

**Theorem**

Any  $A$  an  $m \times n$  matrix can be written as a product

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

where:

- $U = [\mathbf{u}_1 \ \dots \ \mathbf{u}_m]$  is an  $m \times m$  orthogonal matrix.
- $V = [\mathbf{v}_1 \ \dots \ \mathbf{v}_n]$  is an  $n \times n$  orthogonal matrix.
- $\Sigma$  is an  $m \times n$  matrix of the following form:

$$\begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & \sigma_n \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} \sigma_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & \sigma_m & 0 & \dots & 0 \end{bmatrix}$$

(if  $n \leq m$ ) (if  $n \geq m$ )

where  $\sigma_1 \geq \sigma_2 \geq \dots \geq 0$ .

**Note.**

- The numbers  $\sigma_1, \sigma_2, \dots$  are called *singular values* of  $A$ .
- The vectors  $\mathbf{u}_1, \dots, \mathbf{u}_m$  are called *left singular vectors* of  $A$ .
- Then the vectors  $\mathbf{v}_1, \dots, \mathbf{v}_n$  are called *right singular vectors* of  $A$ .
- The formula  $A = U\Sigma V^T$  is called a *singular value decomposition (SVD)* of  $A$ .
- The matrix  $\Sigma$  is uniquely determined, but  $U$  and  $V$  depend on some choices.

### Theorem

Let  $A$  be a matrix with a singular value decomposition

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

If

$$U = [ \mathbf{u}_1 \quad \dots \quad \mathbf{u}_m ] \quad V = [ \mathbf{v}_1 \quad \dots \quad \mathbf{v}_n ]$$

and  $\sigma_1, \dots, \sigma_r$  are singular values of  $A$  then then

$$A = \sigma_1(\mathbf{u}_1\mathbf{v}_1^T) + \sigma_2(\mathbf{u}_2\mathbf{v}_2^T) + \dots + \sigma_r(\mathbf{u}_r\mathbf{v}_r^T)$$

## Application: Image compression



- The size of this image is  $800 \times 700$  pixels.
- The color of each pixel is represented by an integer between 0 (black) and 255 (white).
- The whole image is described by a matrix  $A$  consisting of  $800 \times 700 = 560,000$  numbers.
- Each number is stored in 1 byte, so the image file size is 560,000 bytes ( $\approx 0.53$  MB).

**How to make the image file smaller:**

1) Compute SVD of the matrix  $A$ :

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

where

$$U = \begin{bmatrix} \mathbf{u}_1 & \dots & \mathbf{u}_m \end{bmatrix} \quad V = \begin{bmatrix} \mathbf{v}_1 & \dots & \mathbf{v}_n \end{bmatrix}$$

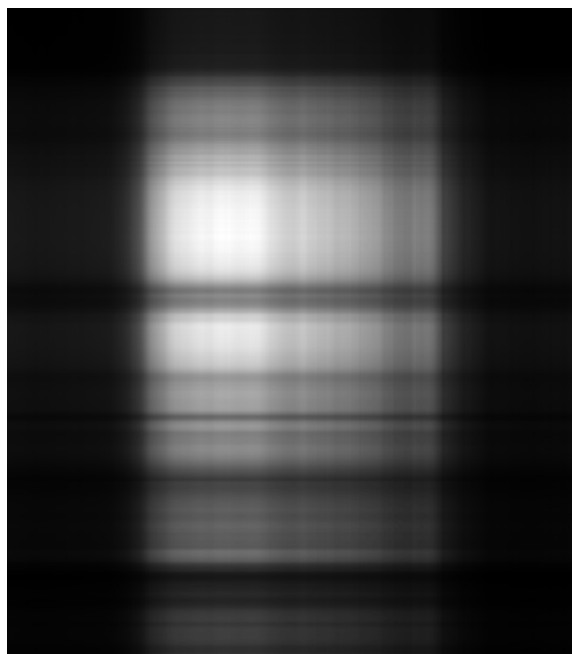
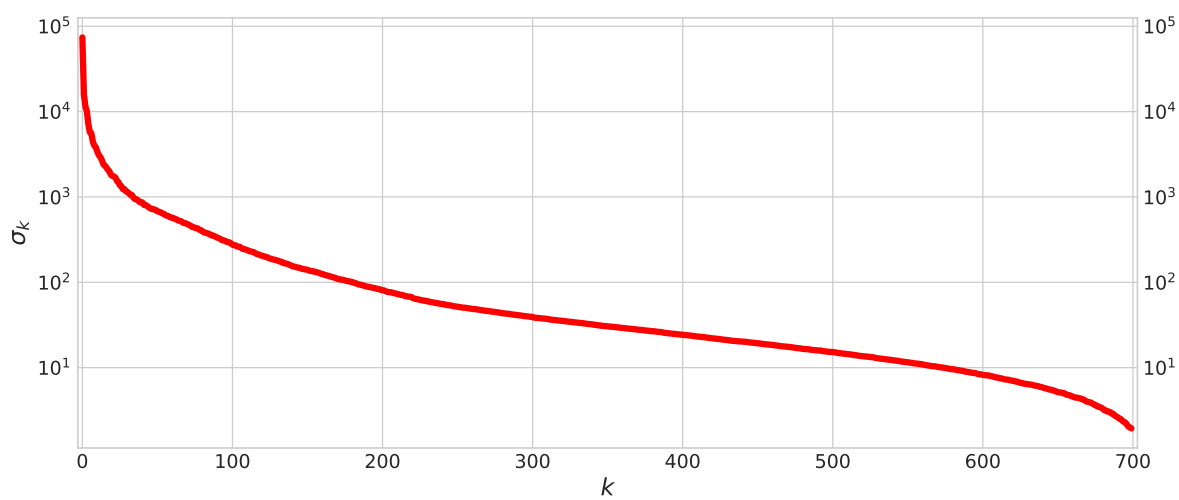
and  $\sigma_1, \dots, \sigma_r$  are singular values of  $A$ .

2) Replace  $A$  by the matrix

$$B_k = \sigma_1(\mathbf{u}_1\mathbf{v}_1^T) + \dots + \sigma_k(\mathbf{u}_k\mathbf{v}_k^T)$$

for some  $1 \leq k \leq 700$ . This matrix can be stored using  $k \cdot (800 + 700 + 1)$  numbers.

## Singular values of the matrix A



**matrix  $B_1$**   
1501 bytes  
compression 374:1



**matrix  $B_5$**   
7905 bytes  
compression 75:1



**matrix  $B_{10}$**   
15,010 bytes  
compression 37:1



**matrix  $B_{20}$**   
30,020 bytes  
compression 18:1



**matrix  $B_{50}$**   
75,050 bytes  
compression 7:1



**matrix  $B_{100}$**   
150,100 bytes  
compression 4:1

## How to compute SVD of a matrix $A$

## How to compute SVD of a matrix $A$

- 1) Compute an orthogonal diagonalization of the symmetric  $n \times n$  matrix  $A^T A$ :

$$A^T A = Q D Q^T$$

such that eigenvalues on the diagonal of the matrix  $D$  are arranged from the largest to the smallest. We set  $V = Q$ .

- 2) If

$$D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

then  $\sigma_i = \sqrt{\lambda_i}$ . This gives the matrix  $\Sigma$ .

Note: if  $n > m$  then we use only  $\lambda_1, \dots, \lambda_m$ . The remaining eigenvalues  $\lambda_{m+1}, \dots, \lambda_n$  of  $D$  will be equal to 0 in this case.

- 3) Let  $V = [\mathbf{v}_1 \ \dots \ \mathbf{v}_n]$ , and let  $\sigma_1, \dots, \sigma_r$  be non-zero singular values of  $A$ . The first  $r$  columns of the matrix  $U = [\mathbf{u}_1 \ \dots \ \mathbf{u}_m]$  are given by

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

The remaining columns  $\mathbf{u}_{r+1}, \dots, \mathbf{u}_m$  can be added arbitrarily so that  $U$  is an orthogonal matrix (i.e.  $\{\mathbf{u}_1, \dots, \mathbf{u}_m\}$  is an orthonormal basis of  $\mathbb{R}^m$ ).

**Example.** Find SVD of the following matrix:

$$A = \begin{bmatrix} -1 & 0 \\ 1 & -1 \\ 0 & 1 \end{bmatrix}$$