatrix} \begin{pmatrix} - \hat{v}_1^T - \\ - \hat{v}_2^T - \end{pmatrix} \begin{pmatrix} \\ \\ \end{pmatrix}$$

$$s_1(\hat{u}_1\hat{v}_1^T)\vec{x}$$

$$+$$

$$s_2(\hat{u}_2\hat{v}_2^T)\vec{x}$$

$$\begin{bmatrix} s_1(\hat{v}_1^T \bullet \vec{x}) \\ s_2(\hat{v}_2^T \bullet \vec{x}) \end{bmatrix} \begin{bmatrix} \hat{v}_1^T \bullet \vec{x} \\ \hat{v}_2^T \bullet \vec{x} \end{bmatrix}$$



Singular Value Decomposition

For an $m \times n$ matrix $\mathbf{A}$ of rank r there exists a factorization (Singular Value Decomposition = **SVD**) as follows:

$$A = U \Sigma V^T$$

$m \times m$ | $m \times n$ | V is $n \times n$

The columns of ***V*** are orthogonal eigenvectors of $\boldsymbol{A^T A}$.

Eigenvalues $\lambda_1 \ldots \lambda_r$ of $\boldsymbol{AA^T}$ are the eigenvalues of $\boldsymbol{A^T A}$.

$$\sigma_i = \sqrt{\lambda_i}$$

Singular values of A

$$\Sigma = diag(\sigma_1 \ldots \sigma_r)$$

Assume they are arranged in decreasing order and r is rank of A

The columns of ***U*** are orthogonal eigenvectors of $\boldsymbol{AA^T}$.

Its first r rows equal:

$$\mathbf{u}_i = \frac{1}{\|A\mathbf{v}_i\|} A\mathbf{v}_i = \frac{1}{\sigma_i} A\mathbf{v}_i$$

Singular Value Decomposition

The decomposition of A involves an $m \times n$ "diagonal" matrix Σ of the form

$$\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \quad (3)$$

$\leftarrow m - r$ rows

$n - r$ columns

where D is an $r \times r$ diagonal matrix for some r not exceeding the smaller of m and n. (If r equals m or n or both, some or all of the zero matrices do not appear.)

Singular Value Decomposition

The decomposition of A involves an $m \times n$ "diagonal" matrix Σ of the form

$$\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \quad (3)$$

← $m - r$ rows

↑ $n - r$ columns

where D is an $r \times r$ diagonal matrix for some r not exceeding the smaller of m and n. (If r equals m or n or both, some or all of the zero matrices do not appear.)

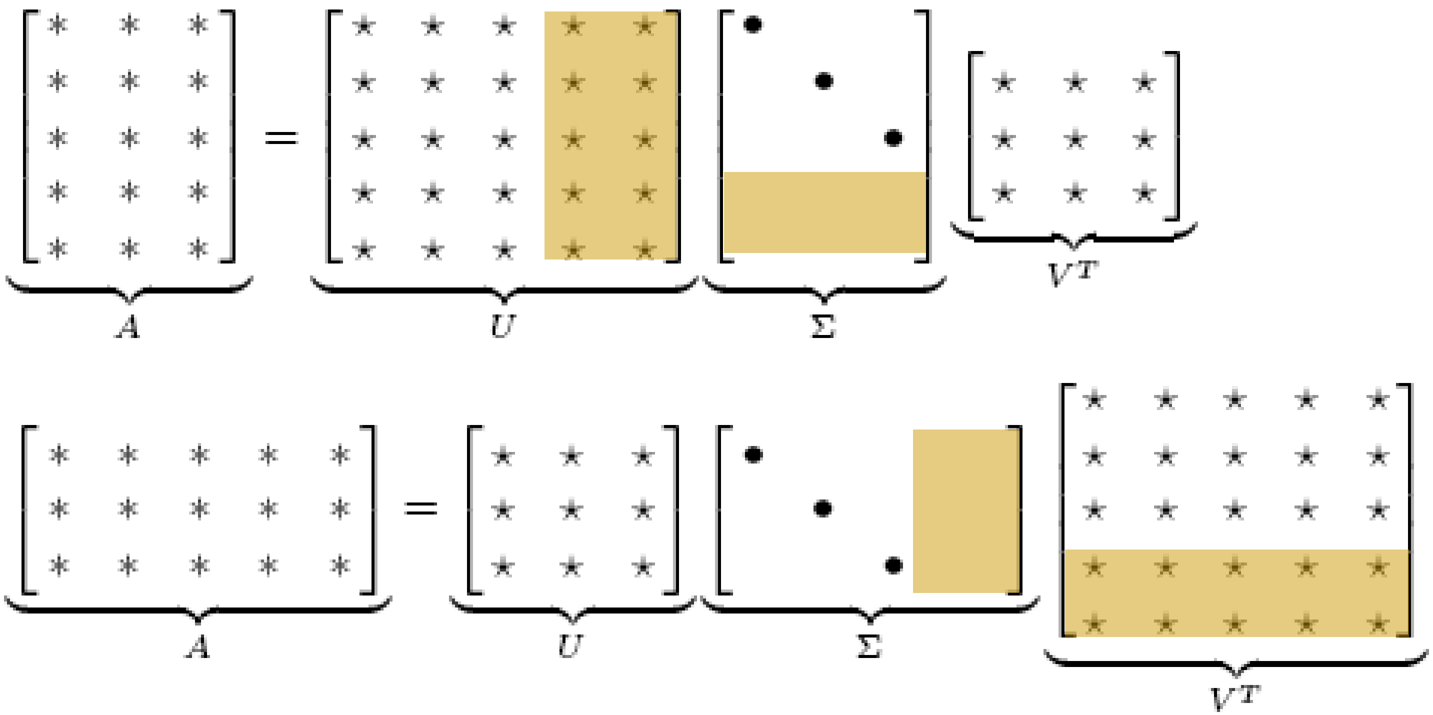
The Singular Value Decomposition

Let A be an $m \times n$ matrix with rank r. Then there exists an $m \times n$ matrix Σ as in (3) for which the diagonal entries in D are the first r singular values of A, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$, and there exist an $m \times m$ orthogonal matrix U and an $n \times n$ orthogonal matrix V such that

$$A = U \Sigma V^T$$

Singular Value Decomposition

- Illustration of SVD dimensions and sparseness



SVD example

Let $A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$

***Step 1.* Find an orthogonal diagonalization of A^TA.**

$$A^TA = \begin{bmatrix} 4 & 8 \\ 11 & 7 \\ 14 & -2 \end{bmatrix} \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} = \begin{bmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{bmatrix}$$

$\lambda_1 = 360, \lambda_2 = 90, \text{and } \lambda_3 = 0$

$$\mathbf{v}_1 = \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 2/3 \\ -2/3 \\ 1/3 \end{bmatrix}$$

SVD example

Let $A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$

Step 2. Set up V and Σ

$$V = [\,\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3\,] = \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{bmatrix}$$

$$\sigma_1 = 6\sqrt{10}, \quad \sigma_2 = 3\sqrt{10}, \quad \sigma_3 = 0$$

$$D = \begin{bmatrix} 6\sqrt{10} & 0 \\ 0 & 3\sqrt{10} \end{bmatrix}, \qquad \Sigma = [\,D \quad 0\,] = \begin{bmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{bmatrix}$$

SVD example

Let $$A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$$

Step 3. ***Construct U.***

$$\mathbf{u}_1 = \frac{1}{\sigma_1} A\mathbf{v}_1 = \frac{1}{6\sqrt{10}} \begin{bmatrix} 18 \\ 6 \end{bmatrix} = \begin{bmatrix} 3/\sqrt{10} \\ 1/\sqrt{10} \end{bmatrix}$$

$$\mathbf{u}_2 = \frac{1}{\sigma_2} A\mathbf{v}_2 = \frac{1}{3\sqrt{10}} \begin{bmatrix} 3 \\ -9 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{10} \\ -3/\sqrt{10} \end{bmatrix}$$

$$A = \underbrace{\begin{bmatrix} 3/\sqrt{10} & 1/\sqrt{10} \\ 1/\sqrt{10} & -3/\sqrt{10} \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} 1/3 & 2/3 & 2/3 \\ -2/3 & -1/3 & 2/3 \\ 2/3 & -2/3 & 1/3 \end{bmatrix}}_{V^T}$$

■