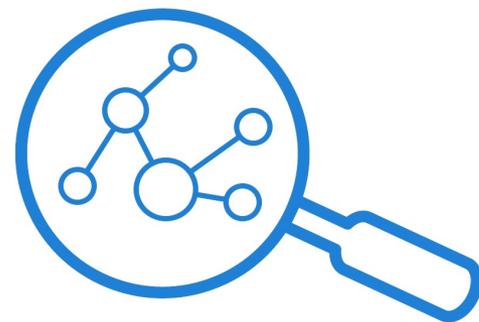


数据科学与大数据技术 的数学基础



第九讲



计算机学院

余皓然

2024/5/20

课程内容

Part1 随机化方法

一致性哈希 布隆过滤器 CM Sketch方法 最小哈希
欧氏距离下的相似搜索 Jaccard相似度下的相似搜索

Part2 谱分析方法

主成分分析 奇异值分解 谱图论

Part3 最优化方法

压缩感知



奇异值分解

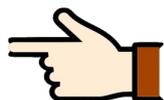
奇异值分解的概念



本讲内容

奇异值分解 (SVD)

- SVD的定义
- 对比SVD与特征分解
- SVD的存在性
- SVD的计算
- SVD的等价形式
- SVD的另一种定义

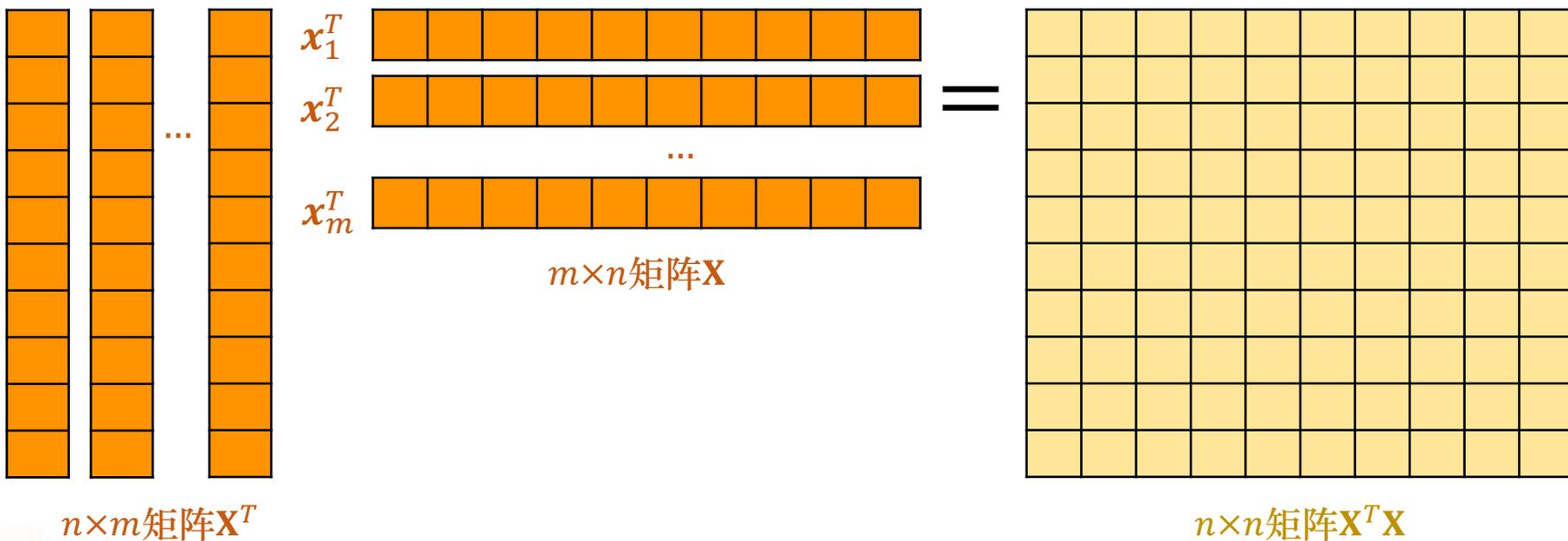


奇异值分解

回顾主成分分析

给定 $x_1, \dots, x_m \in \mathbb{R}^n$ ，先计算 $X^T X$ ，然后计算若干特征值对应的单位特征向量

其中运用到**对称矩阵**的分解： $X^T X = QDQ^T$



奇异值分解

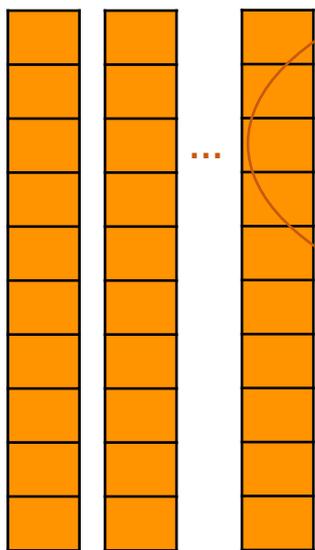


能否对非对称矩阵（包括非方阵的矩阵）进行分解？
比如直接对 $m \times n$ 矩阵 \mathbf{X} 分解？

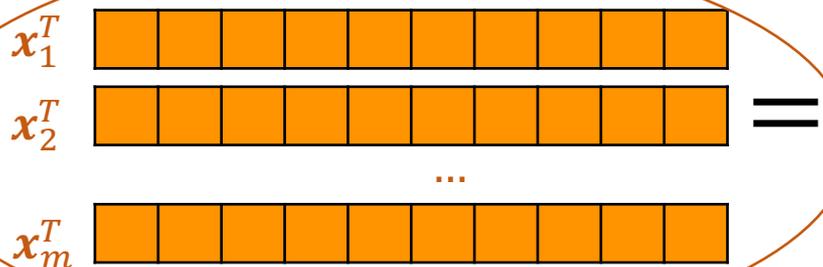
回顾主成分分析

给定 $\mathbf{x}_1, \dots, \mathbf{x}_m \in \mathbb{R}^n$ ，先计算 $\mathbf{X}^T \mathbf{X}$ ，然后计算若干特征值对应的单位特征向量

其中运用到**对称矩阵**的分解： $\mathbf{X}^T \mathbf{X} = \mathbf{Q} \mathbf{D} \mathbf{Q}^T$

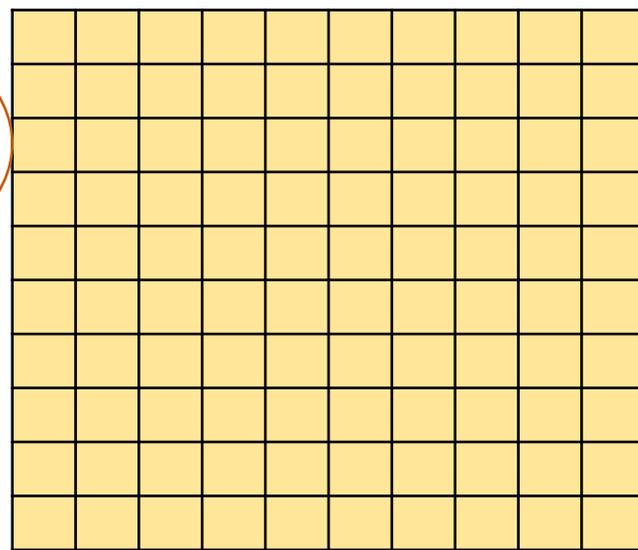


$n \times m$ 矩阵 \mathbf{X}^T



$m \times n$ 矩阵 \mathbf{X}

=



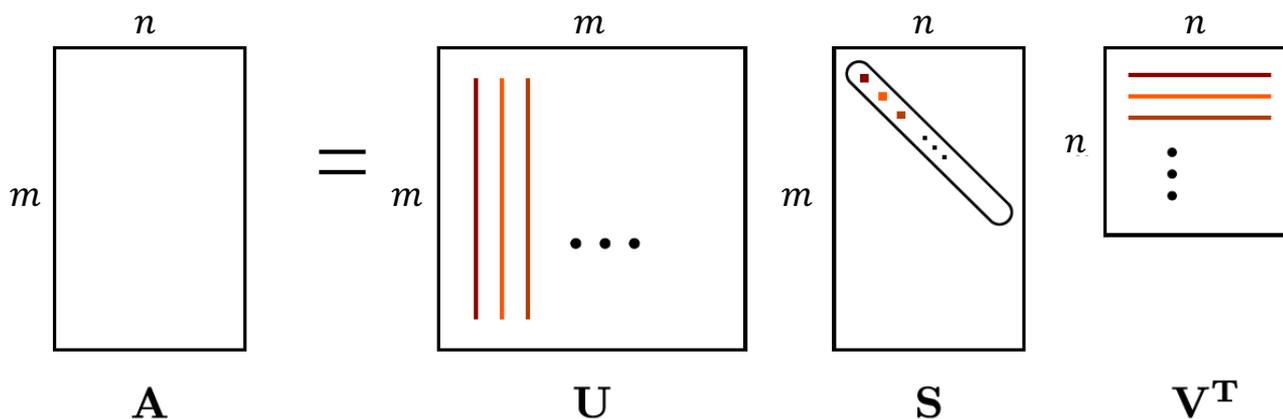
$n \times n$ 矩阵 $\mathbf{X}^T \mathbf{X}$

奇异值分解

奇异值分解 (Singular Value Decomposition)

给定 $m \times n$ 矩阵 A ，将其分解为三个矩阵之积： $A = USV^T$ ，其中：

- (1) U 是 $m \times m$ 正交矩阵；
- (2) V 是 $n \times n$ 正交矩阵；
- (3) S 是 $m \times n$ (矩形) 对角矩阵且所有元素非负。



奇异值分解

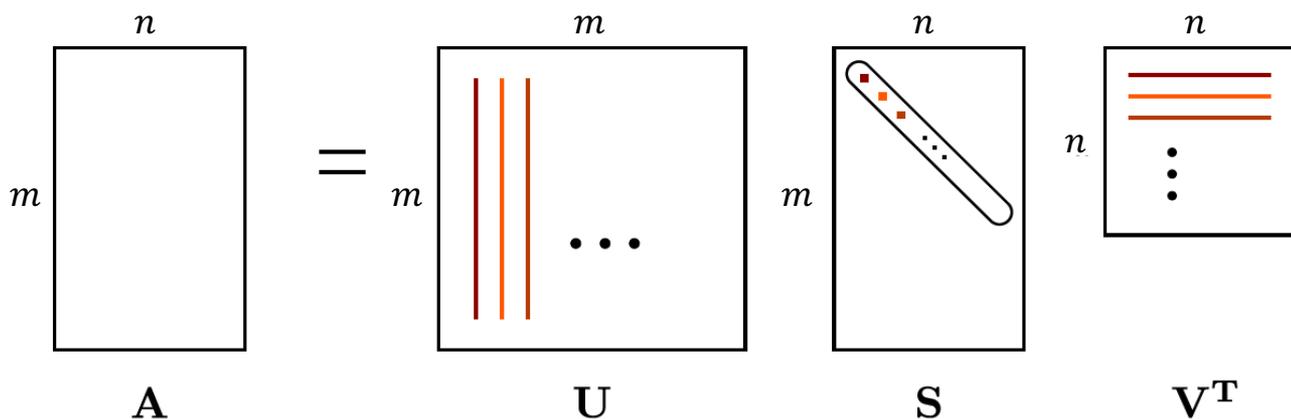
例

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ 2 & 0 & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

奇异值分解 (Singular Value Decomposition)

给定 $m \times n$ 矩阵 A ，将其分解为三个矩阵之积： $A = USV^T$ ，其中：

- (1) U 是 $m \times m$ 正交矩阵；列向量（行向量）构成标准正交向量组
- (2) V 是 $n \times n$ 正交矩阵；
- (3) S 是 $m \times n$ （矩形）对角矩阵且所有元素非负。

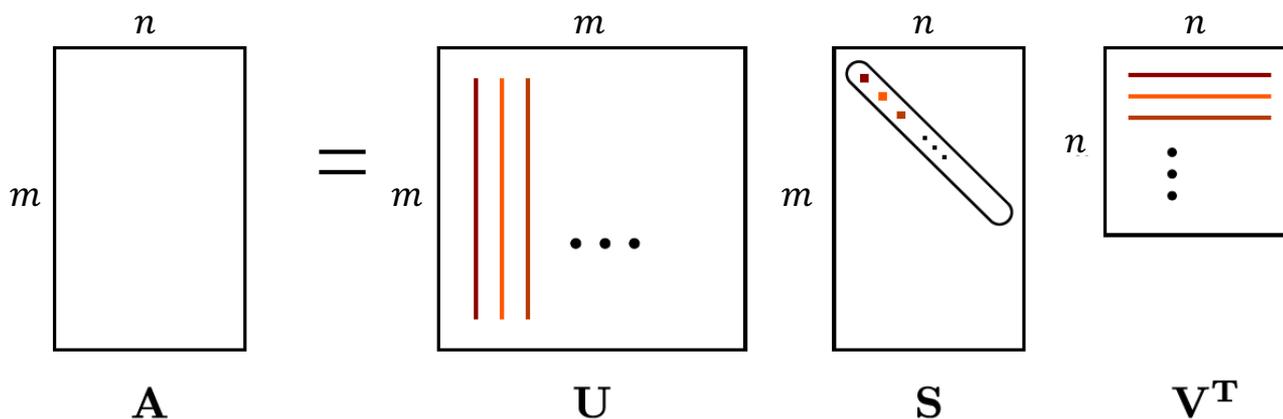


奇异值分解

奇异值分解 (Singular Value Decomposition)

给定 $m \times n$ 矩阵 A ，将其分解为三个矩阵之积： $A = USV^T$ ，其中：

- (1) U 是 $m \times m$ 正交矩阵；
- (2) V 是 $n \times n$ 正交矩阵；
- (3) S 是 $m \times n$ (矩形) 对角矩阵且所有元素非负。



* 不妨设 S 的对角线元素由高到低排列

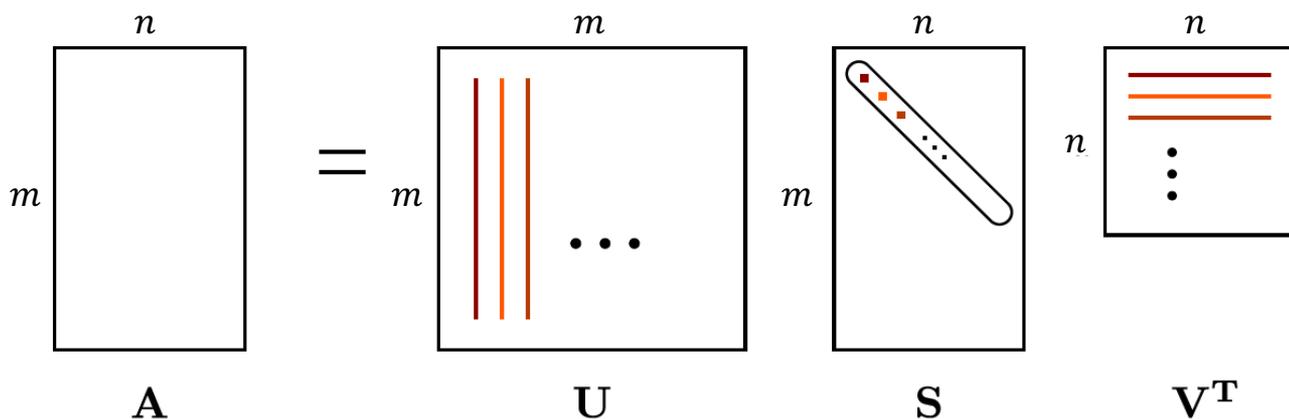
(在调整对角线元素顺序的同时相应调整 U 的列向量、 V^T 的行向量的顺序，结果不变)

奇异值分解

奇异值分解 (Singular Value Decomposition)

给定 $m \times n$ 矩阵 A ，将其分解为三个矩阵之积： $A = USV^T$ ，其中：

- (1) U 是 $m \times m$ 正交矩阵；
- (2) V 是 $n \times n$ 正交矩阵；
- (3) S 是 $m \times n$ (矩形) 对角矩阵且所有元素非负。



U 的 m 个列向量称为 A 的左奇异向量

V 的 n 个列向量 (即 V^T 的 n 个行向量) 称为 A 的右奇异向量

S 的对角线元素为 A 的奇异值

奇异值分解

奇异值分解 (Singular Value Decomposition)

给定 $m \times n$ 矩阵 A ，将其分解为三个矩阵之积： $A = USV^T$ ，其中：

- (1) U 是 $m \times m$ 正交矩阵；
- (2) V 是 $n \times n$ 正交矩阵；
- (3) S 是 $m \times n$ (矩形) 对角矩阵且所有元素非负。

SVD号称矩阵分解中的“瑞士军刀”/“劳斯莱斯”
 (“Bridging the Gap Between Numerical Linear Algebra,
Theoretical Computer Science, and Data Applications”)

"singular value decomposition"



About 690,000 results (0.06 sec)

On the early history of the singular value decomposition

GW Stewart - SIAM review, 1993 - SIAM

... Thus the advent of the **singular value decomposition** in 1873 ... the early work on the **singular value decomposition**. Most of it ... we will be concerned with the **singular value decomposition** ...

☆ Save Cite Cited by 943 Related articles All 15 versions

Singular value decomposition and least squares solutions

GH Golub, C Reinsch - Linear algebra, 1971 - Springer

... 1 To compute the **singular value decomposition** of a given matrix A , Forsythe and Henrici [2], Hestenes [8], and Kogbetliantz [9] proposed methods based on plane rotations ...

☆ Save Cite Cited by 4152 Related articles All 13 versions

[PDF] Singular value decomposition tutorial

K Baker - The Ohio State University, 2005 - site.iugaza.edu.ps

... It's about the mechanics of **singular value decomposition**, ... who knew zilch about **singular value decomposition** or any of the ... of **singular value decomposition** before to be able to do it. ...

☆ Save Cite Cited by 303 Related articles All 13 versions

The singular value decomposition: Its computation and some applications

V Klementa, A Laub - IEEE Transactions on automatic control, 1980 - ieeexplore.ieee.org

We provide a tutorial introduction to certain numerical computations both in linear algebra and linear systems in the context of bounded arithmetic. The essential characteristics of ...

☆ Save Cite Cited by 1801 Related articles All 10 versions

奇异值分解

奇异值分解 (Singular Value Decomposition)

给定 $m \times n$ 矩阵 A ，将其分解为三个矩阵之积： $A = USV^T$ ，其中：

- (1) U 是 $m \times m$ 正交矩阵；
- (2) V 是 $n \times n$ 正交矩阵；
- (3) S 是 $m \times n$ (矩形) 对角矩阵且所有元素非负。

广泛的应用 (下讲内容) :

- 数据压缩
- 矩阵补全
- 实体嵌入
- ...



本讲内容

奇异值分解 (SVD)

➤ SVD的定义

➤ 对比SVD与特征分解



➤ SVD的存在性

➤ SVD的计算

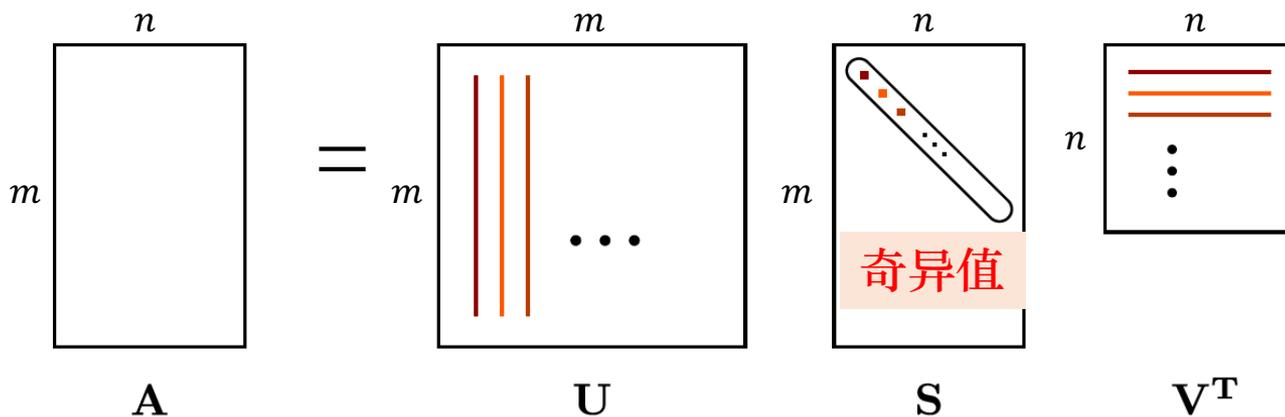
➤ SVD的等价形式

➤ SVD的另一种定义

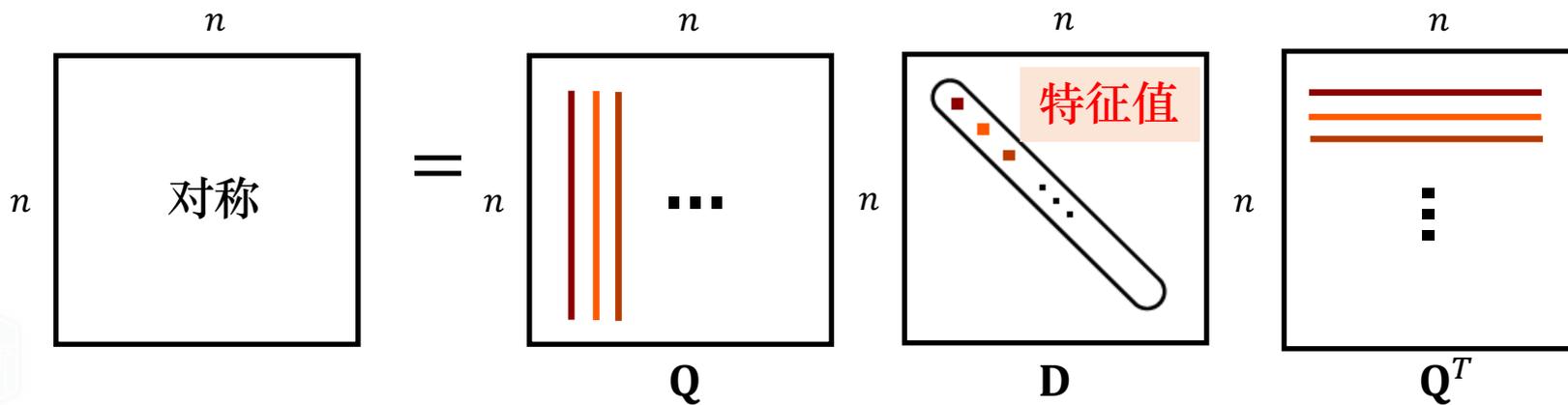


奇异值分解与特征分解

$m \times n$ 矩阵 A 的奇异值分解



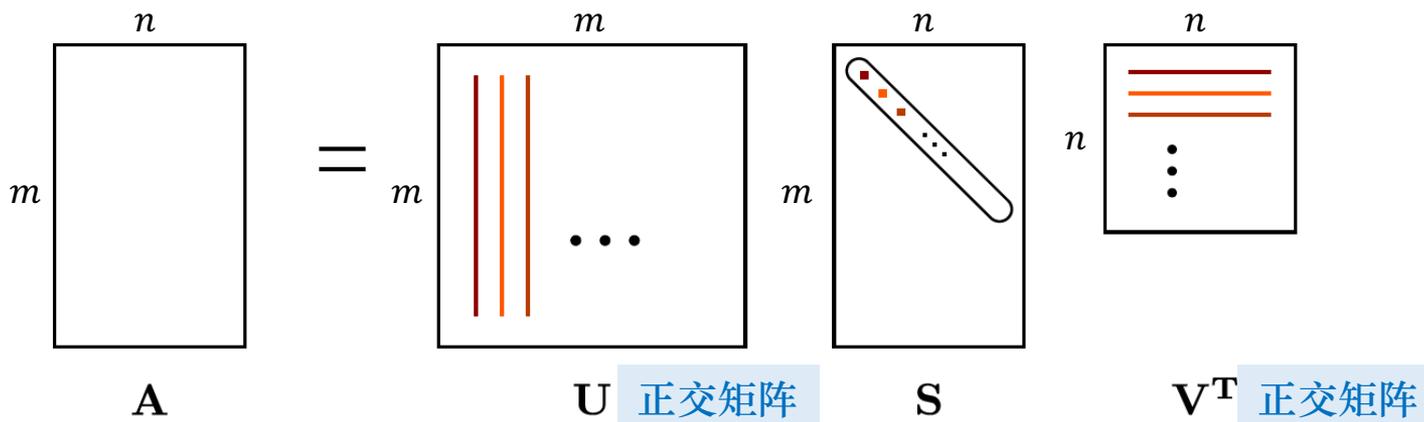
$n \times n$ 对称矩阵的特征分解



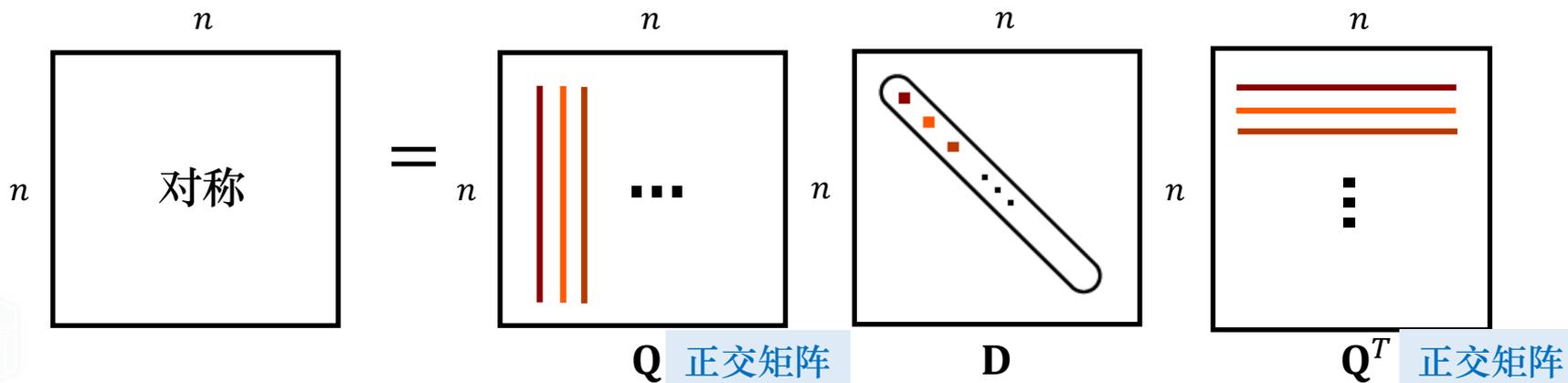
奇异值分解与特征分解

U与V可能不同

$m \times n$ 矩阵A的奇异值分解



$n \times n$ 对称矩阵的特征分解



本讲内容

奇异值分解 (SVD)

- SVD的定义
- 对比SVD与特征分解
- **SVD的存在性**
- SVD的计算
- SVD的等价形式
- SVD的另一种定义



奇异值分解的存在性

为什么对任意 $m \times n$ 矩阵 A ，一定可将其分解成 USV^T

回顾： $n \times n$ 对称矩阵 $A^T A$ 可分解成 QDQ^T ，如何据此说明 A 可写成 USV^T ？



奇异值分解的存在性

为什么对任意 $m \times n$ 矩阵 A ，一定可将其分解成 USV^T

回顾： $n \times n$ 对称矩阵 $A^T A$ 可分解成 QDQ^T ，如何据此说明 A 可写成 USV^T ？

若 $A = USV^T$ ，则有 $A^T A = VS^T U^T USV^T = VS^T SV^T$

所以可以将 V 取为 Q ， $S^T S$ 取为 D

如何求矩阵 U ？



奇异值分解的存在性

为什么对任意 $m \times n$ 矩阵 A ，一定可将其分解成 USV^T

回顾： $n \times n$ 对称矩阵 $A^T A$ 可分解成 QDQ^T ，如何据此说明 A 可写成 USV^T ？

若 $A = USV^T$ ，则有 $A^T A = VS^T U^T USV^T = VS^T SV^T$

所以可以将 V 取为 Q ， $S^T S$ 取为 D

如何求矩阵 U ？

不妨设 $m \geq n$

令 v_1, \dots, v_n 分别为 q_1, \dots, q_n ，即 $A^T A$ 的（标准正交）特征向量

奇异值分解的存在性

为什么对任意 $m \times n$ 矩阵 A ，一定可将其分解成 USV^T

回顾： $n \times n$ 对称矩阵 $A^T A$ 可分解成 QDQ^T ，如何据此说明 A 可写成 USV^T ？

若 $A = USV^T$ ，则有 $A^T A = VS^T U^T USV^T = VS^T SV^T$

所以可以将 V 取为 Q ， $S^T S$ 取为 D

不妨设 $m \geq n$

令 v_1, \dots, v_n 分别为 q_1, \dots, q_n ，即 $A^T A$ 的（标准正交）特征向量

令 S 对角线元素为 $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}$ （非负），即根号下 $A^T A$ 特征值（根号下 D 对角线元素）

奇异值分解的存在性

为什么对任意 $m \times n$ 矩阵 A ，一定可将其分解成 USV^T

回顾： $n \times n$ 对称矩阵 $A^T A$ 可分解成 QDQ^T ，如何据此说明 A 可写成 USV^T ？

若 $A = USV^T$ ，则有 $A^T A = VS^T U^T USV^T = VS^T SV^T$

所以可以将 V 取为 Q ， $S^T S$ 取为 D

不妨设 $m \geq n$

令 v_1, \dots, v_n 分别为 q_1, \dots, q_n ，即 $A^T A$ 的（标准正交）特征向量

令 S 对角线元素为 $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}$ （非负），即根号下 $A^T A$ 特征值（根号下 D 对角线元素）

取 u_1, \dots, u_m 为 ？

奇异值分解的存在性

为什么对任意 $m \times n$ 矩阵 A ，一定可将其分解成 USV^T

回顾： $n \times n$ 对称矩阵 $A^T A$ 可分解成 QDQ^T ，如何据此说明 A 可写成 USV^T ？

若 $A = USV^T$ ，则有 $A^T A = VS^T U^T USV^T = VS^T SV^T$

所以可以将 V 取为 Q ， $S^T S$ 取为 D

不妨设 $m \geq n$

令 v_1, \dots, v_n 分别为 q_1, \dots, q_n ，即 $A^T A$ 的（标准正交）特征向量

令 S 对角线元素为 $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}$ （非负），即根号下 $A^T A$ 特征值（根号下 D 对角线元素）

取 u_1, \dots, u_m 为 ？

由 $A = USV^T$ 可得 $AV = US = \begin{bmatrix} | & & | \\ \sqrt{\lambda_1} u_1 & \dots & \sqrt{\lambda_n} u_n \\ | & & | \end{bmatrix}$ ，即有 $Av_i = \sqrt{\lambda_i} u_i$

奇异值分解的存在性

为什么对任意 $m \times n$ 矩阵 A ，一定可将其分解成 USV^T

回顾： $n \times n$ 对称矩阵 $A^T A$ 可分解成 QDQ^T ，如何据此说明 A 可写成 USV^T ？

若 $A = USV^T$ ，则有 $A^T A = VS^T U^T USV^T = VS^T SV^T$

所以可以将 V 取为 Q ， $S^T S$ 取为 D

不妨设 $m \geq n$

令 v_1, \dots, v_n 分别为 q_1, \dots, q_n ，即 $A^T A$ 的（标准正交）特征向量

令 S 对角线元素为 $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}$ （非负），即根号下 $A^T A$ 特征值（根号下 D 对角线元素）

取 u_1, \dots, u_n 为 $\frac{Av_1}{\sqrt{\lambda_1}}, \dots, \frac{Av_n}{\sqrt{\lambda_n}}$

不难证明 u_1, \dots, u_n 构成标准正交向量组（证明单位向量时用 $\left\| \frac{Av_1}{\sqrt{\lambda_1}} \right\| = \sqrt{\frac{(Av_1)^T Av_1}{\sqrt{\lambda_1} \sqrt{\lambda_1}}}$ ）

需要 U 是 $m \times m$ 矩阵，即 m 个列向量，该如何选取 u_{n+1}, \dots, u_m ？

奇异值分解的存在性

例

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ 2 & 0 & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

为什么对任意 $m \times n$ 矩阵 A ，一定可将其分解成 USV^T

回顾： $n \times n$ 对称矩阵 $A^T A$ 可分解成 QDQ^T ，如何据此说明 A 可写成 USV^T ？

若 $A = USV^T$ ，则有 $A^T A = VS^T U^T USV^T = VS^T SV^T$

所以可以将 V 取为 Q ， $S^T S$ 取为 D

不妨设 $m \geq n$

令 v_1, \dots, v_n 分别为 q_1, \dots, q_n ，即 $A^T A$ 的（标准正交）特征向量

令 S 对角线元素为 $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}$ （非负），即根号下 $A^T A$ 特征值（根号下 D 对角线元素）

取 u_1, \dots, u_n 为 $\frac{Av_1}{\sqrt{\lambda_1}}, \dots, \frac{Av_n}{\sqrt{\lambda_n}}$

不难证明 u_1, \dots, u_n 构成标准正交向量组（证明单位向量时用 $\left\| \frac{Av_1}{\sqrt{\lambda_1}} \right\| = \sqrt{\frac{(Av_1)^T Av_1}{\sqrt{\lambda_1} \sqrt{\lambda_1}}}$ ）

选取 u_{n+1}, \dots, u_m 令与 u_1, \dots, u_n 共同构成标准正交向量组即可。找到用于奇异值分解的 U, S, V

奇异值分解的存在性

为什么对任意 $m \times n$ 矩阵 A ，一定可将其分解成 USV^T

回顾： $n \times n$ 对称矩阵 $A^T A$ 可分解成 QDQ^T ，如何据此说明 A 可写成 USV^T ?

若 $A = USV^T$ ，则有 $A^T A = VS^T U^T USV^T = VS^T SV^T$

所以可以将 V 取为 Q ， $S^T S$ 取为 D

不妨设 $m \geq n$

令 v_1, \dots, v_n 分别为 q_1, \dots, q_n ，即 $A^T A$ 的（标准正交）特征向量

令 S 对角线元素为 $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}$ （非负），即根号下 $A^T A$ 特征值（根号下 D 对角线元素）

取 u_1, \dots, u_n 为 $\frac{Av_1}{\sqrt{\lambda_1}}, \dots, \frac{Av_n}{\sqrt{\lambda_n}}$

不难证明 u_1, \dots, u_n 构成标准正交向量组（证明单位向量时用 $\left\| \frac{Av_1}{\sqrt{\lambda_1}} \right\| = \sqrt{\frac{(Av_1)^T Av_1}{\sqrt{\lambda_1} \sqrt{\lambda_1}}}$ ）

选取 u_{n+1}, \dots, u_m 令与 u_1, \dots, u_n 共同构成标准正交向量组即可。找到用于奇异值分解的 U, S, V

矩阵的奇异值分解是否唯一?

不唯一, $A^T A$ 的特征向量不唯一

奇异值分解的存在性

为什么对任意 $m \times n$ 矩阵 A , 一定可将其分解成 USV^T

回顾: $n \times n$ 对称矩阵 $A^T A$ 可分解成 QDQ^T , 如何据此说明 A 可写成 USV^T ?

若 $A = USV^T$, 则有 $A^T A = VS^T U^T USV^T = VS^T SV^T$

所以可以将 V 取为 Q , $S^T S$ 取为 D

不妨设 $m \geq n$

令 v_1, \dots, v_n 分别为 q_1, \dots, q_n , 即 $A^T A$ 的 (标准正交) 特征向量

令 S 对角线元素为 $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}$ (非负), 即根号下 $A^T A$ 特征值 (根号下 D 对角线元素)

取 u_1, \dots, u_n 为 $\frac{Av_1}{\sqrt{\lambda_1}}, \dots, \frac{Av_n}{\sqrt{\lambda_n}}$

不难证明 u_1, \dots, u_n 构成标准正交向量组 (证明单位向量时用 $\left\| \frac{Av_1}{\sqrt{\lambda_1}} \right\| = \sqrt{\frac{(Av_1)^T Av_1}{\sqrt{\lambda_1} \sqrt{\lambda_1}}}$)

选取 u_{n+1}, \dots, u_m 令与 u_1, \dots, u_n 共同构成标准正交向量组即可。找到用于奇异值分解的 U, S, V

本讲内容

奇异值分解 (SVD)

- SVD的定义
- 对比SVD与特征分解
- SVD的存在性
- **SVD的计算** 
- SVD的等价形式
- SVD的另一种定义



奇异值分解的计算

给定 $m \times n$ 矩阵 A ($m \geq n$)，计算其分解 USV^T 的基本方法：

(1) 计算 $A^T A$

(2) 对 $A^T A$ 进行特征分解： $A^T A = QDQ^T$ ，取 $V = Q$

(3) 计算 AV

(4) 取 $s_i = \|Av_i\|$ 且 $u_i = \frac{Av_i}{\|Av_i\|}$

(5) 补齐正交矩阵



奇异值分解的计算

给定 $m \times n$ 矩阵 \mathbf{A} ($m \geq n$)，计算其分解 \mathbf{USV}^T 的基本方法：

(1) 计算 $\mathbf{A}^T \mathbf{A}$

(2) 对 $\mathbf{A}^T \mathbf{A}$ 进行特征分解： $\mathbf{A}^T \mathbf{A} = \mathbf{QDQ}^T$ ，取 $\mathbf{V} = \mathbf{Q}$

(3) 计算 \mathbf{AV}

(4) 取 $s_i = \|\mathbf{A}\mathbf{v}_i\|$ 且 $\mathbf{u}_i = \frac{\mathbf{A}\mathbf{v}_i}{\|\mathbf{A}\mathbf{v}_i\|}$

(5) 补齐正交矩阵

若 $\mathbf{A} = \begin{bmatrix} 3 & 0 \\ 4 & 5 \end{bmatrix}$ ，可得 $\mathbf{A}^T \mathbf{A} = \begin{bmatrix} 25 & 20 \\ 20 & 25 \end{bmatrix}$

$\mathbf{A}^T \mathbf{A}$ 的特征值满足 $\left| \begin{bmatrix} 25 & 20 \\ 20 & 25 \end{bmatrix} - \lambda \mathbf{I} \right| = 0$ ，即 $\lambda_1 = 45, \lambda_2 = 5$ ，有 $s_1 = \sqrt{45}, s_2 = \sqrt{5}$

λ_1, λ_2 对应特征向量满足 $\mathbf{A}^T \mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ ，即 $\mathbf{v}_1, \mathbf{v}_2$ 满足 $v_{11} = v_{12}, v_{21} = -v_{22}$

化为单位向量得 $\mathbf{v}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 1 \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ 1 \\ \frac{1}{\sqrt{2}} \end{bmatrix}$ ，继而有 $\mathbf{u}_1 = \begin{bmatrix} \frac{1}{\sqrt{10}} \\ 3 \\ \frac{1}{\sqrt{10}} \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -\frac{3}{\sqrt{10}} \\ 1 \\ \frac{1}{\sqrt{10}} \end{bmatrix}$

奇异值分解的计算

给定 $m \times n$ 矩阵 A ($m \geq n$)，计算其分解 USV^T 的基本方法：

(1) 计算 $A^T A$

(2) 对 $A^T A$ 进行特征分解： $A^T A = QDQ^T$ ，取 $V = Q$

(3) 计算 AV

(4) 取 $s_i = \|Av_i\|$ 且 $u_i = \frac{Av_i}{\|Av_i\|}$

(5) 补齐正交矩阵

若 $A = \begin{bmatrix} 3 & 0 \\ 4 & 5 \end{bmatrix}$ ，可得 $A^T A = \begin{bmatrix} 25 & 20 \\ 20 & 25 \end{bmatrix}$

$$A = \begin{bmatrix} 1 & 3 \\ \sqrt{10} & \sqrt{10} \\ 3 & 1 \\ \sqrt{10} & \sqrt{10} \end{bmatrix} \begin{bmatrix} \sqrt{45} & 0 \\ 0 & \sqrt{5} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ \sqrt{2} & \sqrt{2} \\ -1 & 1 \\ -\sqrt{2} & \sqrt{2} \end{bmatrix}$$

$A^T A$ 的特征值满足 $\begin{vmatrix} 25 & 20 \\ 20 & 25 \end{vmatrix} - \lambda I = 0$ ，即 $\lambda_1 = 45, \lambda_2 = 5$ ，有 $s_1 = \sqrt{45}, s_2 = \sqrt{5}$

λ_1, λ_2 对应特征向量满足 $A^T A v = \lambda v$ ，即 v_1, v_2 满足 $v_{11} = v_{12}, v_{21} = -v_{22}$

化为单位向量得 $v_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 1 \\ \frac{1}{\sqrt{2}} \end{bmatrix}, v_2 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ 1 \\ \frac{1}{\sqrt{2}} \end{bmatrix}$ ，继而有 $u_1 = \begin{bmatrix} \frac{1}{\sqrt{10}} \\ 3 \\ \frac{1}{\sqrt{10}} \end{bmatrix}, u_2 = \begin{bmatrix} -\frac{3}{\sqrt{10}} \\ 1 \\ \frac{1}{\sqrt{10}} \end{bmatrix}$

奇异值分解的计算

给定 $m \times n$ 矩阵 A ($m \geq n$)，计算其分解 USV^T 的基本方法：

(1) 计算 $A^T A$ $(O(mn^2))$

(2) 对 $A^T A$ 进行特征分解： $A^T A = QDQ^T$ ，取 $V = Q$ $(O(n^3))$

(3) 计算 AV $(O(mn^2))$

(4) 取 $s_i = \|Av_i\|$ 且 $u_i = \frac{Av_i}{\|Av_i\|}$ $(O(mn))$

(5) 补齐正交矩阵

$k_1 \times k_2$ 矩阵乘 $k_2 \times k_3$ 矩阵的复杂度 $O(k_1 k_2 k_3)$

复杂度： $O(mn^2)$ 对于实际应用场景太大

实际中采用更快的方法计算（如无需计算 $A^T A$ ）

本讲内容

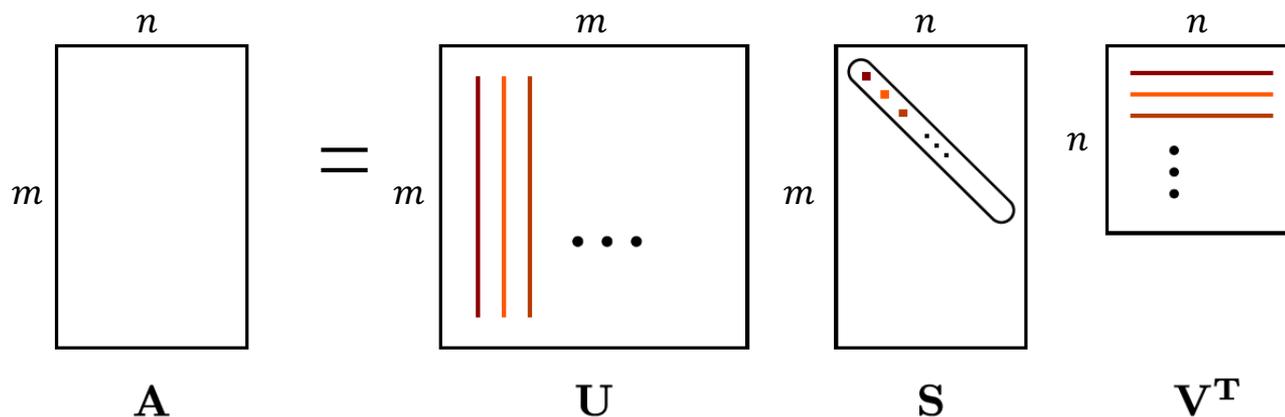
奇异值分解 (SVD)

- SVD的定义
- 对比SVD与特征分解
- SVD的存在性
- SVD的计算
- SVD的等价形式 
- SVD的另一种定义



等价形式

$m \times n$ 矩阵 A 的奇异值分解



$A = USV^T$ 等价于

$$A = \sum_{i=1}^{\min\{m,n\}} s_i \cdot u_i v_i^T$$

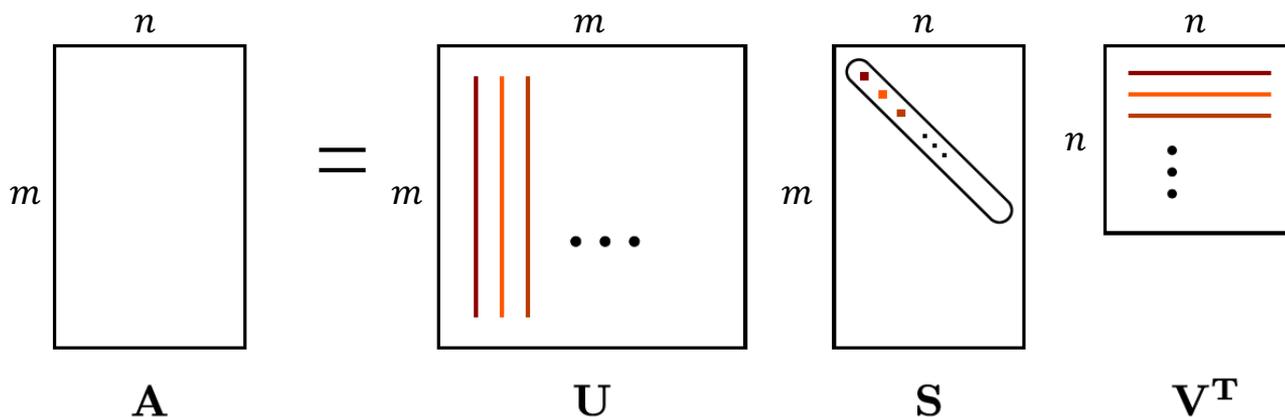
$u_i v_i^T$ 是秩为1的 $m \times n$ 矩阵

第 i 个奇异值 V^T 的第 i 个行向量



等价形式

$m \times n$ 矩阵 A 的奇异值分解



$A = USV^T$ 等价于

$$A = \sum_{i=1}^{\min\{m,n\}} s_i \cdot u_i v_i^T$$

第 i 个奇异值

V^T 的第 i 个行向量

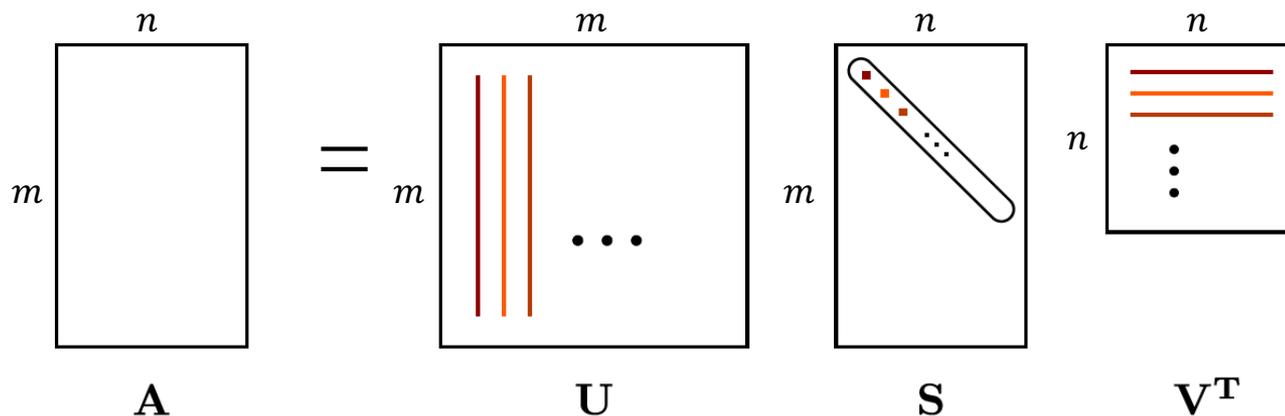
$u_i v_i^T$ 是秩为 1 的 $m \times n$ 矩阵

$$\begin{bmatrix} -u_{i1} v_i^T & - \\ -u_{i2} v_i^T & - \\ \dots & \dots \\ -u_{im} v_i^T & - \end{bmatrix}$$

所有 m 个行向量可写作 1 个向量的线性组合

等价形式

$m \times n$ 矩阵 A 的奇异值分解



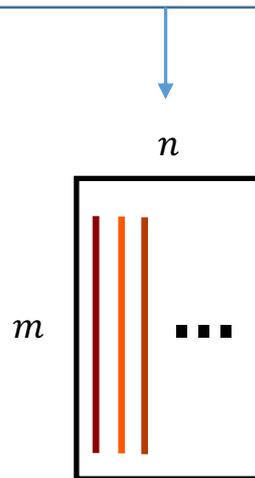
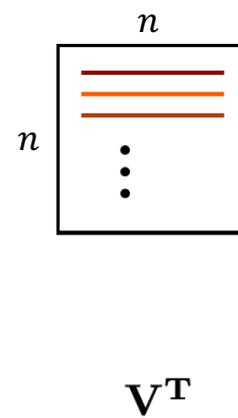
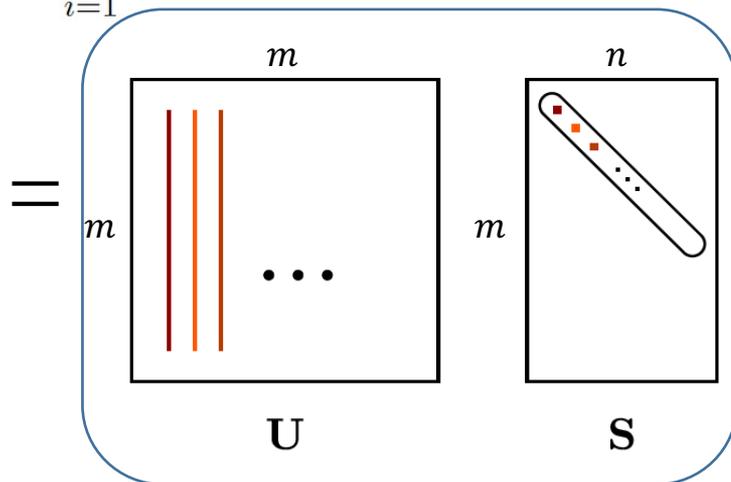
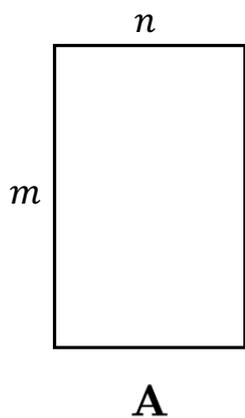
$A = USV^T$ 等价于

$$A = \sum_{i=1}^{\min\{m,n\}} s_i \cdot \mathbf{u}_i \mathbf{v}_i^T$$

奇异值分解将 $m \times n$ 矩阵 A 展成 $\min\{m,n\}$ 个秩为 1 的矩阵的 **非负** 线性组合

等价形式的证明

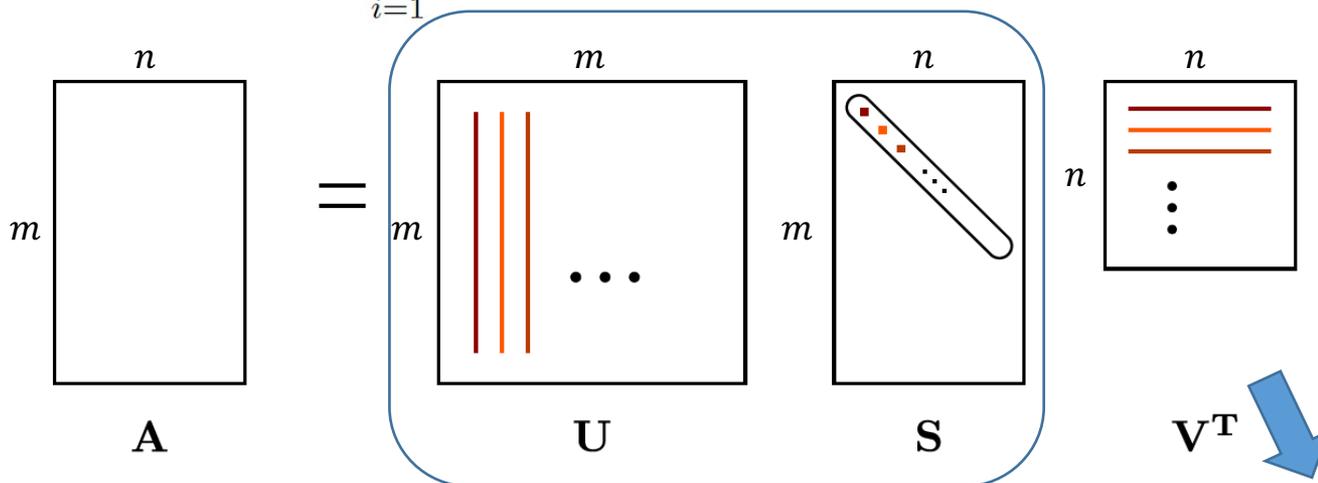
$A = USV^T$ 等价于 $A = \sum_{i=1}^{\min\{m,n\}} s_i \cdot \mathbf{u}_i \mathbf{v}_i^T$ 此例中 $m \geq n$



第 k 列为 $s_k \mathbf{u}_k$ (仅 n 列)

等价形式的证明

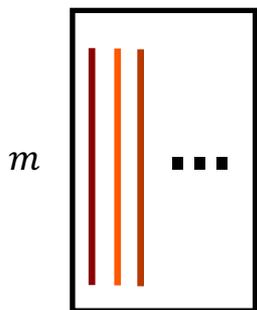
$A = USV^T$ 等价于 $A = \sum_{i=1}^{\min\{m,n\}} s_i \cdot \mathbf{u}_i \mathbf{v}_i^T$ 此例中 $m \geq n$



相乘得到矩阵的第*i*行第*j*列为

$$\sum_{k=1}^n s_k (\mathbf{u}_k \text{ 第 } i \text{ 元素}) (\mathbf{v}_k \text{ 第 } j \text{ 元素})$$

第*k*列为 $s_k \mathbf{u}_k$ (仅*n*列)

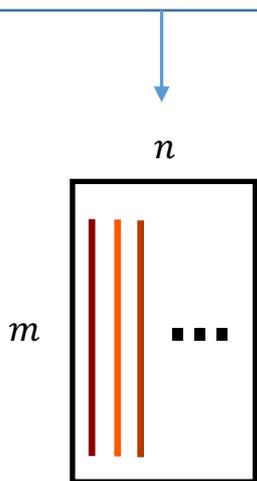
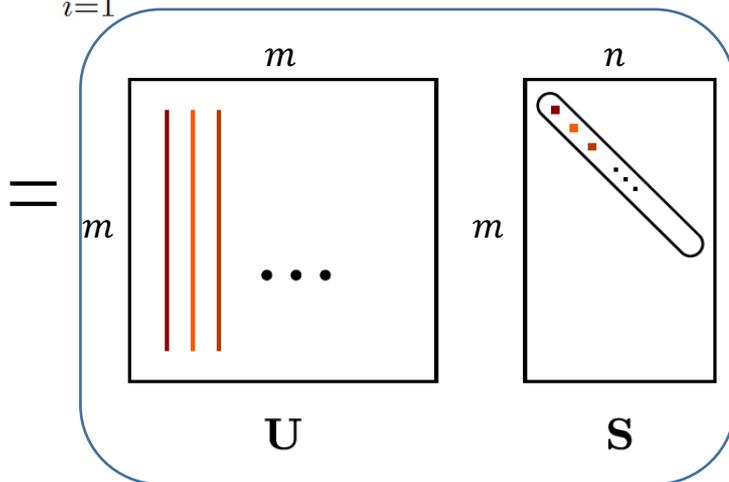
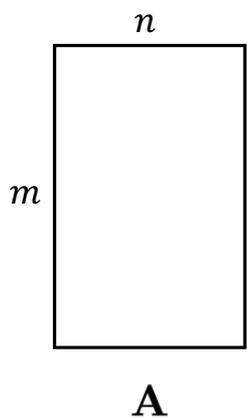


等价形式的证明

第*i*行第*j*列为

$$\sum_{k=1}^n s_k (\mathbf{u}_k \text{ 第 } i \text{ 元素}) (\mathbf{v}_k \text{ 第 } j \text{ 元素})$$

$A = \mathbf{U}\mathbf{S}\mathbf{V}^T$ 等价于 $A = \sum_{i=1}^{\min\{m,n\}} s_i \cdot \mathbf{u}_i \mathbf{v}_i^T$ 此例中 $m \geq n$



相乘得到矩阵的第*i*行第*j*列为

$$\sum_{k=1}^n s_k (\mathbf{u}_k \text{ 第 } i \text{ 元素}) (\mathbf{v}_k \text{ 第 } j \text{ 元素})$$

第*k*列为 $s_k \mathbf{u}_k$ (仅*n*列)



本讲内容

奇异值分解 (SVD)

- SVD的定义
- 对比SVD与特征分解
- SVD的存在性
- SVD的计算
- SVD的等价形式
- SVD的另一种定义



奇异值分解的另一种定义

$$\mathbf{A} = \mathbf{USV}^T \text{ 等价于 } \mathbf{A} = \sum_{i=1}^{\min\{m,n\}} s_i \cdot \mathbf{u}_i \mathbf{v}_i^T$$



奇异值分解的另一种定义

$$\mathbf{A} = \mathbf{USV}^T \text{ 等价于 } \mathbf{A} = \sum_{i=1}^{\min\{m,n\}} s_i \cdot \mathbf{u}_i \mathbf{v}_i^T$$

回顾： $m \geq n$ 时， s_1, \dots, s_n 为 $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}$ ，即根号下 $\mathbf{A}^T \mathbf{A}$ 特征值

若恰有 $r \leq n$ 个非零特征值（即 $\lambda_1, \dots, \lambda_r > 0, \lambda_{r+1}, \dots, \lambda_n = 0$ ），则 $\mathbf{A} = \sum_{i=1}^r s_i \cdot \mathbf{u}_i \mathbf{v}_i^T$

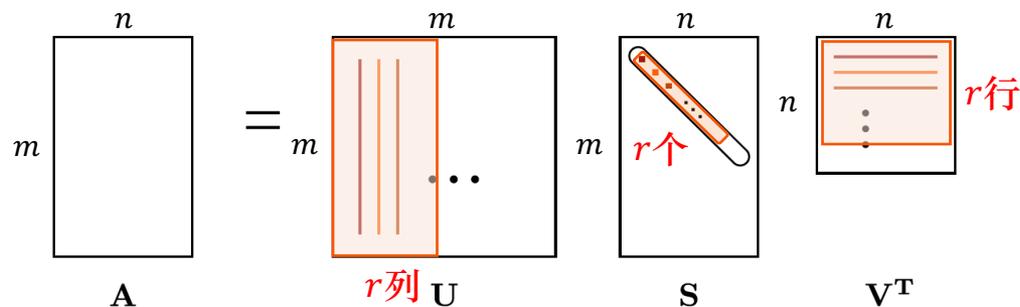


奇异值分解的另一种定义

$$\mathbf{A} = \mathbf{USV}^T \text{ 等价于 } \mathbf{A} = \sum_{i=1}^{\min\{m,n\}} s_i \cdot \mathbf{u}_i \mathbf{v}_i^T$$

回顾: $m \geq n$ 时, s_1, \dots, s_n 为 $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}$, 即根号下 $\mathbf{A}^T \mathbf{A}$ 特征值

若恰有 $r \leq n$ 个非零特征值 (即 $\lambda_1, \dots, \lambda_r > 0, \lambda_{r+1}, \dots, \lambda_n = 0$), 则 $\mathbf{A} = \sum_{i=1}^r s_i \cdot \mathbf{u}_i \mathbf{v}_i^T$



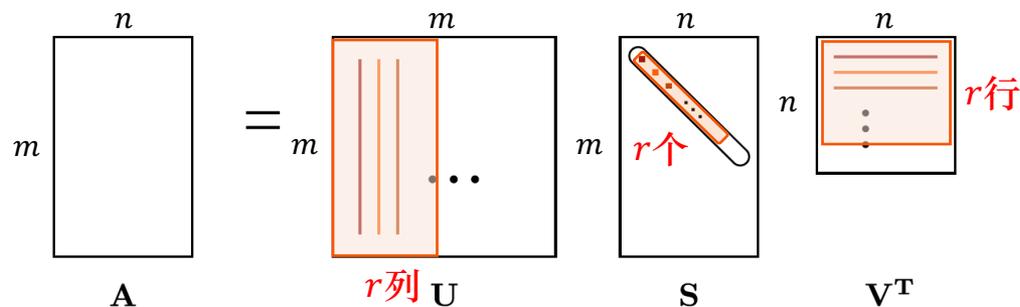
其余列/行向量是
为了凑成正交矩阵

奇异值分解的另一种定义

$$\mathbf{A} = \mathbf{USV}^T \text{ 等价于 } \mathbf{A} = \sum_{i=1}^{\min\{m,n\}} s_i \cdot \mathbf{u}_i \mathbf{v}_i^T$$

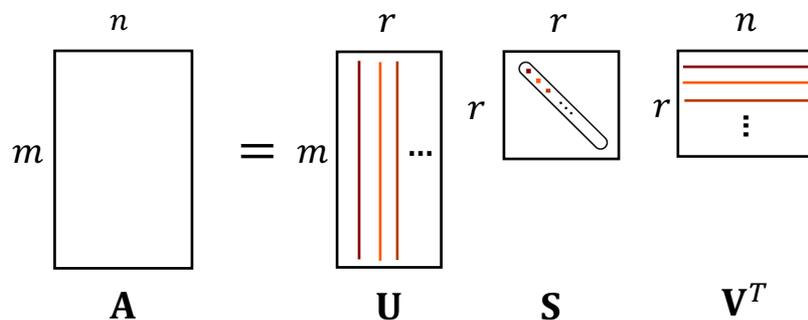
回顾: $m \geq n$ 时, s_1, \dots, s_n 为 $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}$, 即根号下 $\mathbf{A}^T \mathbf{A}$ 特征值

若恰有 $r \leq n$ 个非零特征值 (即 $\lambda_1, \dots, \lambda_r > 0, \lambda_{r+1}, \dots, \lambda_n = 0$), 则 $\mathbf{A} = \sum_{i=1}^r s_i \cdot \mathbf{u}_i \mathbf{v}_i^T$



其余列/行向量是
为了凑成正交矩阵

SVD的另一种定义

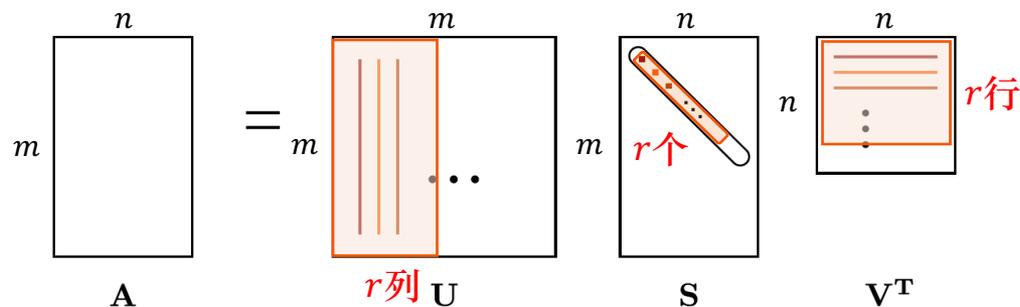


奇异值分解的另一种定义

$$\mathbf{A} = \mathbf{USV}^T \text{ 等价于 } \mathbf{A} = \sum_{i=1}^{\min\{m,n\}} s_i \cdot \mathbf{u}_i \mathbf{v}_i^T$$

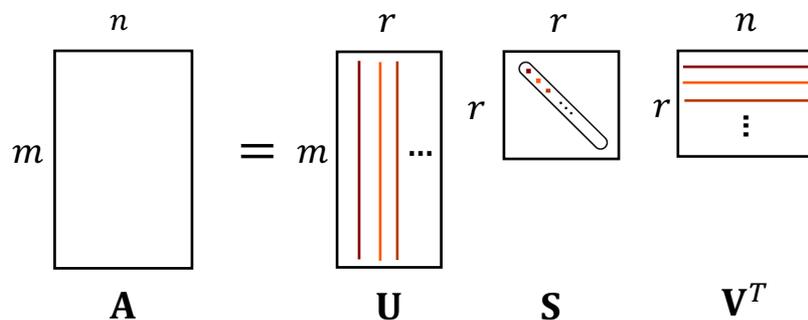
回顾: $m \geq n$ 时, s_1, \dots, s_n 为 $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}$, 即根号下 $\mathbf{A}^T \mathbf{A}$ 特征值

若恰有 $r \leq n$ 个非零特征值 (即 $\lambda_1, \dots, \lambda_r > 0, \lambda_{r+1}, \dots, \lambda_n = 0$), 则 $\mathbf{A} = \sum_{i=1}^r s_i \cdot \mathbf{u}_i \mathbf{v}_i^T$



其余列/行向量是
为了凑成正交矩阵

SVD的另一种定义



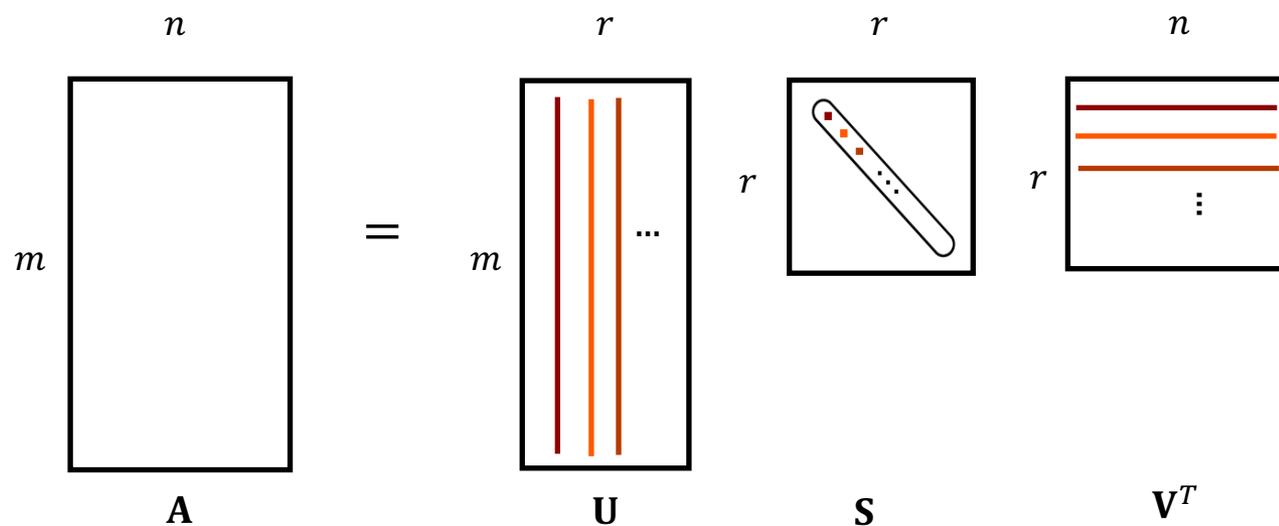
可以证明 r 为矩阵 \mathbf{A} 的秩

奇异值分解的另一种定义

奇异值分解 (Singular Value Decomposition)

给定秩为 r 的 $m \times n$ 矩阵 A ，将其分解为三个矩阵之积： $A = USV^T$ ，其中：

- (1) U 是 $m \times r$ 矩阵， r 个列向量构成标准正交向量组；
- (2) V 是 $n \times r$ 矩阵， r 个列向量构成标准正交向量组；
- (3) S 是 $r \times r$ 对角矩阵且**所有对角线上元素为正**。



奇异值分解的另一种定义

定义一

奇异值分解 (Singular Value Decomposition)

给定 $m \times n$ 矩阵 A ，将其分解为三个矩阵之积： $A = USV^T$ ，其中：

- (1) U 是 $m \times m$ 正交矩阵；方阵 矩形矩阵 方阵
- (2) V 是 $n \times n$ 正交矩阵；
- (3) S 是 $m \times n$ (矩形) 对角矩阵且所有元素非负。

定义二

奇异值分解 (Singular Value Decomposition)

给定秩为 r 的 $m \times n$ 矩阵 A ，将其分解为三个矩阵之积： $A = USV^T$ ，其中：

- (1) U 是 $m \times r$ 矩阵， r 个列向量构成标准正交向量组；矩形矩阵 方阵 矩形矩阵
- (2) V 是 $n \times r$ 矩阵， r 个列向量构成标准正交向量组；
- (3) S 是 $r \times r$ 对角矩阵且所有对角线上元素为正。

本讲小结



奇异值分解的两种定义及等价形式



奇异值分解的存在性与计算

主要参考资料

Tim Roughgarden and Gregory Valiant <CS 168 - The Modern Algorithmic Toolbox> Lecture Notes

Cameron Musco <COMPSCI 514 - Algorithms for Data Science> Slides

谢谢!

