Singular Values and Singular Vectors

Linear Algebra

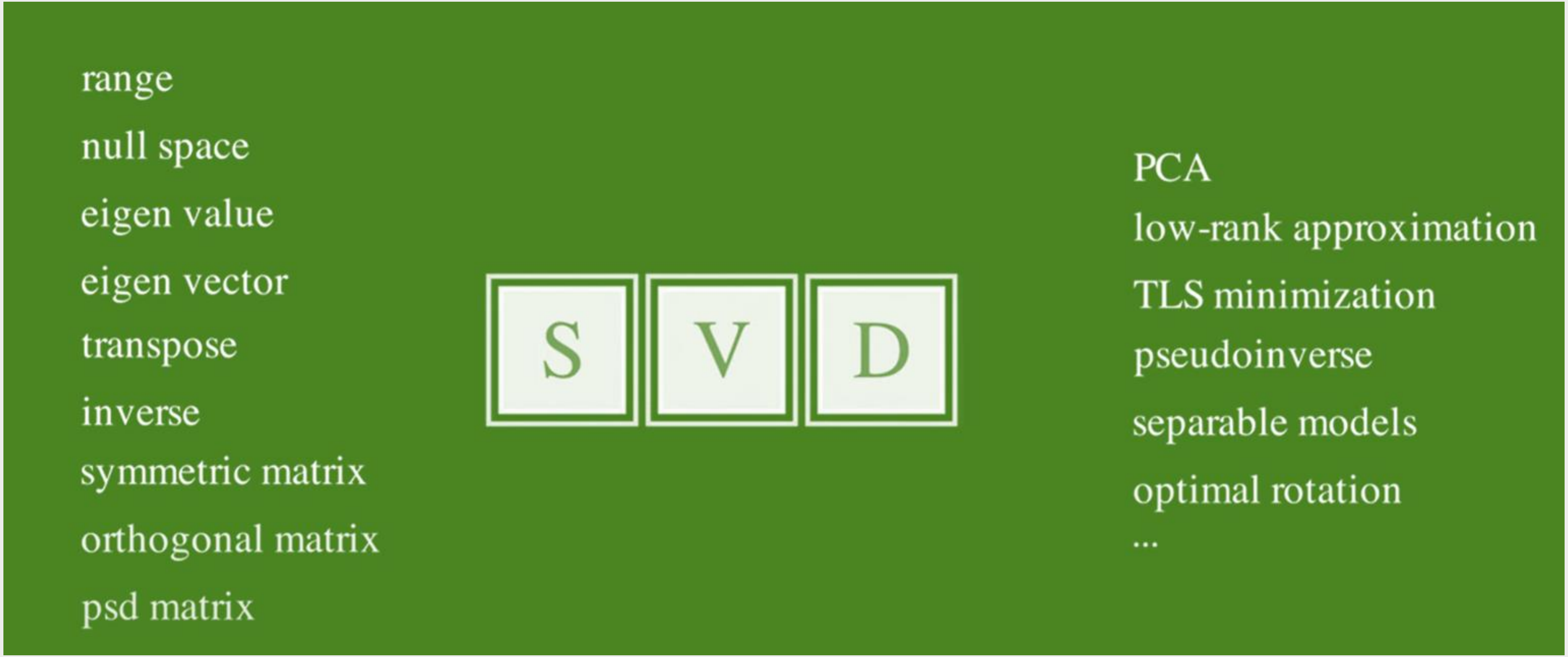
Department of Computer Engineering

Sharif University of Technology

Hamid R. Rabiee rabiee@sharif.edu

Maryam Ramezani maryam.ramezani@sharif.edu

Introduction


range
null space
eigen value
eigen vector
transpose
inverse
symmetric matrix
orthogonal matrix
psd matrix
S V D
PCA
low-rank approximation
TLS minimization
pseudoinverse
separable models
optimal rotation
...

Introduction

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Singular Value

Singular value and eigenvalue

- $S_{m\times n}$ Non−Square!!
- $\sigma_i = \sqrt{\lambda_i} \quad \lambda_i \in \sigma(S^T S), i = 1, \dots, n$
- $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{m-1} \geq \sigma_m$

Example

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

$$S^T S = \begin{bmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{bmatrix} \Rightarrow \lambda(S^T S) = \{360, 90, 0\}$$

$$\Rightarrow \begin{cases} \sigma_1 = \sqrt{360} = 6\sqrt{10} \\ \sigma_2 = \sqrt{90} = 3\sqrt{10} \\ \sigma_3 = 0 \end{cases}$$

Singular value and eigenvector

Theorem

$\{v_1, \dots, v_n\}$ are orthonormal eigenvectors of matrix $S^T S$ then singular values of matrix S are norm of Sv_i vectors:

$$||Sv_i|| = \sigma_i$$

Proof?

Singular value and eigenvector

Example

$$S = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} \rightarrow S^T S = \begin{bmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{bmatrix} \rightarrow \sigma_1 = \sqrt{360}, \sigma_2 = \sqrt{90}, \sigma_3 = 0$$

$$v_1 = \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix}, v_2 = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}, v_3 = \begin{bmatrix} 2/3 \\ -2/3 \\ 1/3 \end{bmatrix}$$

$$Sv_1 = \begin{bmatrix} 18 \\ 6 \end{bmatrix} \Rightarrow \|Sv_1\| = \sqrt{18^2 + 6^2} = \sigma_1$$

$$Sv_2 = \begin{bmatrix} 3 \\ -9 \end{bmatrix} \Rightarrow \|Sv_2\| = \sqrt{3^2 + (-9)^2} = \sigma_2$$

$$Sv_3 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \|Sv_3\| = 0 = \sigma_3$$

Singular value and Rank

Theorem

$\{v_1, \dots, v_n\}$ are orthonormal eigenvectors of matrix $S^T S$ and S has r non-zero singular value:

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0, \qquad \sigma_{r+1} = \cdots = \sigma_n = 0$$

$\{Sv_1, \dots, Sv_r\}$ is a orthogonal basis for range of S

rank(S)=r

Rank of Matrix = Number of nonzero singular values

How to find $\{u_1, \dots, u_r\}$ is a orthonormal basis for range of S

SVD

Singular Value Decomposition (SVD)

- Given any $m \times n$ matrix **A**, algorithm to find matrices **U**, **V**, and $\sum$ such that (**always exists**)
- $A = U\Sigma V^T$ $A = U\Sigma V^*$

  U is $m \times m$ and orthogonal (always real)

  $\sum$ is $m \times n$ and diagonal with non-negative (always real) called singular values.

  V is $n \times n$ and orthogonal (always real)
- Columns of U are eigenvectors of AA^T (called the left singular vectors).
- Columns of V are eigenvectors of A^TA (called the right singular vectors).
- The non-zero singular vectors are the positive square roots of non-zero eigenvalues of AA^T or A^TA.

SVD Introduction

- Generalization of the spectral decomposition that applies to all matrices, rather than just normal matrices.
- Applications:
  - Compute the size of a matrix (in a way that typically makes more sense than norm)
  - Provide a new geometric interpretation of linear transformations
  - Solve optimization problems
  - Construct an “almost inverse” for matrices that do not have an inverse.

SVD Comparison

SVD	Diagonalization	Spectral Decomposition
applies to every single matrix (even rectangular ones).	only applies to matrices with a basis of eigenvectors	only applies to normal matrices
matrix $\sum$ in the middle of the SVD is diagonal (and even has real non-negative entries)	do not guarantee an entrywise non-negative matrix	do not guarantee an entrywise non-negative matrix
It requires two unitary matrices U and V	only required one invertible matrix	only required one unitary matrix

- The Σ_i are called the singular values of **A**
- If **A** is singular, some of the Σ_i will be 0
- In general *rank*(**A**) = number of nonzero Σ_i
- SVD is mostly unique (up to permutation of singular values, or if some Σ_i are equal)

SVD for Square Matrix

- The SVD is a factorization of a m x n matrix into

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

Where U is a m x m orthogonal matrix, V^T is a n x n orthogonal matrix and Σ is a m x n diagonal matrix.

For a square matrix (m=n):

$$A = \begin{pmatrix} \vdots & \cdots & \vdots \\ u_1 & \cdots & u_n \\ \vdots & \cdots & \vdots \end{pmatrix} \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{pmatrix} \begin{pmatrix} \cdots & v_1^T & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & v_n^T & \cdots \end{pmatrix}$$

$$A = \begin{pmatrix} \vdots & \cdots & \vdots \\ u_1 & \cdots & u_n \\ \vdots & \cdots & \vdots \end{pmatrix} \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{pmatrix} \begin{pmatrix} \vdots & \cdots & \vdots \\ v_1 & \cdots & v_n \\ \vdots & \cdots & \vdots \end{pmatrix}^T$$

Reduced SVD

$$[Sv_1 \quad \dots \quad Sv_r \quad 0 \quad \dots \quad 0]_{m\times n} = [\sigma_1 u_1 \quad \dots \quad \sigma_r u_r \quad 0 \quad \dots \quad 0]_{m\times n}$$
$$[Sv_1 \quad \dots \quad Sv_r \quad Sv_{r+1} \quad \dots \quad Sv_n]_{m\times n} = [\sigma_1 u_1 \quad \dots \quad \sigma_r u_r \quad 0 \quad \dots \quad 0]_{m\times n}$$

$$S[v_1 \quad \dots \quad v_n] = [u_1 \quad \dots \quad u_m]\left[\begin{array}{ccc|c} \sigma_1 & \cdots & 0 & \\ \vdots & & \vdots & 0 \\ 0 & \cdots & \sigma_r & \\ \hline & 0 & & 0 \end{array}\right]$$

$$S_{m\times n}V_{n\times n} = U_{m\times m}\Sigma_{m\times n}$$

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

Reduced SVD

- what happens when A is not a square matrix?
- n > m

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

$$= \begin{pmatrix} \vdots & \cdots & \vdots \\ u_1 & \cdots & u_m \\ \vdots & \cdots & \vdots \end{pmatrix}_{m\times m} \begin{pmatrix} \sigma_1 & & & 0 & \\ & \ddots & & & \ddots & \\ & & \sigma_m & & & 0 \end{pmatrix}_{m\times n} \begin{pmatrix} \cdots & v_1^T & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & v_m^T & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & v_n^T & \cdots \end{pmatrix}_{n\times n}$$

We can instead rewrite the above as:

$$A = U\Sigma_R V_R^T$$

where V_R is n x m matrix and Σ_R is a m x m matrix

In general:

$$A = U_R \Sigma_R V_R^T$$

U_R is a m x k matrix
Σ_R is a k x k matrix
V_R is a n x k matrix

k = min(m, n)

Now U and V are not orthogonal.
But their columns are orthonormal.

Reduced SVD

- m > n

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

$$= \begin{pmatrix} \vdots & \cdots & \vdots & \cdots & \vdots \\ u_1 & \cdots & u_n & \cdots & u_m \\ \vdots & \cdots & \vdots & \cdots & \vdots \end{pmatrix}_{m\times m} \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}_{m\times n} \begin{pmatrix} \cdots & v_1^T & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & v_n^T & \cdots \end{pmatrix}_{n\times n}$$

We can instead rewrite the above as:

$$A = U\Sigma_R V_R^T$$

Now U and V are not orthogonal.
But their columns are orthonormal.

where U_R is m x n matrix and Σ_R is a n x n matrix

Reduced SVD

- Let's take a look at the product of $\Sigma^T\Sigma$ where Σ has the singular values of a A, a m x n matrix.

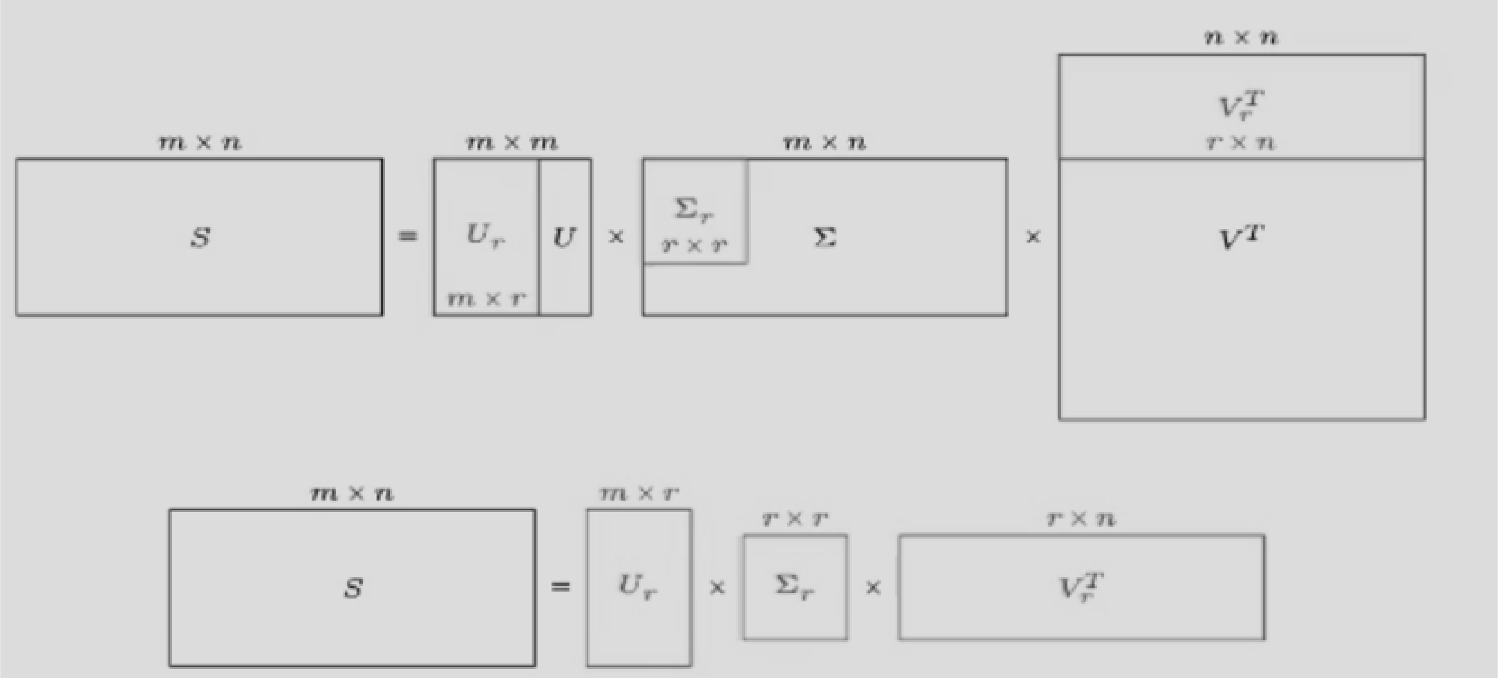
  - **m > n:**

$$\Sigma^T\Sigma = \begin{pmatrix} \sigma_1 & & & 0 & & \\ & \ddots & & & \ddots & \\ & & \sigma_n & & & 0 \end{pmatrix}_{n\times m} \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}_{m\times n} = \begin{pmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 \end{pmatrix}_{n\times n}$$

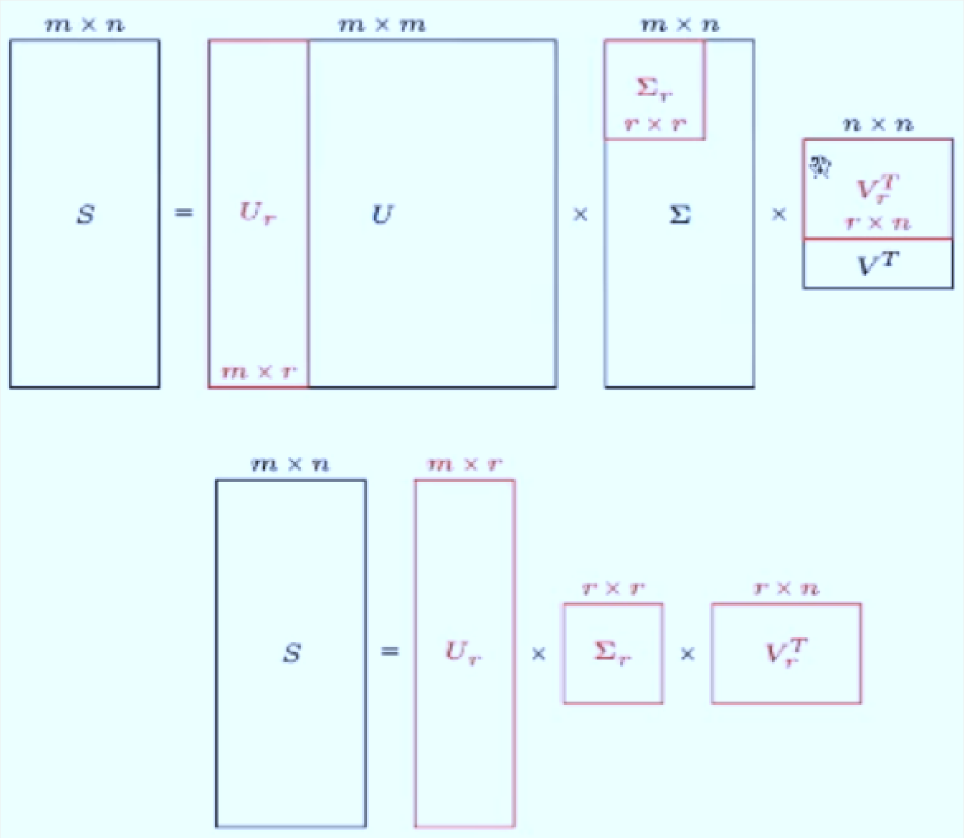
  - **n > m:**

$$\Sigma^T\Sigma = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_m \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}_{n\times m} \begin{pmatrix} \sigma_1 & & & 0 & & \\ & \ddots & & & \ddots & \\ & & \sigma_m & & & 0 \end{pmatrix}_{m\times n} = \begin{pmatrix} \sigma_1^2 & & & 0 & & \\ & \ddots & & & \ddots & \\ & & \sigma_m^2 & & & 0 \\ 0 & & & 0 & & \\ & \ddots & & & \ddots & \\ & & 0 & & & 0 \end{pmatrix}_{n\times n}$$

- Wide Matrix


$m \times n$
S
$=$
$m \times m$
U_r
$m \times r$
U
$\times$
$m \times n$
Σ_r
$r \times r$
Σ
$\times$
$n \times n$
V_r^T
$r \times n$
V^T
$m \times n$
S
$=$
$m \times r$
U_r
$\times$
$r \times r$
Σ_r
$\times$
$r \times n$
V_r^T

- Tall Matrix



How can we compute an SVD of a matrix A?

- Assume A with singular value decomposition $A = U\Sigma V^T$. Let's take a look at the eigenpairs corresponding to $A^T A$:

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

Hence $A^T A = V\Sigma^2 V^T$

- Recall that columns of V are all linear independent (orthogonal matrix), then from diagonalization ($B = XDX^{-1}$), we get:
  - The columns of V are the eigenvectors of the matrix $A^T A$
  - The diagonal entries of Σ^2 are the eigenvalues of $A^T A$
- Let's call λ the eigenvalues of $A^T A$, then $\sigma_i^2 = \lambda_i$

How can we compute an SVD of a matrix A?

- In a similar way,

$$AA^T = (U\Sigma V^T)(U\Sigma V^T)^T$$
$$(U\Sigma V^T)(V^T)^T(\Sigma)^T U^T = U\Sigma V^T V\Sigma^T U^T = U\Sigma\Sigma^T U^T$$

Hence $AA^T = U\Sigma^2 U^T$

- Recall that columns of U are all linear independent (orthogonal matrix), then from diagonalization ($B = XDX^{-1}$), we get:
  - The columns of U are the eigenvectors of the matrix AA^T

How can we compute an SVD of a matrix A?

1. Evaluate the n eigenvectors v_i and eigenvalues λ_i of $A^T A$
2. Make a matrix V from the normalized vectors v_i. The columns are called "right singular vectors".

$$\mathrm{V} = \begin{pmatrix} \vdots & \cdots & \vdots \\ v_1 & \cdots & v_n \\ \vdots & \cdots & \vdots \end{pmatrix}$$

3. Make a diagonal matrix from the square roots of the eigenvalues.

$$\Sigma = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{pmatrix} \quad \sigma_i = \sqrt{\lambda_i} \quad \text{and } \sigma_1 \geq \sigma_2 \geq \cdots$$

4. Find U: $A = U\Sigma V^T \Rightarrow U\Sigma = AV \Rightarrow U = AV\Sigma^{-1}$. The columns are called "left singular values".

Example

$$S = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} \to S^T S = \begin{bmatrix} 9 & -9 \\ -9 & 9 \end{bmatrix}, rank(S) = 1$$

$$\Delta(\lambda) = \lambda^2 - 18\lambda = 0 \Rightarrow \sigma_1 = \sqrt{18}, \sigma_2 = 0 \Rightarrow \boldsymbol{\Sigma} = \begin{bmatrix} \mathbf{3\sqrt{2}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

$$v_1 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}, v_2 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \Rightarrow \boldsymbol{V} = \frac{\mathbf{1}}{\sqrt{\mathbf{2}}} \begin{bmatrix} \mathbf{1} & \mathbf{1} \\ \mathbf{-1} & \mathbf{1} \end{bmatrix}$$

$$Sv_1 = \begin{bmatrix} 2/\sqrt{2} \\ -4/\sqrt{2} \\ 4/\sqrt{2} \end{bmatrix} \Rightarrow u_1 = \frac{1}{\sigma_1} Sv_1 = \frac{1}{3\sqrt{2}} \begin{bmatrix} 2/\sqrt{2} \\ -4/\sqrt{2} \\ 4/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix}$$

$$u_2 = \begin{bmatrix} 2/3 \\ 2/3 \\ 1/3 \end{bmatrix}, u_3 = \begin{bmatrix} -2/3 \\ 1/3 \\ 2/3 \end{bmatrix} \Rightarrow \boldsymbol{U} = \frac{\mathbf{1}}{\mathbf{3}} \begin{bmatrix} \mathbf{1} & \mathbf{2} & \mathbf{-2} \\ \mathbf{-2} & \mathbf{2} & \mathbf{1} \\ \mathbf{2} & \mathbf{1} & \mathbf{2} \end{bmatrix}$$

$$S = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 & 2 & -2 \\ -2 & 2 & 1 \\ 2 & 1 & 2 \end{bmatrix} \begin{bmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = U\Sigma V^T$$

Unitary Freedom of PSD Decompositions

Suppose $B, C \in \mathcal{M}_{m,n}(\mathbb{F})$. The following are equivalent:

a. There exists a unitary matrix $U \in \mathcal{M}_m(\mathbb{F})$ such that $C = UB$,

b. $B^*B = C^*C$,

c. $(B\mathbf{v}).(B\mathbf{w}) = (C\mathbf{v}).(C\mathbf{w})$ for all $\mathbf{v}, \mathbf{w} \in \mathbb{F}^n$, and

d. $||B\mathbf{v}|| = ||C\mathbf{v}||$ for all $\mathbf{v} \in \mathbb{F}^n$.

Example

$$\begin{bmatrix} 3 & 2 \\ -2 & 0 \end{bmatrix}$$

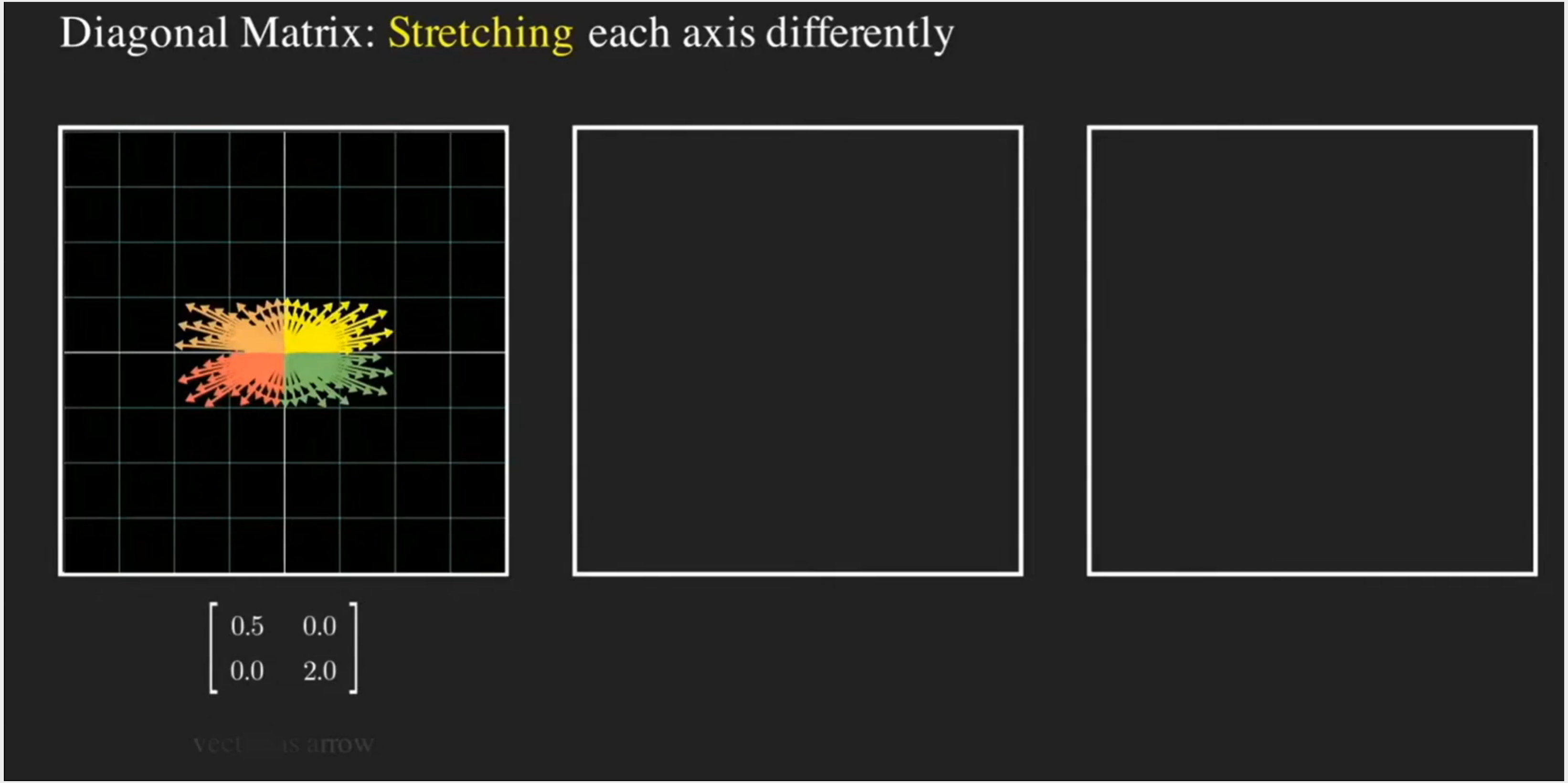
$$\begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 1 \\ 3 & 2 & 1 \end{bmatrix}$$

SVD Proof

- If $m \neq n$ then A^*A, AA^* have different sizes, but they still have essentially the same eigenvalues—whichever one is larger just has some extra 0 eigenvalues.
- The same is actually true of AB and BA for any A and B.

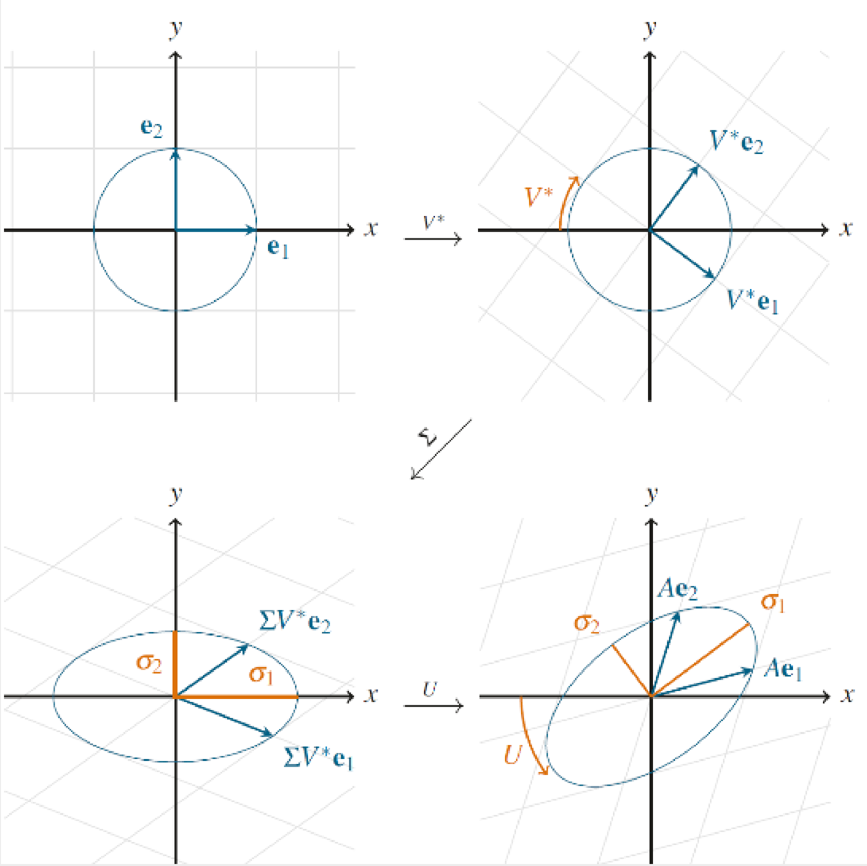
- Proof SVD in another view!!

Review


Diagonal Matrix: Stretching each axis differently
0.5 0.0
0.0 2.0

$A = U\Sigma V^*$

The product of a matrix's singular values equals the absolute value of its determinant



Determining the rank of a matrix

- Suppose A is a m x n rectangular matrix where m > n:

$$A = \begin{pmatrix} \vdots & \cdots & \vdots & \cdots & \vdots \\ u_1 & \cdots & u_n & \cdots & u_m \\ \vdots & \cdots & \vdots & \cdots & \vdots \end{pmatrix}_{m\times m} \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}_{m\times n} \begin{pmatrix} \cdots & v_1^T & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & v_n^T & \cdots \end{pmatrix}_{n\times n}$$

$$A = \begin{pmatrix} \vdots & \cdots & \vdots \\ u_1 & \cdots & u_n \\ \vdots & \cdots & \vdots \end{pmatrix} \begin{pmatrix} \cdots & \sigma_1 v_1^T & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & \sigma_n v_n^T & \cdots \end{pmatrix} = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_n u_n v_n^T$$

$$A = \sum_{i=1}^{n} \sigma_i u_i v_i^T$$

$A_1 = \sigma_1 u_1 v_1^T$ what is rank$(A_1) = ?$

In general, rank$(A_k) = k$

Conclusion

- Let $A \in \mathcal{M}_{m,n}$ be a matrix with rank(A) = r and the singular value decomposition $A = U\Sigma V^T$, where

$$U = [u_1 \mid u_2 \mid \ \ldots \mid u_m] \text{ and } V = [v_1 \mid v_2 \mid \ \ldots \mid v_n]$$

Then

a. $\{u_1, u_2, \ldots, u_r\}$ is an orthonormal basis of range(A),

b. $\{u_{r+1}, u_{r+2}, \ldots, u_m\}$ is an orthonormal basis of null(A^*),

c. $\{v_1, v_2, \ldots, v_r\}$ is an orthonormal basis of range(A^*), and

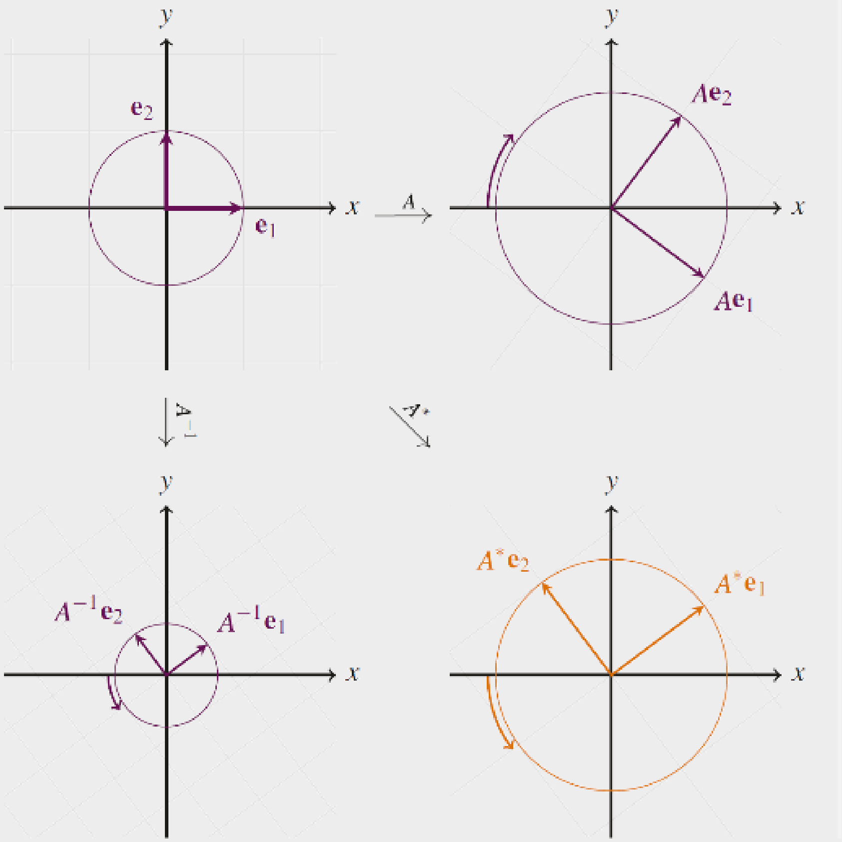
d. $\{v_{r+1}, v_{r+2}, \ldots, v_n\}$ is an orthonormal basis of null(A)

A Geometric Interpretation

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

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

$$A^{-1} = V\Sigma^{-1}U^*$$



Applications

Orthogonal Rank-One Sum Decomposition

- Suppose $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$, and $A \in \mathcal{M}_{m,n}(\mathbb{F})$ has $\text{rank}(A) = r$. There exist orthonormal sets of vectors $\{u_j\}_{j=1}^{r} \subset \mathbb{F}^m$ and $\{v_j\}_{j=1}^{r} \subset \mathbb{F}^n$ such that

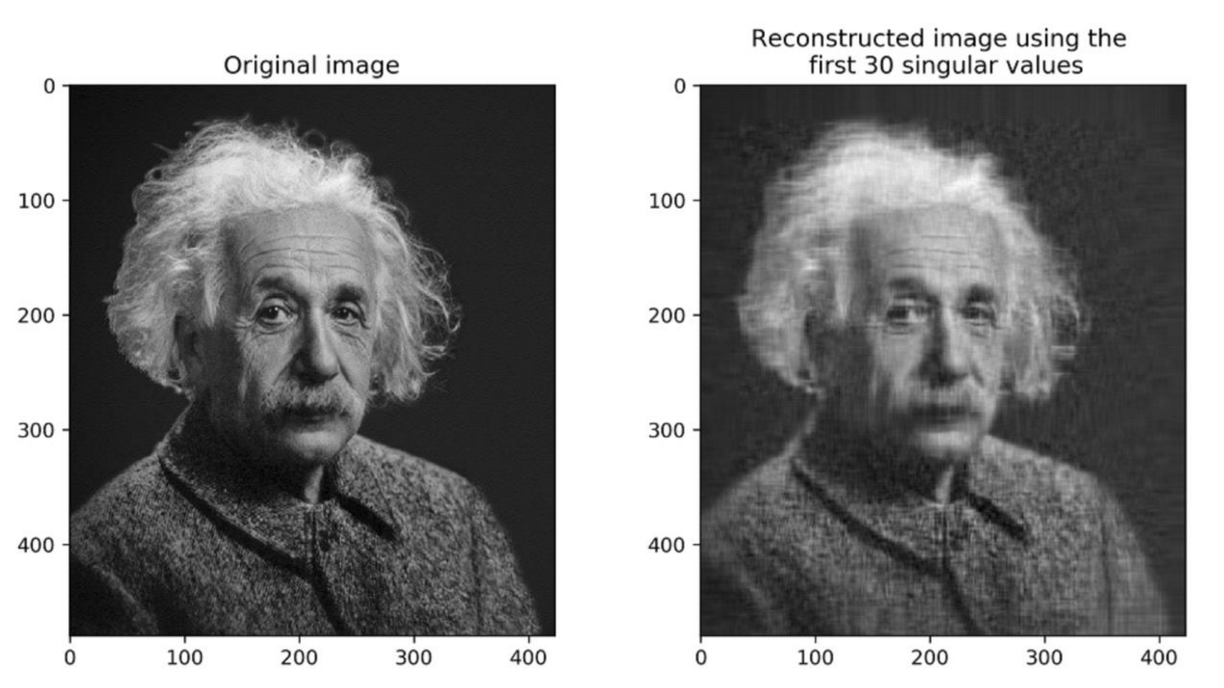
$$A = \sum_{i=1}^{r} \sigma_i u_i v_i^*,$$

where $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ are the non-zero singular values of A.

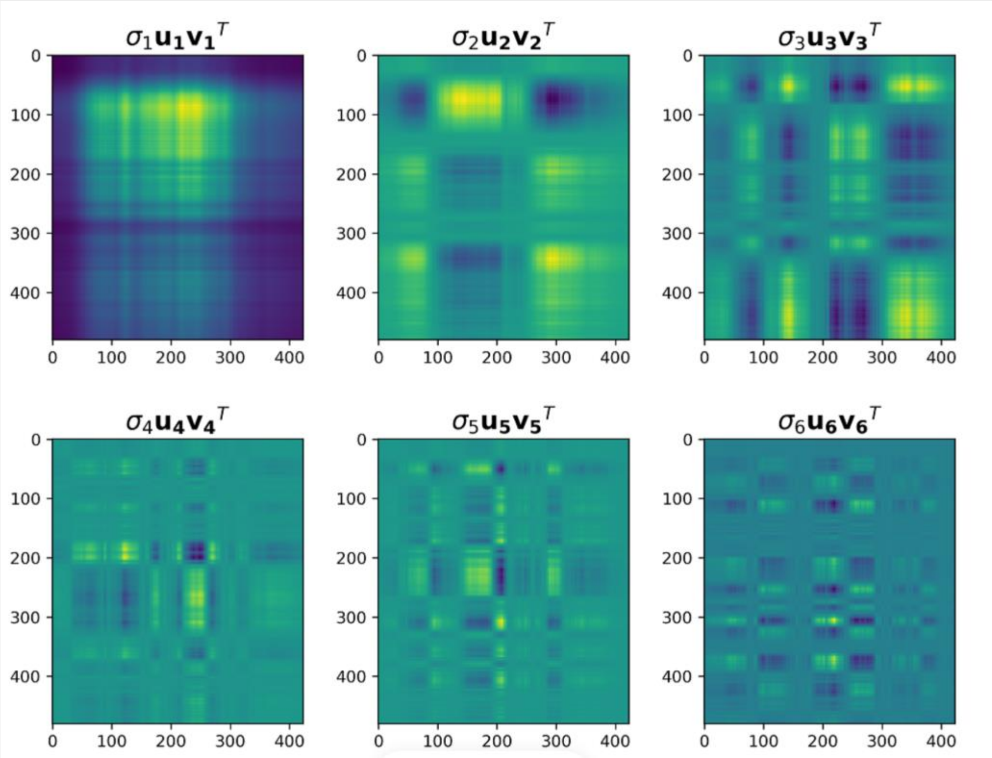
SVD and Matrix Similarity

- Suppose you want to find best rank−k approximation to **A**
  - Answer: set all but the largest k singular values to zero
- Can form compact representation by eliminating columns of **U** and **V** corresponding to zeroed Σ_i

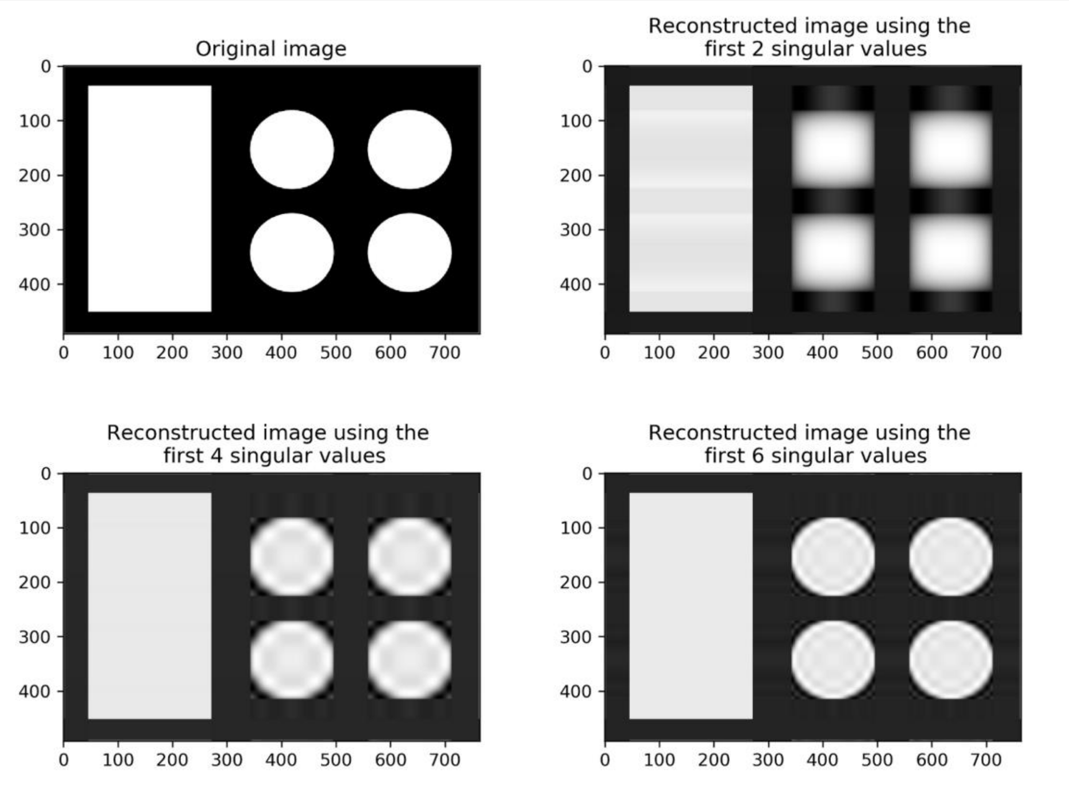
Application: Dimensionality Reduction



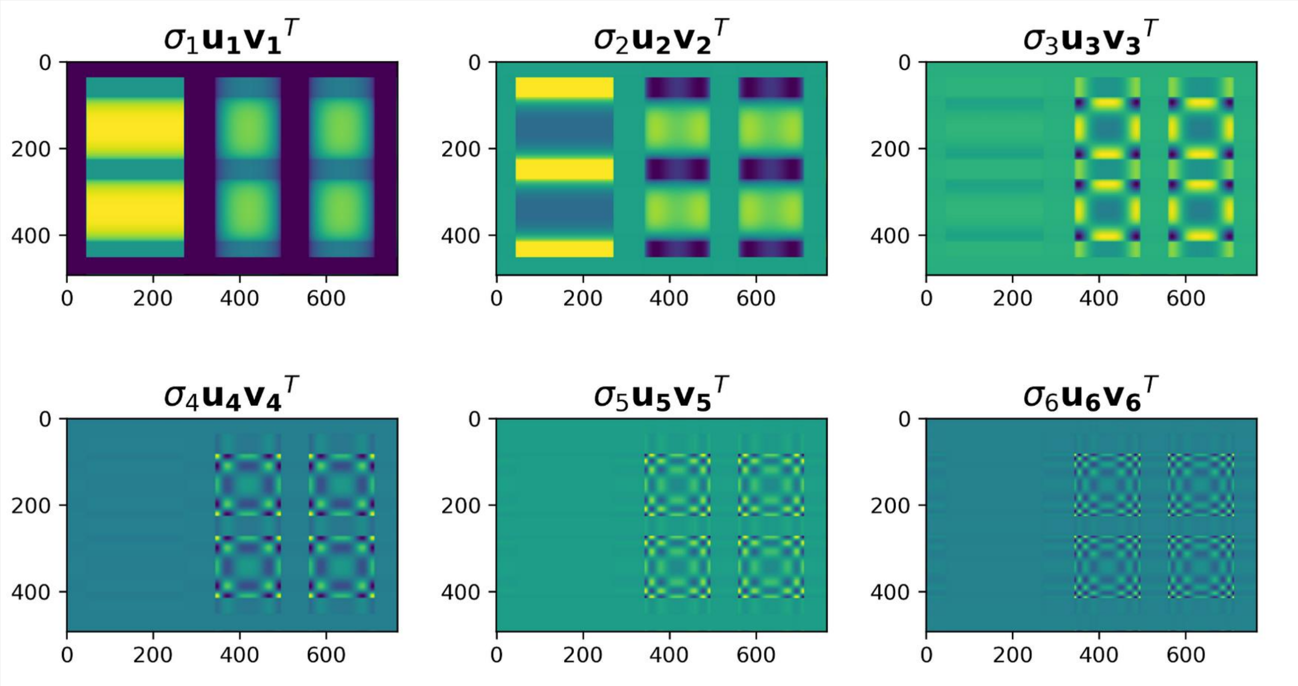
Application: Dimensionality Reduction



Application: Dimensionality Reduction



Application: Dimensionality Reduction



Image

SVD and PSD

- If $A \in \mathcal{M}_n$ is positive semidefinite then its singular values equals its eigenvalues.

SVD and Inverses

- Why is SVD so useful?
- $A^{-1} = V\Sigma^{-1}U^{-1} = V\Sigma^{-1}U^T$
  - Using fact that inverse = transpose for orthogonal matrices
  - Since Σ is diagonal, Σ^{-1} also diagonal with reciprocals of entries of Σ
- This fails when some Σ_i are 0
  - It's *supposed* to fail – singular matrix
- Pseudoinverse: if $\Sigma_i = 0$, set $\frac{1}{\Sigma_i}$ to 0 (!)
  - "Closest" matrix to inverse
  - Defined for all (even non-square, singular, etc.) matrices
  - Equal to $(A^TA)^{-1}A^T$ if A^TA invertible

Pseudo Inverse

- Problem:
  if A is rank-deficient, Σ is not invertible.
- How to fix it:
  Define the Pseudo Inverse
- Pseudo Inverse of a diagonal matrix:

$$(\Sigma^+)_i = \begin{cases} \dfrac{1}{\sigma_i}, & if\ \sigma_i \neq 0 \\ 0, & if\, \sigma_i = 0 \end{cases}$$

- Pseudo Inverse of a matrix A:

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

Pseudo Inverse

- If a matrix A has the singular value decomposition

$$A = UWV^T$$

then the pseudo-inverse or Moore-Penrose inverse of A is

$$A^+ = V^T W^{-1} U$$

- If A is 'tall' (m > n) and has full rank

$$A^+ = (A^T A)^{-1} A^T$$

(it gives the least-squares solution $x_{lsq} = A^+ b$)

- If A is 'short' (n > m) and has full rank

$$A^+ = A^T (AA^T)^{-1}$$

(it gives the least-norm solution $x_{l-n} = A^+ b$)

- In general, $x_{pinv} = A^+ b$ is the minimum-norm, least-square solution.

SVD and Norm

- One common definition for the norm of a matrix is the Frobenius norm:

$$||A||_F^2 = \sum_{i=1:m} \sum_{j=1:n} a_{ij}^2$$

- Frobenius norm can be computed from SVD

$$||A||_{\mathrm{F}}^2 = \sum_{i=1:p} \Sigma_i{}^2 \quad where\, p = \min(n, m)$$

- So changes to a matrix can be evaluated by looking at changes to singular values

SVD and Norm

- 2-norm:

$$\|A\|_2 = \sup_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \sigma_1$$

SVD and Least Square

- $x^* = A^{-1}b = (UDV^T)^{-1}b,$

$$(UDV^T)^{-1} = V^{-T} D^{-1} U^{-1}$$

Moore–Penrose pseudoinverse

$$x^* = A^{-1}b = \boxed{VD^{-1}U^T}b$$

- Invert the diagonal entries in D that are nonzero, but leave the other diagonal entries alone as zeros.