

Lecture 14 - A

§8.2 Singular Value Decomposition

T. Takahashi

Department of Mathematics and Statistics
University of New Mexico



Goals

We will answer these questions:

- What are singular values and why are they more versatile than eigenvalues?
- How do you find the singular value decomposition?
- What is an overdetermined system and when does it arise in applications?
- What is the standard way to solve an overdetermined system?
- How can you use the SVD to solve overdetermined systems and why would you want to?



Linear Algebra Review



Matrix Multiplication $C = AB$

Column Views

The product $C = AB$ can be viewed by columns

Column View

Each column of C is the product of A with the corresponding column of B :

$$\text{Col}_j(C) = A \cdot \text{Col}_j(B)$$

$$\begin{bmatrix} | & & | \\ c_1 & \dots & c_n \\ | & & | \end{bmatrix} = A \begin{bmatrix} | & & | \\ b_1 & \dots & b_n \\ | & & | \end{bmatrix}$$
$$c_j = Ab_j$$

This will be useful later.



Multiply by e_i

- Let I be an $n \times n$ identity matrix

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

- We define e_i to be the i^{th} column of I
- Consider the multiplication Ae_1

$$Ae_1 = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ d \\ g \end{bmatrix}$$

General Formula

$$Ae_i = a_i,$$

where a_i is the i^{th} column of A

Definition: Eigenvalues and Eigenvectors

For a matrix A , the scalar-vector pairs (λ, v) such that

$$Av = \lambda v, \quad v \neq 0$$

are **eigenvalue-eigenvectors**.

- To find the eigenvalues, we find the roots of the **characteristic polynomial**

$$p(\lambda) := \det(\lambda I - A)$$

- We can only find eigenvalues for square matrices



Orthogonal Matrix

Definition: Orthogonality

A square matrix U is **orthogonal** if $U^T U = U U^T = I$.

Some Properties

- Each row and column of U has norm 1
- The columns of U are an orthonormal set of vectors
- The rows of U are an orthonormal set of vectors
- $U^{-1} = U^T$
- They are norm preserving: $\|Uv\| = \|v\|$

This will be useful later.



Orthogonal Matrix

Example

Consider the matrix

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

- The inverse is

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

- If we let $x = \begin{bmatrix} 3 & 4 \end{bmatrix}^T$, then

$$\|Ax\| = \left\| \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} \right\| = \left\| \begin{bmatrix} -4 \\ 3 \end{bmatrix} \right\| = \|u\|$$

- The matrix A is orthogonal



Limitations of Eigenvalues

- Image files are actually matrices



- Eigenvalues of images used to be prominent in face recognition
- Similar eigenvalues, eigenvectors $\implies$ similar faces

Limitations of Eigenvalues

Question

What if the image was not $n \times n$? Does that mean facial recognition is impossible?

- Finding eigenvalues is only possible for $n \times n$ matrices
- If we can't find eigenvalues, facial recognition is not possible



Limitations of Eigenvalues

Question

What if the image was not $n \times n$? Does that mean facial recognition is impossible?

- Finding eigenvalues is only possible for $n \times n$ matrices
- If we can't find eigenvalues, facial recognition is not possible
- However... AA^T and $A^T A$ are both always square matrices.

Key Idea

Even if an image is rectangular, we can still find the eigenvalues of AA^T which is always a square.



Singular Value Decomposition



Symmetric Matrices

- Consider the rectangular matrix A

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$$

- The matrix AA^T is square

$$AA^T = \begin{bmatrix} 5 & 11 & 17 \\ 11 & 25 & 39 \\ 17 & 39 & 61 \end{bmatrix}$$

- The matrix A^TA is also square

$$A^TA = \begin{bmatrix} 35 & 44 \\ 44 & 56 \end{bmatrix}$$

Key Idea

Even if A doesn't have eigenvalues, AA^T and A^TA do have eigenvalues.

Singular Values

Definition: Singular Values

The **singular values** of an $m \times n$ matrix A are

$$\sigma_1, \sigma_2, \dots, \sigma_r,$$

where $r = \min(m, n)$ and the σ 's are square roots of the eigenvalues of $A^T A$ or AA^T .

How to Find Singular Values of A

- 1 Calculate AA^T or $A^T A$ (whichever is smaller).
- 2 Find the eigenvalues of that matrix.
- 3 Take the square root of those eigenvalues.



Singular Value Example

Example

Find the singular values of

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 0 & 1 \end{bmatrix}$$

- ① Calculate $B = A^T A$ since its 2×2

$$B = A^T A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$$

- ② Find eigenvalues of B

$$\det(\lambda B - I) = \begin{vmatrix} \lambda - 2 & -2 \\ -2 & \lambda - 2 \end{vmatrix} = \lambda^2 - 4\lambda + 4 - 4 = \lambda(\lambda - 4) = 0 \implies \lambda = 0, 4$$

- ③ Take the square root of the eigenvalues

$$\sigma_1 = \sqrt{4} = 2, \quad \sigma_2 = \sqrt{0} = 0$$



Singular Value Decomposition (SVD)

Definition: SVD

SVD is the factorization of $A \in \mathbb{R}^{m \times n}$ in the form $A = USV^T$, where U ($m \times m$) and V ($n \times n$) are orthogonal and S ($m \times n$) is a diagonal matrix containing the **singular values** of A

$$\underbrace{\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r}_{\text{unspecified}} > \underbrace{\sigma_{r+1} = \dots = \sigma_p = 0}_{\text{zero}}$$

where $r = \text{rank}(A)$ and $p = \min(m, n)$.

$$A = \begin{bmatrix} | & & | \\ u_1 & \dots & u_m \\ | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & \ddots \\ & & & & 0 \end{bmatrix} \begin{bmatrix} - & v_1^T & - \\ & \vdots & \\ - & v_n^T & - \end{bmatrix}$$



Visualizing the Shapes

SVD is possible for any size

- If A is a 4×2 matrix

$$A = USV^T = \begin{bmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ u_{21} & u_{22} & u_{23} & u_{24} \\ u_{31} & u_{32} & u_{33} & u_{34} \\ u_{41} & u_{42} & u_{43} & u_{44} \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix}^T$$

- If A is a 2×3 matrix

$$A = USV^T = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \end{bmatrix} \begin{bmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{bmatrix}^T$$

- A has $p = 2$ singular values σ_1 and σ_2
- If A is rank 1, then $\sigma_2 = 0$



Applications of Singular Values

- 1 Search Technology: find closely related documents or images
- 2 Clustering: aggregate documents or images into similar groups
- 3 Compression: efficient image storage
- 4 Principal axis: find the main axis of a solid (engineering/graphics)
- 5 Key words: identify best search tags for text
- 6 Graphs: partition graphs into subgraphs (graphics, analysis)
- 7 **Least squares**



Calculating the SVD



Calculating the SVD

Goal

For a matrix $A \in \mathbb{R}^{m \times n}$, calculate the decomposition

$$A = USV^T$$

where

- 1 S is diagonal and contains the singular values of A
- 2 U is orthogonal
- 3 V is orthogonal

Let's derive a formula for each matrix



Calculating S

We know that S contains the singular values on the diagonal

Steps

- 1 Calculate $B = A^T A$ or $B = A A^T$ (choose the smaller one)
- 2 Find all eigenvalues of B by solving

$$\det(\lambda I - B) = 0$$

- 3 Calculate the square root of the eigenvalues

$$\sigma = \sqrt{\lambda}$$

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



What is U ?

Part 1

- 1 Begin with the SVD of A :

$$A = USV^T$$

- 2 Next, calculate AA^T :

$$\begin{aligned} AA^T &= USV^T(USV^T)^T \\ &= USV^T V S^T U^T \\ &= US(V^T V) S^T U^T \\ &= US S^T U^T \end{aligned}$$

- 3 Therefore, (letting $S^2 := SS^T$) we have

$$AA^T U = US^2$$



What is U ?

Part 2

- ④ Consider the RHS of $AA^T U = US^2$

$$\begin{aligned} US^2 &= U \begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_m^2 \end{bmatrix} \\ &= U [e_1 \sigma_1 | e_2 \sigma_2 | \cdots | e_m \sigma_m] \\ &= [u_1 \sigma_1 | u_2 \sigma_2 | \cdots | u_m \sigma_m] \end{aligned}$$

- ⑤ The column view gives of $AA^T U = US^2$ is

$$AA^T u_i = \sigma_i^2 u_i$$



What is U ?

Part 3

Question

What is the significance of this equation?

$$AA^T u_i = \sigma_i^2 u_i$$

AA^T
matrix

u_i
vector

=

σ_i^2
scalar

u_i
same vector



What is U ?

Part 3

Question

What is the significance of this equation?

$$AA^T u_i = \sigma_i^2 u_i$$

$\underbrace{AA^T}_{\text{matrix}}$	$\underbrace{u_i}_{\text{vector}}$	=	$\underbrace{\sigma_i^2}_{\text{scalar}}$	$\underbrace{u_i}_{\text{same vector}}$
B	v	=	λ	v

Key Idea

The columns of U are the eigenvectors of AA^T !



Calculating U

Steps to find the left singular vectors

Steps

1 Compute $B = AA^T$, ensuring $B \in \mathbb{R}^{m \times m}$.

2 Solve the eigenvalue equation for B :

$$\det(\lambda I - B) = 0$$

to find the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_m$.

3 Compute the eigenvectors of B , denoted as $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$.

4 Normalize each eigenvector:

$$\mathbf{u}_i = \frac{\mathbf{u}_i}{\|\mathbf{u}_i\|}$$

5 Form U as the matrix with these normalized eigenvectors as columns:

$$U = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_m]$$

What is V ?

- ① The SVD of A :

$$A = USV^T$$

implies:

$$A^T A = (USV^T)^T USV^T = VS^T SV^T$$

- ② Therefore, (letting $S^2 := S^T S$) we have

$$A^T AV = VS^2$$

- ③ The column view gives us

$$A^T Av_i = \sigma_i^2 v_i$$

- ④ By definition, v_i is an eigenvector for $A^T A$



Calculating V

Steps to find the right singular vectors

Steps

① Compute $B = A^T A$, ensuring $B \in \mathbb{R}^{n \times n}$.

② Solve the eigenvalue equation for B :

$$\det(\lambda I - B) = 0$$

to find the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$.

③ Compute the eigenvectors of B , denoted as $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$.

④ Normalize each eigenvector:

$$\mathbf{v}_i = \frac{\mathbf{v}_i}{\|\mathbf{v}_i\|}$$

⑤ Form V as the matrix with these normalized eigenvectors as columns:

$$V = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_n]$$

Summary on $A = USV^T$

$$A = \underbrace{\begin{bmatrix} | & & | \\ u_1 & \dots & u_m \\ | & & | \end{bmatrix}}_{\text{eigenvectors of } AA^T} \underbrace{\begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & \ddots \\ & & & & 0 \end{bmatrix}}_{\text{eigenvalues of both}} \underbrace{\begin{bmatrix} - & v_1^T & - \\ & \vdots & \\ - & v_n^T & - \end{bmatrix}}_{\text{eigenvectors of } A^T A}$$

- 1 Compute $A^T A$
- 2 Find the eigenvalues of $A^T A$ (these are σ_i^2)
- 3 Find the orthonormal eigenvectors of $A^T A$ (these are the rows of V^T)
- 4 Compute AA^T
- 5 Find the orthonormal eigenvectors of AA^T (these are the columns of U)



Example

Example

Find the SVD of

$$A = \begin{bmatrix} 2 & -2 \\ 1 & 1 \end{bmatrix}$$

First construct $A^T A$:

$$A^T A = \begin{bmatrix} 2 & 1 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 2 & -2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 5 & -3 \\ -3 & 5 \end{bmatrix}$$

Eigenvalues: $\lambda_1 = 8$ and $\lambda_2 = 2$. So

$$S^2 = \begin{bmatrix} 8 & 0 \\ 0 & 2 \end{bmatrix} \Rightarrow S = \begin{bmatrix} 2\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix}$$

Key Idea

It's important to sort the singular values from greatest to least.

Example

Now find V^T and U . The columns of V^T are the eigenvectors of $A^T A$.

- $\lambda_1 = 8$: $(A^T A - \lambda_1 I)v_1 = 0$

$$\Rightarrow \begin{bmatrix} -3 & -3 \\ -3 & -3 \end{bmatrix} v_1 = 0 \quad \Rightarrow \quad \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} v_1 = 0 \quad \Rightarrow \quad v_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} \sqrt{2}/2 \\ -\sqrt{2}/2 \end{bmatrix}$$

- $\lambda_2 = 2$: $(A^T A - \lambda_2 I)v_2 = 0$

$$\Rightarrow \begin{bmatrix} 3 & -3 \\ -3 & 3 \end{bmatrix} v_2 = 0 \quad \Rightarrow \quad \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} v_2 = 0 \quad \Rightarrow \quad v_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}$$

- Finally:

$$V = \begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}$$

Key Idea

You must make sure you normalize your eigenvectors. If an eigenvector has mixed signs, stick to a convention, e.g. first element shall be positive.

Example

Now find U . The columns of U are the eigenvectors of AA^T .

- $\lambda_1 = 8$: $(AA^T - \lambda_1 I)u_1 = 0$

$$\Rightarrow \begin{bmatrix} 0 & 0 \\ 0 & -6 \end{bmatrix} u_1 = 0 \quad \Rightarrow \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} u_1 = 0 \quad \Rightarrow \quad u_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

- $\lambda_2 = 2$: $(AA^T - \lambda_2 I)u_2 = 0$

$$\Rightarrow \begin{bmatrix} 6 & 0 \\ 0 & 0 \end{bmatrix} u_2 = 0 \quad \Rightarrow \quad \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} u_2 = 0 \quad \Rightarrow \quad u_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

- Finally:

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

- Together:

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}$$



Remark on Normalization

- There is ambiguity in normalizing a vector $v = [-3 \ 4]$, it can be
 - $v = [-3/5 \ 4/5]$
 - $v = [3/5 \ -4/5]$
- In the previous example,

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

are both orthogonal matrix, but only one will yield the correct SVD

Tips

- Don't forget to sort the eigenvalues of AA^T
- Stick to convention, e.g. "first element of eigenvectors must be negative"
- Verify that $A = USV^T$



Question

What makes SVD such a useful algorithm?



Question

What makes SVD such a useful algorithm?

- SVD sorts the eigenvectors from most important to least important

$$A = \begin{bmatrix} | & & | \\ u_1 & \dots & u_m \\ | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & \ddots \\ & & & & 0 \end{bmatrix} \begin{bmatrix} - & v_1^T & - \\ & \vdots & \\ - & v_n^T & - \end{bmatrix}$$

- Since $\sigma_1 > \sigma_2 > \dots > \sigma_r$, the most important eigenvectors are $(u_1, v_1), (u_2, v_2), \dots$



Compression



Memory Cost of an Image

- A black and white photo is just a matrix $A \in \mathbb{R}^{m \times n}$, n pixels wide and m pixels tall:

$A =$



- The storage cost for A is mn bytes
- If we decompose it as $A = USV^T$, then we the cost is $mm + nn + r$ bytes

Question

How can we use SVD to save on memory cost?



Compression with SVD

- Notice what $A = USV^T$ looks like:

$$A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_r u_r v_r^T + 0 u_{r+1} v_{r+1}^T + \cdots + 0 u_p v_p^T$$

- This is easily truncated to

$$A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_r u_r v_r^T$$

- We can discard useless eigenvectors:

$$A = \underbrace{\begin{bmatrix} | & & | \\ u_1 & \cdots & u_m \\ | & & | \end{bmatrix}}_{\text{keep } r \text{ columns}} \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & \ddots \\ & & & & 0 \end{bmatrix} \underbrace{\begin{bmatrix} - & v_1^T & - \\ & \vdots & \\ - & v_n^T & - \end{bmatrix}}_{\text{keep } r \text{ rows}}$$



Storage Cost Comparison

- Original matrix $A \in \mathbb{R}^{m \times n}$:

Cost: mn bytes

- If $r = \text{rank}(A)$ then

Cost: $r(m + n + 1)$ bytes

- We can compress more aggressively
- Keep only the top k singular values $\sigma_1, \dots, \sigma_k$



Deleting Singular Values

- Choose some integer $k < \text{rank}(A)$
- We can discard least important singular values

$$A = \underbrace{\begin{bmatrix} | & & | & & \\ u_1 & \cdots & u_k & 0 & \cdots \\ | & & | & & \end{bmatrix}}_{\text{keep } k \text{ columns}} \begin{bmatrix} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_k & & \\ & & & 0 & \\ & & & & \ddots \end{bmatrix} \underbrace{\begin{bmatrix} - & v_1^T & - \\ & \vdots & \\ - & v_k^T & - \\ & 0 & \\ - & \vdots & \end{bmatrix}}_{\text{keep } k \text{ rows}}$$

- Cost is $k(m + n + 1)$ bytes



Compression Example

- In Matlab, `[U, S, V] = svd('lena.bmp')`
- Uncompressed 512 by 512 image file: 262144
- Out of the 512 singular values... keep only 50
- Compressed image: $50 \times (512 + 512 + 1) = 51250$



- Compressed image is roughly 20% of the original



Analogy



Figure: Using SVD to compress an image is like breaking off the legs and converting it to a fold-out table.



Deep Fried

Compressed Image

