Singular Value Decomposition

Motivatation

- The diagonalization theorem play a part in many interesting applications.
- Unfortunately not all matrices can be factored as $A = PDP^{-1}$
- However a factorization $A = QDP^{-1}$ is possible for any $m \times n$ matrix A.
- A special factorization of this type, called singular value decomposition, is one of the most useful matrix factorizations in applied linear algebra.
- As we will see one application is finding minimum value of Ax for solving Ax=0.

Singular Value Decomposition

- Let A be a $m \times n$ matrix of rank r. We cannot find eigenvalues and eigenvectors for non-square matrices.
- However, A can be diagonalized in such a way that

$$\begin{gathered} A\boldsymbol{v}_1 = \sigma_1 \boldsymbol{u}_1 \\ A\boldsymbol{v}_2 = \sigma_2 \boldsymbol{u}_2 \\ \vdots \\ A\boldsymbol{v}_r = \sigma_r \boldsymbol{u}_r \end{gathered}$$

- The singular vectors $\boldsymbol{v}_1 \ldots \boldsymbol{v}_r$ are orthogonal and are the basis for the row space of A.
- The output vectors $\boldsymbol{u}_1 \ldots \boldsymbol{u}_r$ are orthogonal and in column space of A.
- The singular values $\sigma_1 \ldots \sigma_r$ are all positive numbers in non-increasing order.

Singular value decomposition

- $A\boldsymbol{v}_i = \sigma_i \boldsymbol{u}_i$ leads to $AV = U\Sigma$:

$$A\begin{bmatrix} \boldsymbol{v}_1 & \dots & \boldsymbol{v}_r \end{bmatrix} = \begin{bmatrix} \boldsymbol{u}_1 & \dots & \boldsymbol{u}_r \end{bmatrix}\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix}$$

- Note that in above equation the dimensions are

$$(m \times n)(n \times r) = (m \times r)(r \times r)$$

- We always can find orthogonal basis for $Nul(A)$ and $Nul(A^T)$ and augment those vectors in matrices U and V, respectively to make them into square orthogonal matrices. We can fill zeroes in the rest of diagonal elements in matrix Σ. (row space is orthogonal to null space)
- Then the matrices dimension in the equation $AV = U\Sigma$ become:

$$(m \times n)(n \times n) = (m \times m)(m \times n)$$

Singular value decomposition

- After adding null space vectors the equation $AV = U\Sigma$ becomes:
- $A\begin{bmatrix} \boldsymbol{v}_1 & \dots & \boldsymbol{v}_m \end{bmatrix} = \begin{bmatrix} \boldsymbol{u}_1 & \dots & \boldsymbol{u}_n \end{bmatrix}\begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & 0 \end{bmatrix}$
- How to find U and V?
- $AV = U\Sigma \rightarrow A = U\Sigma V^{-1} = U\Sigma V^T$, since V is orthogonal.
- $AA^T = U\Sigma V^T (U\Sigma V^T)^T = U\Sigma V^T V\Sigma^T U^T = U\Sigma^2 U^T$, since $V^T V = I$ and $\Sigma\Sigma^T = \Sigma^2$
- $A^T A = (U\Sigma V^T)^T U\Sigma V^T = V\Sigma^T U^T U\Sigma V^T = V\Sigma^2 V^T$, since $U^T U = I$ and $\Sigma\Sigma^T = \Sigma^2$

Singular value decomposition

- $A = U\Sigma V^T$
- $AA^T = U\Sigma^2 U^T$
- $A^T A = V\Sigma^2 V^T$
- Recall that AA^T and $A^T A$ are symmetric matrices where their sizes are $m \times m$ and $n \times n$, respectively.
- Recall that symmetric matrices are diagonalizable and their eigenvector matrix are orthogonal.
- The eigenvalues of AA^T are the same as eigenvalues of $A^T A$ since the eigenvalues of AB are the same as eigenvalues of BA.
- In other words, Σ^2 are the eigenvalues and columns of V are the eigenvectors of $A^T A$, and $U$are the eigenvectors of AA^T

Singular value decomposition

- Example: Find matrix $A = \begin{bmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{bmatrix}$ to its equivalent singular vector decomposition.

  - Find $AA^T = UDU^T$,
  - $AA^T = \begin{bmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ 1 & 3 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 11 & 1 \\ 1 & 11 \end{bmatrix}$
  - $\begin{vmatrix} 11-\lambda & 1 \\ 1 & 11-\lambda \end{vmatrix} = 0, (11-\lambda)^2 - 1 = 0 \ \rightarrow (\lambda - 10)(\lambda - 12) = 0$
  - $\lambda_1 = 12, \lambda_2 = 10.$

Singular value decomposition

- Example Cont'd:

  - $\begin{bmatrix} 11-12 & 1 \\ 1 & 11-12 \end{bmatrix} \boldsymbol{u}_1 = \boldsymbol{0} \rightarrow \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \boldsymbol{u}_1 = \boldsymbol{0} \rightarrow \boldsymbol{u}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

  - $\begin{bmatrix} 11-10 & 1 \\ 1 & 11-10 \end{bmatrix} \boldsymbol{u}_2 = \boldsymbol{0} \rightarrow \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \boldsymbol{u}_2 = \boldsymbol{0} \rightarrow \boldsymbol{u}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

  - $U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}, D = \begin{bmatrix} 12 & 0 \\ 0 & 10 \end{bmatrix}$

Singular value decomposition

- Example Cont'd:
  - Find $A^T A = VDV^T$,
  - $AA^T = \begin{bmatrix} 3 & -1 \\ 1 & 3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 10 & 0 & -2 \\ 0 & 10 & 4 \\ -2 & 4 & 2 \end{bmatrix}$
  - $\begin{vmatrix} 10-\lambda & 0 & -2 \\ 0 & 10-\lambda & 4 \\ -2 & 4 & 2-\lambda \end{vmatrix} = 0, \rightarrow \lambda(\lambda-10)(\lambda-12) = 0$
  - $\lambda_1 = 12, \lambda_2 = 10, \lambda_3 = 0.$

Singular value decomposition

- Example Cont'd:
  - $\begin{bmatrix} 10-12 & 0 & -2 \\ 0 & 10-12 & 4 \\ -2 & 4 & 2-12 \end{bmatrix} \boldsymbol{v}_1 = \boldsymbol{0} \rightarrow \begin{bmatrix} -2 & 0 & -2 \\ 0 & -2 & 4 \\ -2 & 4 & -10 \end{bmatrix} \boldsymbol{v}_1 = \boldsymbol{0} \rightarrow \boldsymbol{v}_1 = \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}$
  - $\begin{bmatrix} 10-10 & 0 & -2 \\ 0 & 10-10 & 4 \\ -2 & 4 & 2-10 \end{bmatrix} \boldsymbol{v}_2 = \boldsymbol{0} \rightarrow \begin{bmatrix} 0 & 0 & -2 \\ 0 & 0 & 4 \\ -2 & 4 & -8 \end{bmatrix} \boldsymbol{v}_2 = \boldsymbol{0} \rightarrow \boldsymbol{v}_2 = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$
  - $\begin{bmatrix} 10-0 & 0 & -2 \\ 0 & 10-0 & 4 \\ -2 & 4 & 2-0 \end{bmatrix} \boldsymbol{v}_3 = \boldsymbol{0} \rightarrow \begin{bmatrix} 10 & 0 & -2 \\ 0 & 10 & 4 \\ -2 & 4 & 2 \end{bmatrix} \boldsymbol{v}_3 = \boldsymbol{0} \rightarrow \boldsymbol{v}_3 = \begin{bmatrix} 1 \\ -2 \\ 5 \end{bmatrix}$

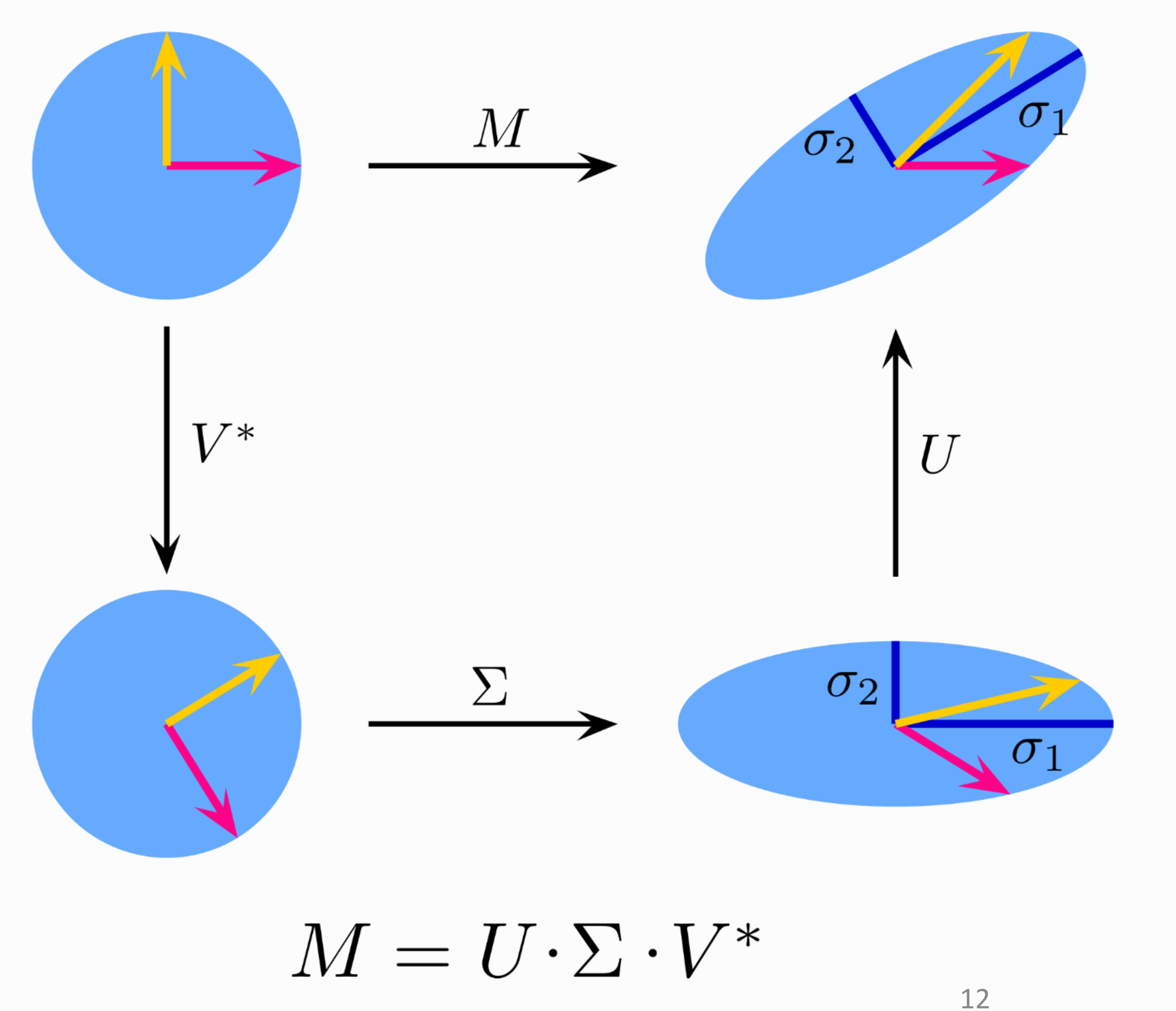
Singular value decomposition

- Example Cont'd:

  - $V = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{30}} \\ -\frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{30}} \\ -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{5}} & \frac{5}{\sqrt{30}} \end{bmatrix}, D = \begin{bmatrix} 12 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 0 \end{bmatrix}$
  - $A = U\Sigma V^T \rightarrow \begin{bmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \sqrt{12} & 0 & 0 \\ 0 & \sqrt{10} & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{30}} & -\frac{2}{\sqrt{30}} & \frac{5}{\sqrt{30}} \end{bmatrix}$

Geometric interpretation

- Here is a geometric interpretation of SVD for a 2×2 matrix M.
- V^T(in figure V^*) rotates unit vectors.
- Σ scales the vectors
- U perform the final rotation.



Singular value decomposition

- Note that we can assume SVD of matrix A as:

$$A = \sum_{i=1}^{r} \sigma_i \boldsymbol{u}_i {\boldsymbol{v}_i}^T$$

Where r is the rank of the matrix.

- Size of each $\sigma_i \boldsymbol{u}_i {\boldsymbol{v}_i}^T$ is $m \times n$. The greater σ_i the greater values added to reconstruct matrix A.

$$A = \sigma_1 \boldsymbol{u}_1 {\boldsymbol{v}_1}^T + \dots + \sigma_r \boldsymbol{u}_r {\boldsymbol{v}_r}^T$$

Solve Ax = 0

- What minimizes Ax. And why?
- As defined previous we can write A as

$$A = U\Sigma V^T = \begin{bmatrix} \boldsymbol{u}_1 & \dots & \boldsymbol{u}_n \end{bmatrix} \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{v}_1 \\ \vdots \\ \boldsymbol{v}_r \end{bmatrix}$$

Where σ_r is the smallest eigenvalue. Now we have

$$\mathrm{Ax} = \sum_{i=1}^{r} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i{}^T \,\mathrm{x}$$

If we choose $x = \boldsymbol{v}_r$. Then since all $v_i^T v_r = 0\ (i \neq r)$ we have

$$Ax = Av_r = \sigma_r \boldsymbol{u}_r \boldsymbol{v}_r{}^T v_r = \sigma_r \boldsymbol{u}_r$$

Smallest value of Ax associate with $x = \boldsymbol{v}_r$. And if $\sigma_r = 0$ then $\boldsymbol{v}_r$ is an answer of Ax=0

Pseudo inverse

- Assume $Ax = b$, Then by singular value decomposition of A we have

$$Ax = b \;\Rightarrow U\Sigma V^T x = b$$
$$\Sigma V^T x = U^T b$$
$$V^T x = \Sigma^* U^T b$$
$$x = (V\Sigma^* U^T) b$$

- $V\Sigma^* U^T$ called the pseudo inverse of A . It is useful for finding inverse of non-square matrices.

Applications

- One of the applications of SVD is dimensionality reduction.
- A $m \times n$ matrix can be thought of gray level of a digital image.
- If the image A is decomposed to its singular values and vectors, we can pick only the most significant $\boldsymbol{u}_i$'s, σ_i's and $\boldsymbol{v}_i$'s.
- By doing this we can compress the information of the image.
- Suppose the image A is $m \times n$.
- $\hat{A} = \sum_{i=1}^{K} \sigma_i \boldsymbol{u}_i {\boldsymbol{v}_i}^T$
- In next slide you will see the original image and its compressed up to K most significant singular values.

Image compression using SVD

Original

K=8

K=32

K=128

K=512