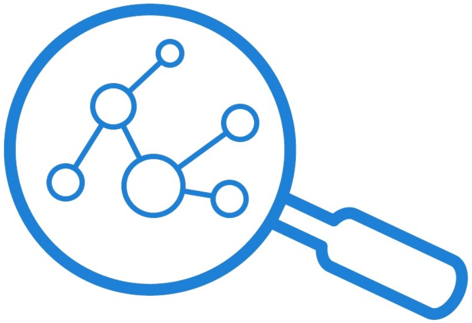
数据科学与大数据技术的数学基础

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）

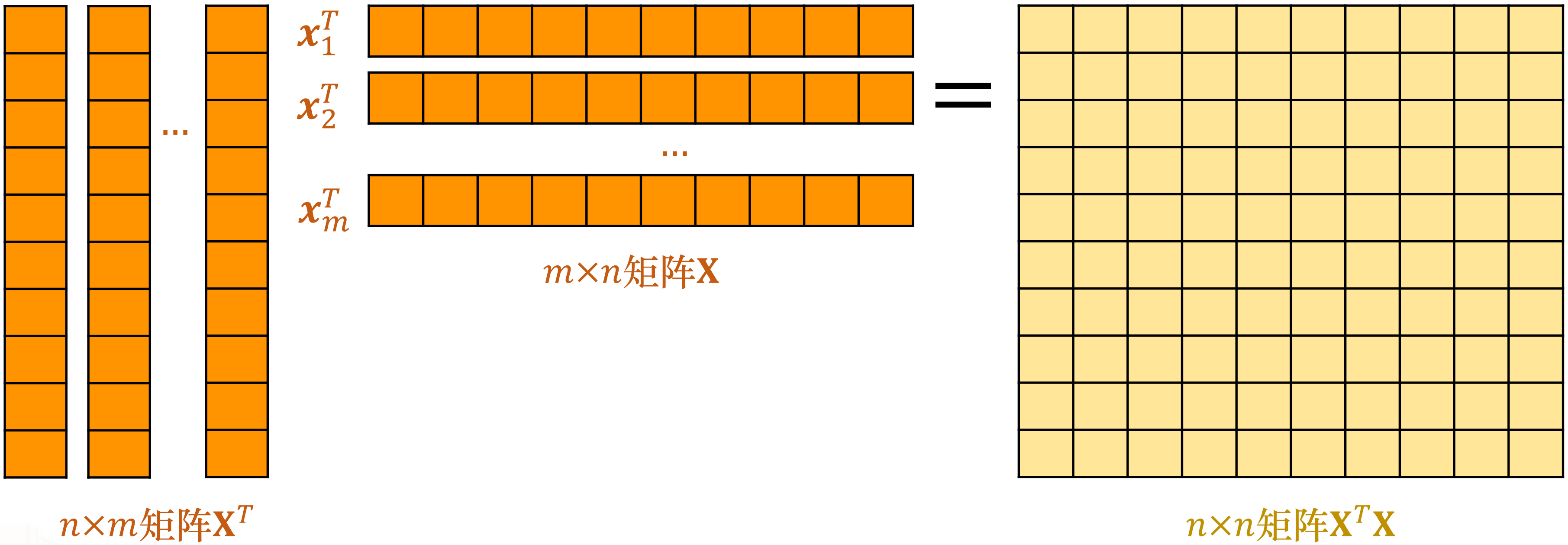
- SVD的定义
- 对比SVD与特征分解
- SVD的存在性
- SVD的计算
- SVD的等价形式
- SVD的另一种定义

奇异值分解

回顾主成分分析

给定$\boldsymbol{x}_1, \dots, \boldsymbol{x}_m \in \mathbb{R}^n$，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：$\mathbf{X}^T\mathbf{X} = \mathbf{QDQ}^T$



奇异值分解

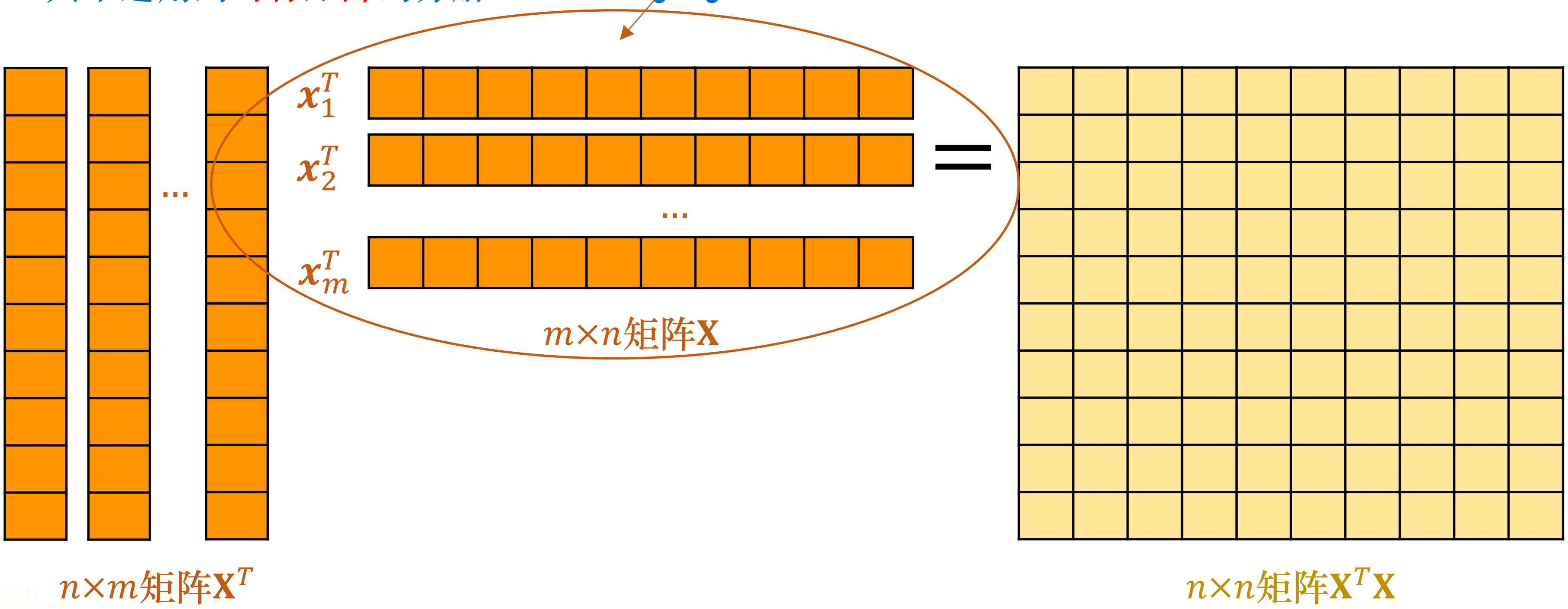
能否对非对称矩阵（包括非方阵的矩阵）进行分解？

比如直接对$m\times n$矩阵$\mathbf{X}$分解？

回顾主成分分析

给定$\boldsymbol{x}_1,\dots,\boldsymbol{x}_m\in\mathbb{R}^n$，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：$\mathbf{X}^T\mathbf{X}=\mathbf{QDQ}^T$



奇异值分解

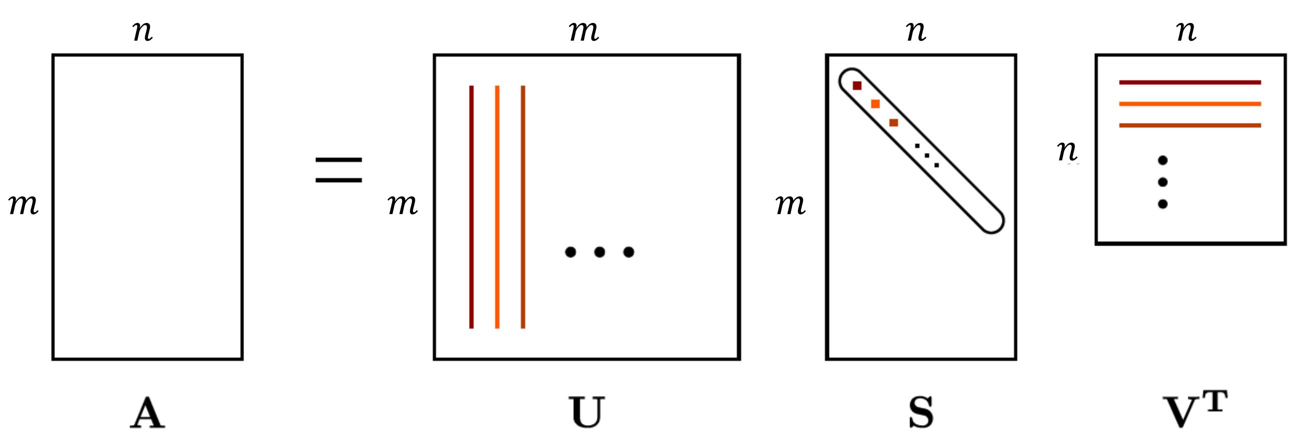
奇异值分解（Singular Value Decomposition）

给定$m \times n$矩阵**A**，将其分解为三个矩阵之积：$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T$，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。



图修改自Tim Roughgarden and Gregory Valiant <CS 168 - The Modern Algorithmic Toolbox>

奇异值分解

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

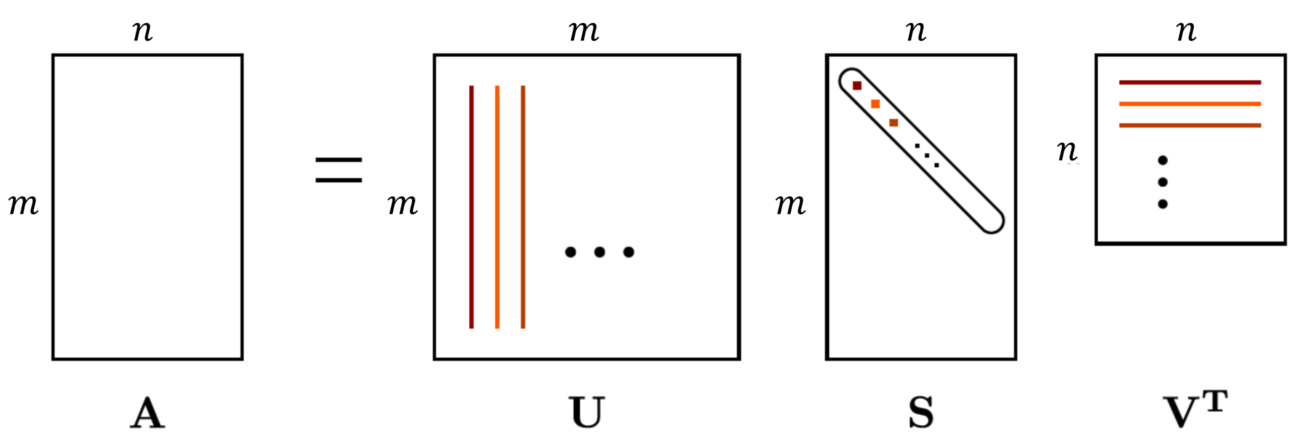
奇异值分解（Singular Value Decomposition）

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

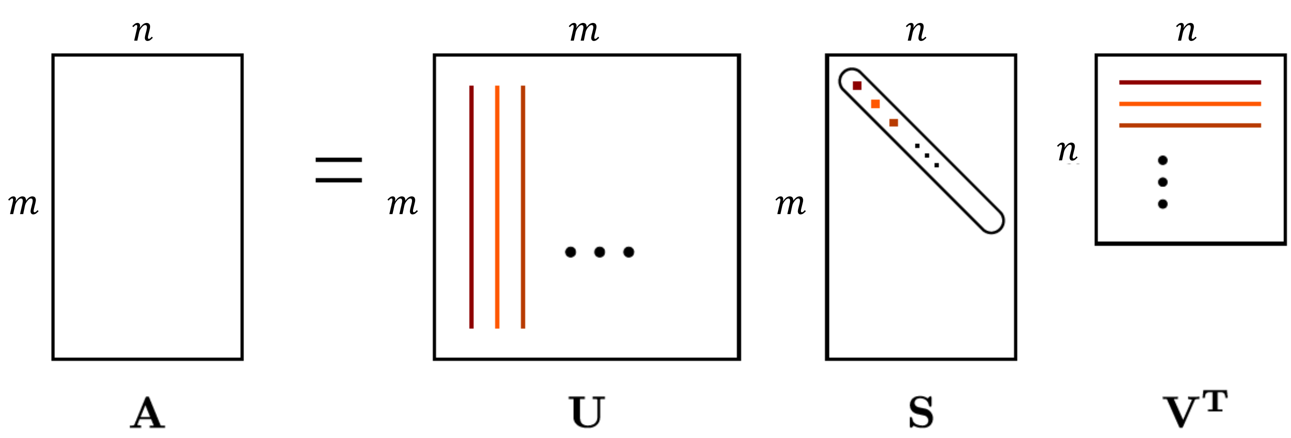
奇异值分解（Singular Value Decomposition）

给定$m \times n$矩阵**A**，将其分解为三个矩阵之积：$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T$，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、$\mathbf{V}^T$的行向量的顺序，结果不变）

奇异值分解

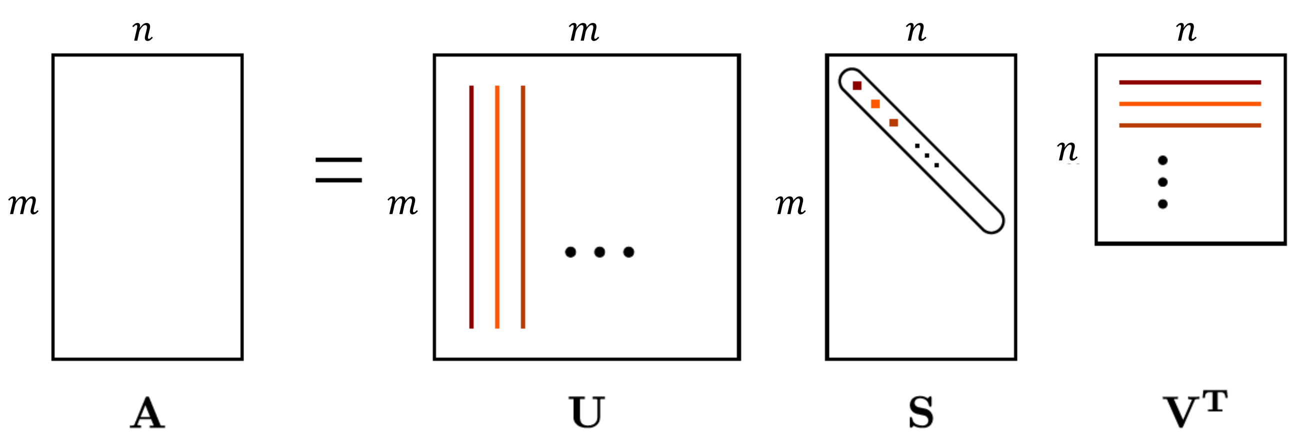
奇异值分解（Singular Value Decomposition）

给定$m \times n$矩阵**A**，将其分解为三个矩阵之积：$\mathbf{A} = \mathbf{USV}^T$，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$\mathbf{V}^T$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

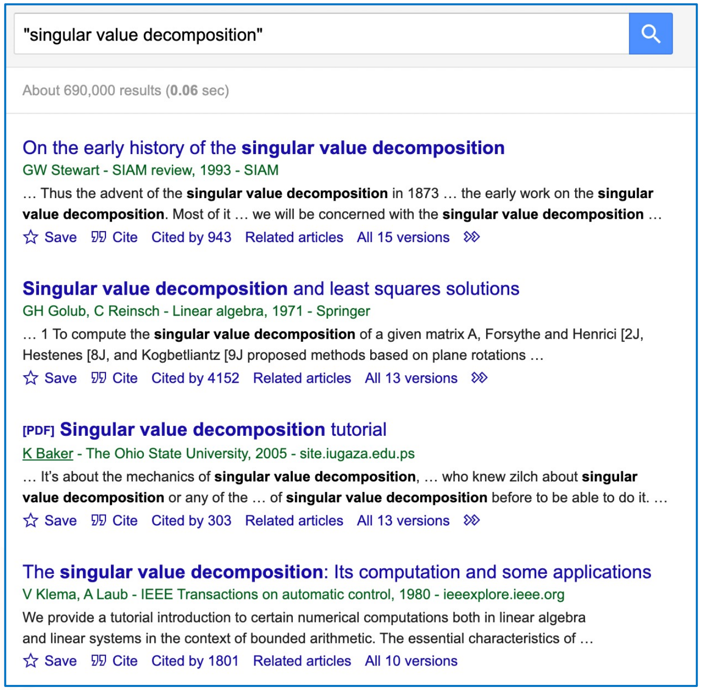
奇异值分解（Singular Value Decomposition）

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

（1）$\mathbf{U}$是$m\times m$正交矩阵；

（2）$\mathbf{V}$是$n\times n$正交矩阵；

（3）$\mathbf{S}$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”）

奇异值分解

奇异值分解（Singular Value Decomposition）

给定$m \times n$矩阵$\mathbf{A}$，将其分解为三个矩阵之积：$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T$，其中：

（1）$\mathbf{U}$是$m \times m$正交矩阵；

（2）$\mathbf{V}$是$n \times n$正交矩阵；

（3）$\mathbf{S}$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
- ...

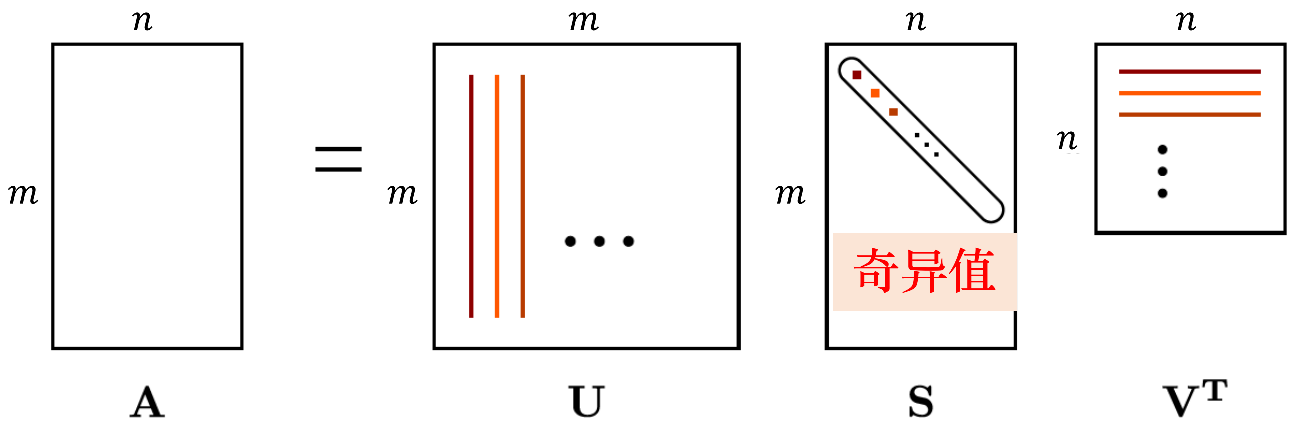
本讲内容

奇异值分解（SVD）

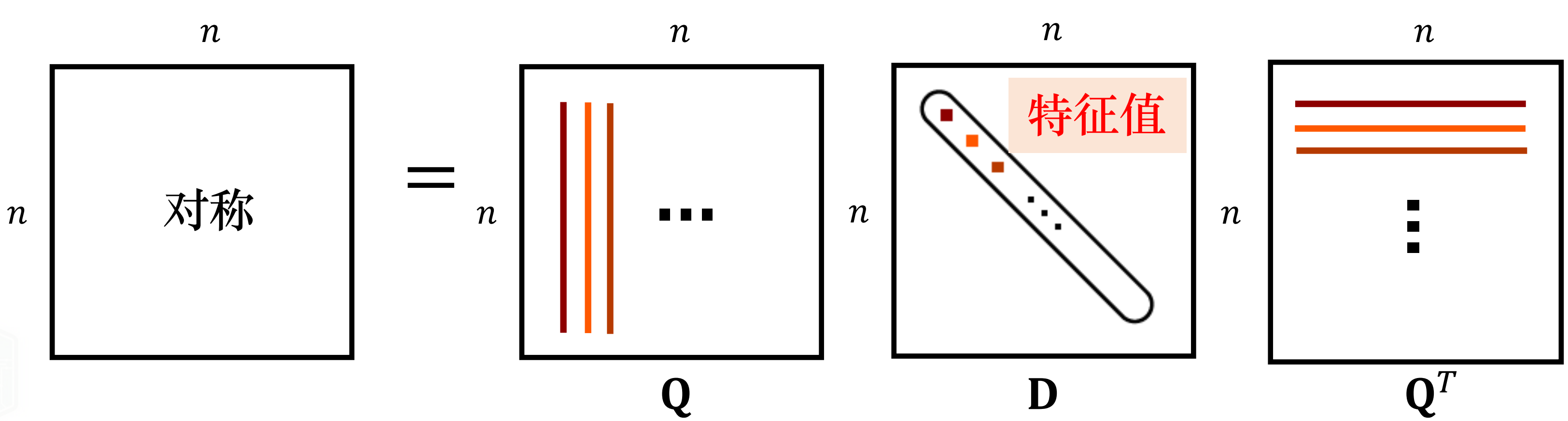
- SVD的定义
- 对比SVD与特征分解
- SVD的存在性
- SVD的计算
- SVD的等价形式
- SVD的另一种定义

奇异值分解与特征分解

$m \times n$矩阵A的奇异值分解



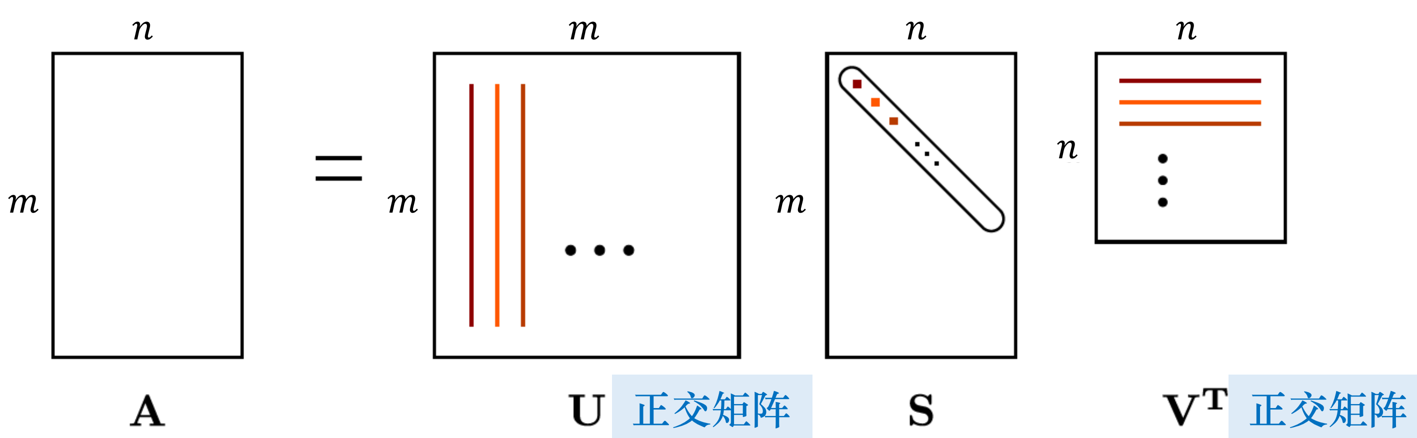
$n \times n$对称矩阵的特征分解



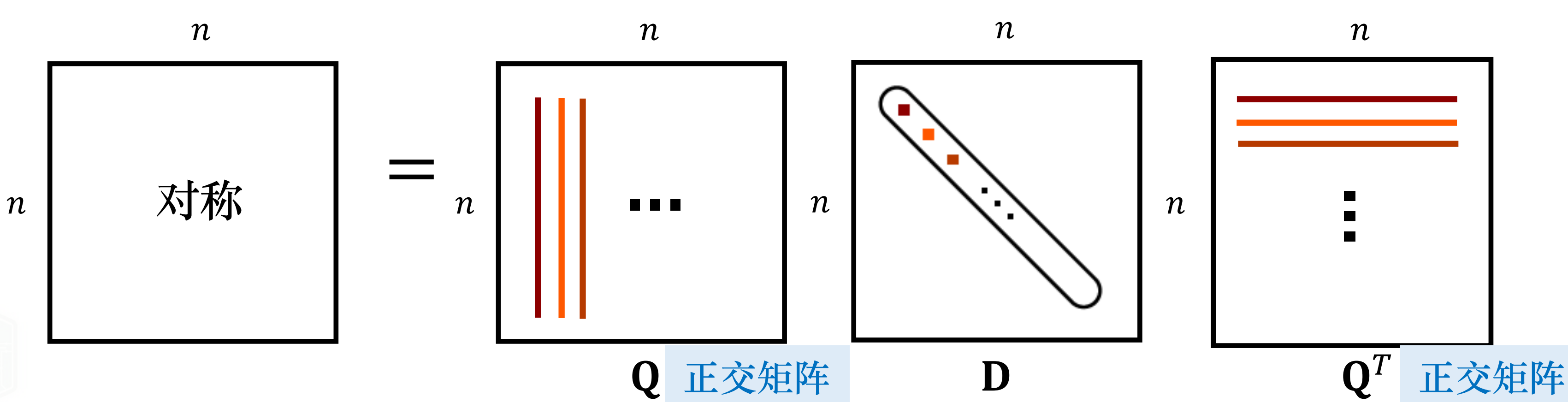
奇异值分解与特征分解

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$\mathbf{A}$，一定可将其分解成$\mathbf{USV}^T$

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

奇异值分解的存在性

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

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

若$\mathbf{A}=\mathbf{USV}^T$，则有$\mathbf{A}^T\mathbf{A}=\mathbf{VS}^T\mathbf{U}^T\mathbf{USV}^T=\mathbf{VS}^T\mathbf{SV}^T$

所以可以将$\mathbf{V}$取为$\mathbf{Q}$，$\mathbf{S}^T\mathbf{S}$取为$\mathbf{D}$

如何求矩阵$\mathbf{U}$？

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$\mathbf{USV}^T$

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

若$\mathbf{A}=\mathbf{USV}^T$，则有$\mathbf{A}^T\mathbf{A}=\mathbf{VS}^T\mathbf{U}^T\mathbf{USV}^T=\mathbf{VS}^T\mathbf{SV}^T$

所以可以将**V**取为**Q**，$\mathbf{S}^T\mathbf{S}$取为**D**

如何求矩阵**U**？

不妨设$m\geq n$

令$\boldsymbol{v}_1,\dots,\boldsymbol{v}_n$分别为$\boldsymbol{q}_1,\dots,\boldsymbol{q}_n$，即$\mathbf{A}^T\mathbf{A}$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$\mathbf{USV}^T$

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

若$\mathbf{A} = \mathbf{USV}^T$，则有$\mathbf{A}^T\mathbf{A} = \mathbf{VS}^T\mathbf{U}^T\mathbf{USV}^T = \mathbf{VS}^T\mathbf{SV}^T$

所以可以将**V**取为**Q**，$\mathbf{S}^T\mathbf{S}$取为**D**

不妨设$m \geq n$

令$\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$分别为$\boldsymbol{q}_1, \dots, \boldsymbol{q}_n$，即$\mathbf{A}^T\mathbf{A}$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$\mathbf{A}^T\mathbf{A}$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$\mathbf{USV}^T$

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

若$\mathbf{A}=\mathbf{USV}^T$，则有$\mathbf{A}^T\mathbf{A}=\mathbf{VS}^T\mathbf{U}^T\mathbf{USV}^T=\mathbf{VS}^T\mathbf{SV}^T$

所以可以将**V**取为**Q**，$\mathbf{S}^T\mathbf{S}$取为**D**

不妨设$m\geq n$

令$\boldsymbol{v}_1,\ldots,\boldsymbol{v}_n$分别为$\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n$，即$\mathbf{A}^T\mathbf{A}$的（标准正交）特征向量

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

取$\boldsymbol{u}_1,\ldots,\boldsymbol{u}_m$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$\mathbf{USV}^T$

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

若$\mathbf{A} = \mathbf{USV}^T$，则有$\mathbf{A}^T\mathbf{A} = \mathbf{VS}^T\mathbf{U}^T\mathbf{USV}^T = \mathbf{VS}^T\mathbf{SV}^T$

所以可以将**V**取为**Q**，$\mathbf{S}^T\mathbf{S}$取为**D**

不妨设$m \geq n$

令$\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$分别为$\boldsymbol{q}_1, \dots, \boldsymbol{q}_n$，即$\mathbf{A}^T\mathbf{A}$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$\mathbf{A}^T\mathbf{A}$特征值（根号下**D**对角线元素）

取$\boldsymbol{u}_1, \dots, \boldsymbol{u}_m$为 ?

由$\mathbf{A} = \mathbf{USV}^T$可得$\mathbf{AV} = \mathbf{US} = \begin{bmatrix} | & & | \\ \sqrt{\lambda_1}\boldsymbol{u}_1 & \dots & \sqrt{\lambda_n}\boldsymbol{u}_n \\ | & & | \end{bmatrix}$，即有$\mathbf{A}\boldsymbol{v}_i = \sqrt{\lambda_i}\boldsymbol{u}_i$

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$\mathbf{USV}^T$

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

若$\mathbf{A}=\mathbf{USV}^T$，则有$\mathbf{A}^T\mathbf{A}=\mathbf{VS}^T\mathbf{U}^T\mathbf{USV}^T=\mathbf{VS}^T\mathbf{SV}^T$

所以可以将**V**取为**Q**，$\mathbf{S}^T\mathbf{S}$取为**D**

不妨设$m\geq n$

令$\boldsymbol{v}_1,\dots,\boldsymbol{v}_n$分别为$\boldsymbol{q}_1,\dots,\boldsymbol{q}_n$，即$\mathbf{A}^T\mathbf{A}$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$\mathbf{A}^T\mathbf{A}$特征值（根号下**D**对角线元素）

取$\boldsymbol{u}_1,\dots,\boldsymbol{u}_n$为$\frac{\mathbf{A}\boldsymbol{v}_1}{\sqrt{\lambda_1}},\dots,\frac{\mathbf{A}\boldsymbol{v}_n}{\sqrt{\lambda_n}}$

不难证明$\boldsymbol{u}_1,\dots,\boldsymbol{u}_n$构成标准正交向量组（证明单位向量时用$\left\|\frac{\mathbf{A}\boldsymbol{v}_1}{\sqrt{\lambda_1}}\right\|=\sqrt{\frac{(\mathbf{A}\boldsymbol{v}_1)^T}{\sqrt{\lambda_1}}\frac{\mathbf{A}\boldsymbol{v}_1}{\sqrt{\lambda_1}}}$）

需要**U**是$m\times m$矩阵，即m个列向量，该如何选取$\boldsymbol{u}_{n+1},\dots,\boldsymbol{u}_m$？

奇异值分解的存在性

例 $\begin{bmatrix}1&0\\1&1\\0&1\end{bmatrix}=\begin{bmatrix}\frac{1}{\sqrt6}&\frac{1}{\sqrt2}&\frac{1}{\sqrt3}\\\frac{2}{\sqrt6}&0&-\frac{1}{\sqrt3}\\\frac{1}{\sqrt6}&-\frac{1}{\sqrt2}&\frac{1}{\sqrt3}\end{bmatrix}\begin{bmatrix}\sqrt3&0\\0&1\\0&0\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt2}&\frac{1}{\sqrt2}\\\frac{1}{\sqrt2}&-\frac{1}{\sqrt2}\end{bmatrix}$

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

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

若$\mathbf{A}=\mathbf{USV}^T$，则有$\mathbf{A}^T\mathbf{A}=\mathbf{VS}^T\mathbf{U}^T\mathbf{USV}^T=\mathbf{VS}^T\mathbf{SV}^T$

所以可以将**V**取为**Q**，$\mathbf{S}^T\mathbf{S}$取为**D**

不妨设$m\geq n$

令$\boldsymbol{v}_1,\dots,\boldsymbol{v}_n$分别为$\boldsymbol{q}_1,\dots,\boldsymbol{q}_n$，即$\mathbf{A}^T\mathbf{A}$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$\mathbf{A}^T\mathbf{A}$特征值（根号下**D**对角线元素）

取$\boldsymbol{u}_1,\dots,\boldsymbol{u}_n$为$\frac{\mathbf{A}\boldsymbol{v}_1}{\sqrt{\lambda_1}},\dots,\frac{\mathbf{A}\boldsymbol{v}_n}{\sqrt{\lambda_n}}$

不难证明$\boldsymbol{u}_1,\dots,\boldsymbol{u}_n$构成标准正交向量组（证明单位向量时用$\left\|\frac{\mathbf{A}\boldsymbol{v}_1}{\sqrt{\lambda_1}}\right\|=\sqrt{\frac{(\mathbf{A}\boldsymbol{v}_1)^T}{\sqrt{\lambda_1}}\frac{\mathbf{A}\boldsymbol{v}_1}{\sqrt{\lambda_1}}}$）

选取$\boldsymbol{u}_{n+1},\dots,\boldsymbol{u}_m$令与$\boldsymbol{u}_1,\dots,\boldsymbol{u}_n$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?

奇异值分解的存在性

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

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

若$\mathbf{A} = \mathbf{USV}^T$，则有$\mathbf{A}^T\mathbf{A} = \mathbf{VS}^T\mathbf{U}^T\mathbf{USV}^T = \mathbf{VS}^T\mathbf{SV}^T$

所以可以将$\mathbf{V}$取为$\mathbf{Q}$，$\mathbf{S}^T\mathbf{S}$取为$\mathbf{D}$

不妨设$m \geq n$

令$\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$分别为$\boldsymbol{q}_1, \dots, \boldsymbol{q}_n$，即$\mathbf{A}^T\mathbf{A}$的（标准正交）特征向量

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

取$\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$为$\frac{\mathbf{A}\boldsymbol{v}_1}{\sqrt{\lambda_1}}, \dots, \frac{\mathbf{A}\boldsymbol{v}_n}{\sqrt{\lambda_n}}$

不难证明$\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$构成标准正交向量组（证明单位向量时用$\left\|\frac{\mathbf{A}\boldsymbol{v}_1}{\sqrt{\lambda_1}}\right\| = \sqrt{\frac{(\mathbf{A}\boldsymbol{v}_1)^T}{\sqrt{\lambda_1}}\frac{\mathbf{A}\boldsymbol{v}_1}{\sqrt{\lambda_1}}}$）

选取$\boldsymbol{u}_{n+1}, \dots, \boldsymbol{u}_m$令与$\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$共同构成标准正交向量组即可。

找到用于奇异值分解的$\mathbf{U}, \mathbf{S}, \mathbf{V}$

矩阵的奇异值分解是否唯一?

不唯一， $\mathbf{A}^T\mathbf{A}$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$\mathbf{USV}^T$

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

若$\mathbf{A}=\mathbf{USV}^T$，则有$\mathbf{A}^T\mathbf{A}=\mathbf{VS}^T\mathbf{U}^T\mathbf{USV}^T=\mathbf{VS}^T\mathbf{SV}^T$

所以可以将V取为Q，$\mathbf{S}^T\mathbf{S}$取为D

不妨设$m\geq n$

令$\boldsymbol{v}_1,\ldots,\boldsymbol{v}_n$分别为$\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n$，即$\mathbf{A}^T\mathbf{A}$的（标准正交）特征向量

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

取$\boldsymbol{u}_1,\ldots,\boldsymbol{u}_n$为$\frac{\mathbf{A}\boldsymbol{v}_1}{\sqrt{\lambda_1}},\ldots,\frac{\mathbf{A}\boldsymbol{v}_n}{\sqrt{\lambda_n}}$

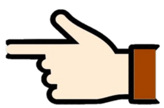
不难证明$\boldsymbol{u}_1,\ldots,\boldsymbol{u}_n$构成标准正交向量组（证明单位向量时用$\left\|\frac{\mathbf{A}\boldsymbol{v}_1}{\sqrt{\lambda_1}}\right\|=\sqrt{\frac{(\mathbf{A}\boldsymbol{v}_1)^T}{\sqrt{\lambda_1}}\frac{\mathbf{A}\boldsymbol{v}_1}{\sqrt{\lambda_1}}}$）

选取$\boldsymbol{u}_{n+1},\ldots,\boldsymbol{u}_m$令与$\boldsymbol{u}_1,\ldots,\boldsymbol{u}_n$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$\mathbf{A}$（$m \geq n$），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$s_i = \|\mathbf{A}\boldsymbol{v}_i\|$且$\boldsymbol{u}_i = \frac{\mathbf{A}\boldsymbol{v}_i}{\|\mathbf{A}\boldsymbol{v}_i\|}$

（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$s_i = \|\mathbf{A}\boldsymbol{v}_i\|$且$\boldsymbol{u}_i = \frac{\mathbf{A}\boldsymbol{v}_i}{\|\mathbf{A}\boldsymbol{v}_i\|}$
（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$\mathbf{A}^T\mathbf{A}\boldsymbol{v} = \lambda\boldsymbol{v}$，即$\boldsymbol{v}_1, \boldsymbol{v}_2$满足$v_{11} = v_{12}, v_{21} = -v_{22}$

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

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$s_i = \|\mathbf{A}\boldsymbol{v}_i\|$且$\boldsymbol{u}_i = \frac{\mathbf{A}\boldsymbol{v}_i}{\|\mathbf{A}\boldsymbol{v}_i\|}$

（5）补齐正交矩阵

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

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$\mathbf{A}^T\mathbf{A}\boldsymbol{v} = \lambda\boldsymbol{v}$，即$\boldsymbol{v}_1, \boldsymbol{v}_2$满足$v_{11} = v_{12}, v_{21} = -v_{22}$

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

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$\mathbf{A}^T\mathbf{A}$ $\quad (O(mn^2))$

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

（3）计算$\mathbf{AV}$ $\quad (O(mn^2))$

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

（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$O(k_1k_2k_3)$

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$\mathbf{A}^T\mathbf{A}$）

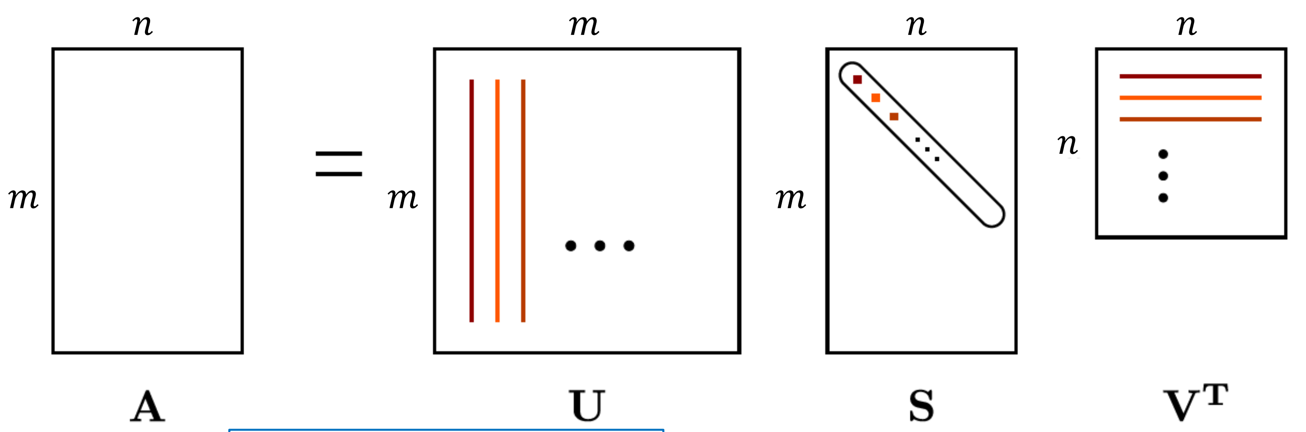
本讲内容

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



$\mathbf{A} = \mathbf{USV}^T$等价于

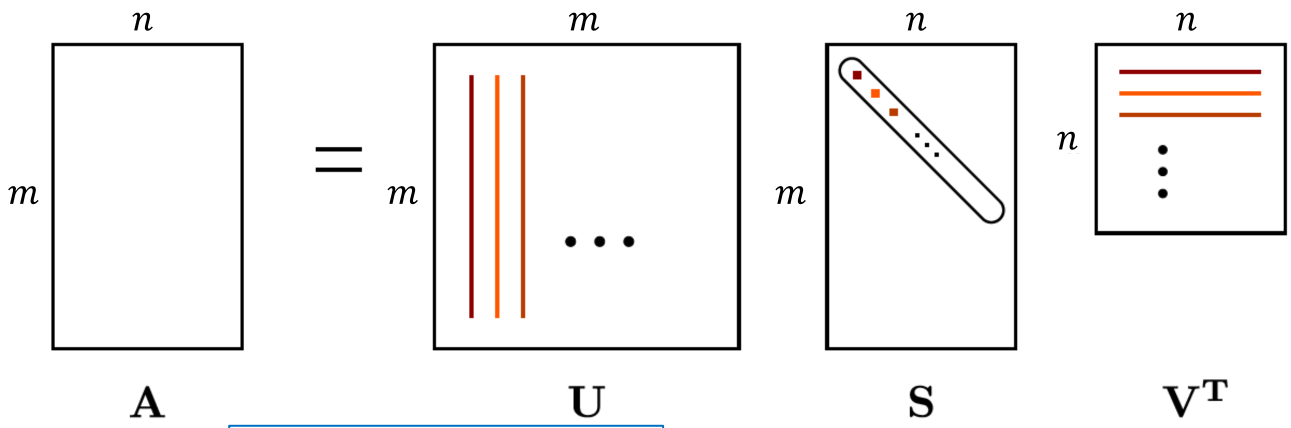
$$\boxed{\mathbf{A} = \sum_{i=1}^{\min\{m,n\}} s_i \cdot \mathbf{u}_i \mathbf{v}_i^\top}$$

$\boldsymbol{u}_i \boldsymbol{v}_i^T$是秩为1的$m \times n$矩阵

第i个奇异值　　$\mathbf{V}^T$的第i个行向量

等价形式

$m\times n$矩阵$\mathbf{A}$的奇异值分解



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

第i个奇异值

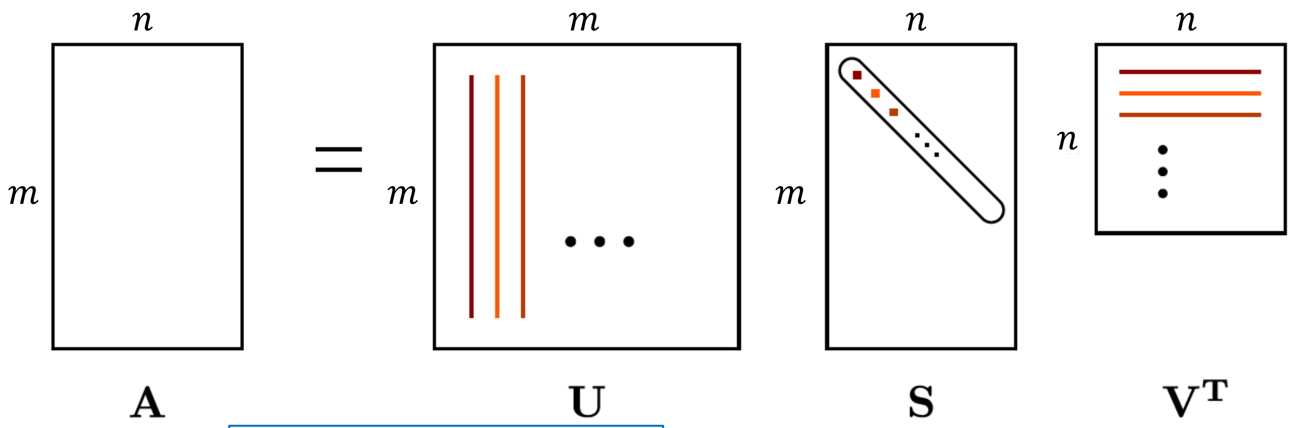
$\mathbf{V}^T$的第i个行向量

$\boldsymbol{u}_i\boldsymbol{v}_i^T$是秩为1的$m\times n$矩阵 $\begin{bmatrix} -u_{i1}\boldsymbol{v}_i^T & - \\ -u_{i2}\boldsymbol{v}_i^T & - \\ \cdots & \\ -u_{im}\boldsymbol{v}_i^T & - \end{bmatrix}$

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

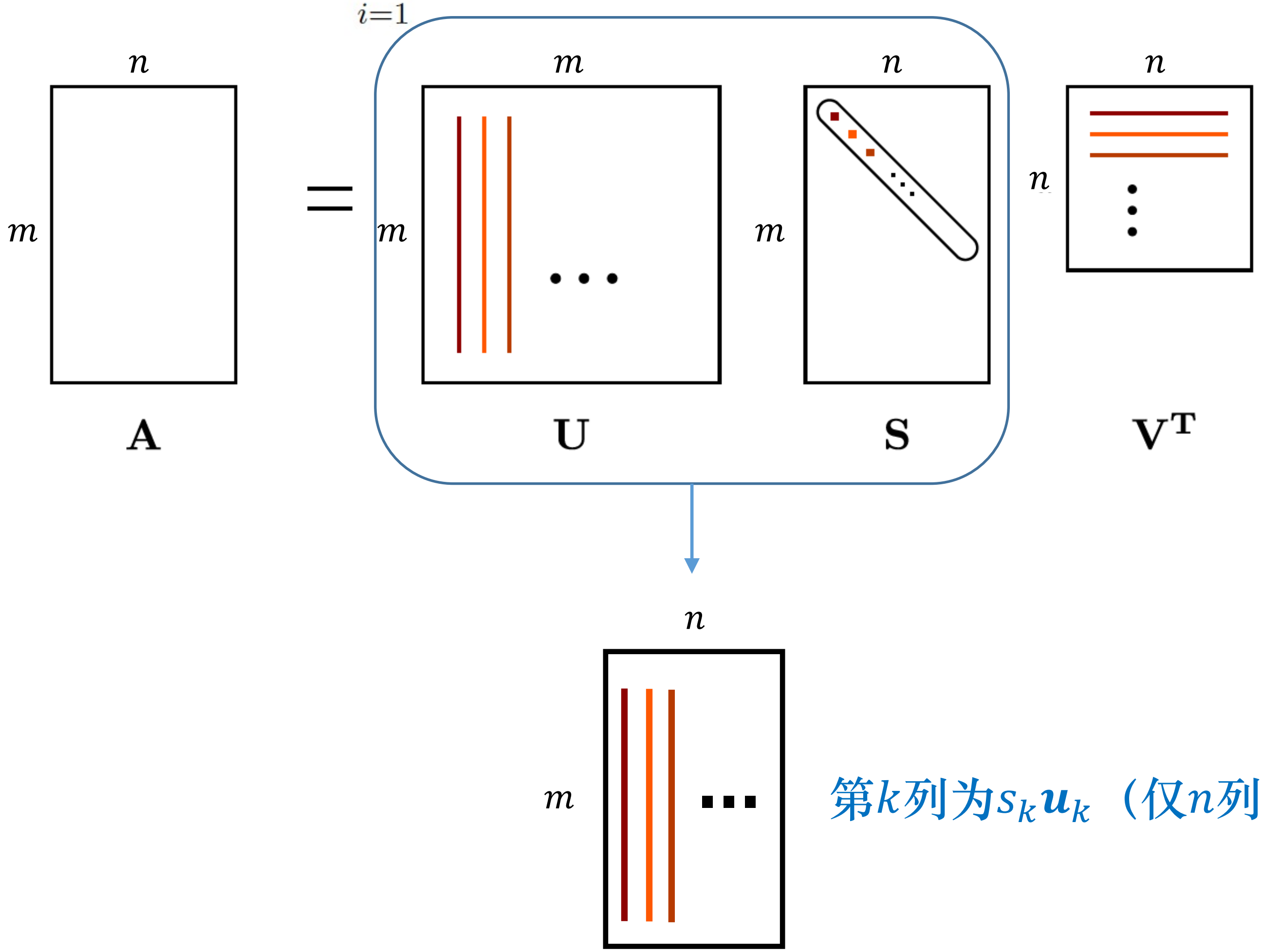


$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T$等价于 $\boxed{\mathbf{A} = \sum_{i=1}^{\min\{m,n\}} s_i \cdot \mathbf{u}_i \mathbf{v}_i^\top}$

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

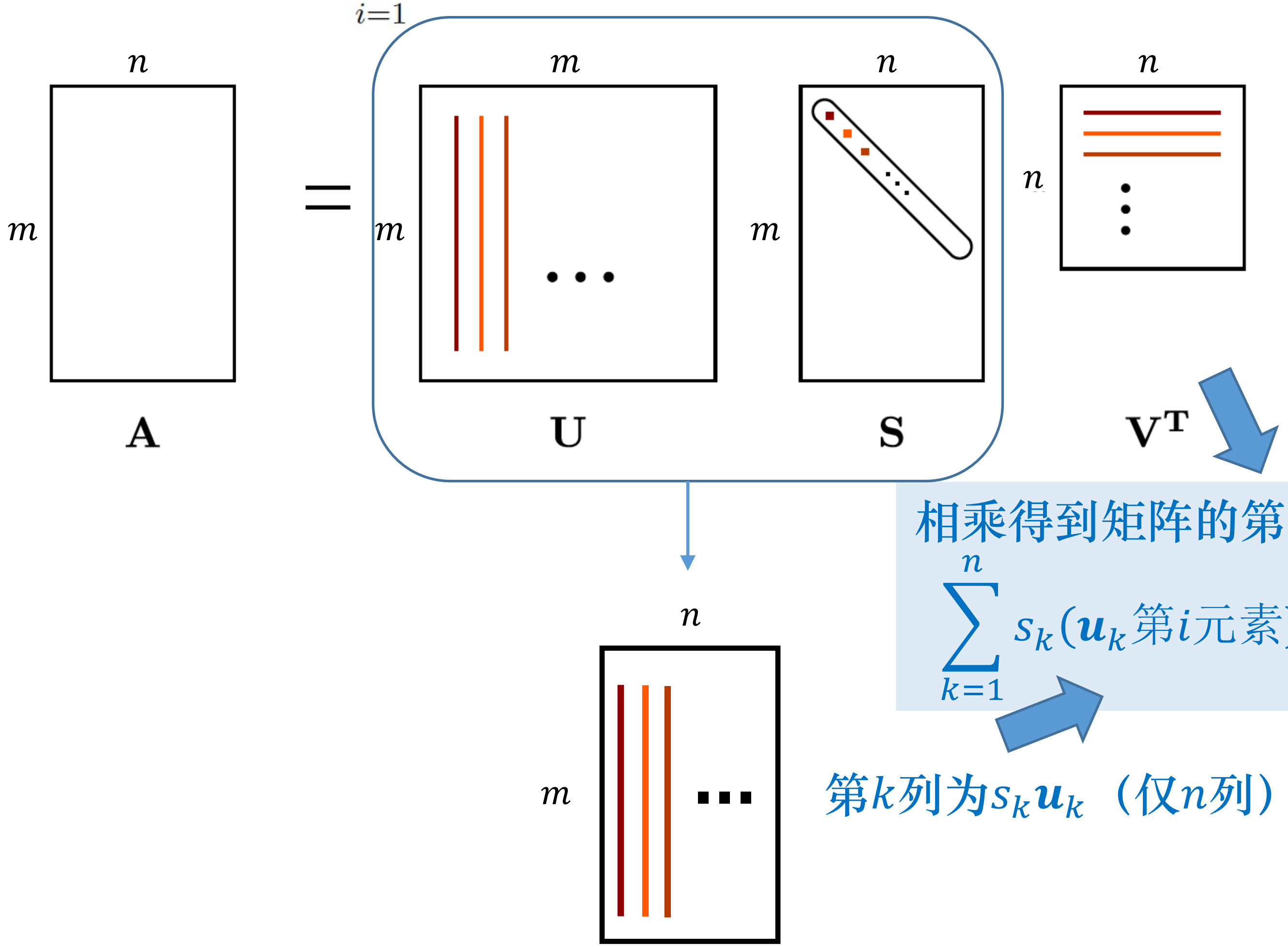
$\mathbf{A}=\mathbf{USV}^T$ 等价于 $\mathbf{A}=\sum_{i=1}^{\min\{m,n\}} s_i \cdot \mathbf{u}_i \mathbf{v}_i^\top$ 此例中 $m \geq n$



第 k 列为 $s_k \boldsymbol{u}_k$（仅 n 列）

等价形式的证明

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



相乘得到矩阵的第i行第j列为

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

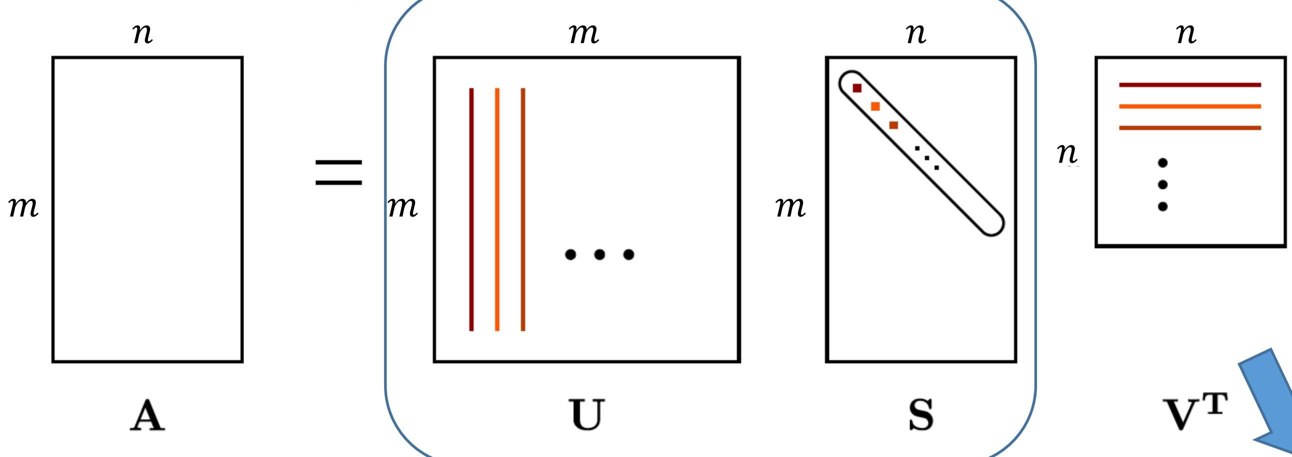
第k列为$s_k\boldsymbol{u}_k$（仅n列）

等价形式的证明

第i行第j列为

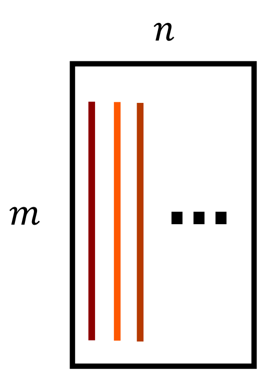
$$\sum_{k=1}^{n} s_k(\boldsymbol{u}_k\text{第}i\text{元素})(\boldsymbol{v}_k\text{第}j\text{元素})$$

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



相乘得到矩阵的第i行第j列为

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



第k列为$s_k\boldsymbol{u}_k$（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

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

奇异值分解的另一种定义

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

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

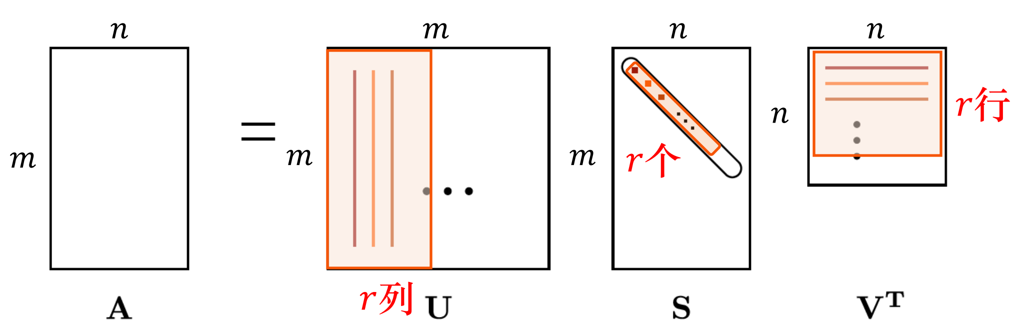
若恰有$r \leq n$个非零特征值（即$\lambda_1,\ldots,\lambda_r>0,\lambda_{r+1},\ldots,\lambda_n=0$），则 $\mathbf{A}=\sum_{i=1}^{r} s_i \cdot \mathbf{u}_i \mathbf{v}_i^\top$

奇异值分解的另一种定义

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

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 $\mathbf{A}=\sum_{i=1}^{r} s_i\cdot\mathbf{u}_i\mathbf{v}_i^\top$



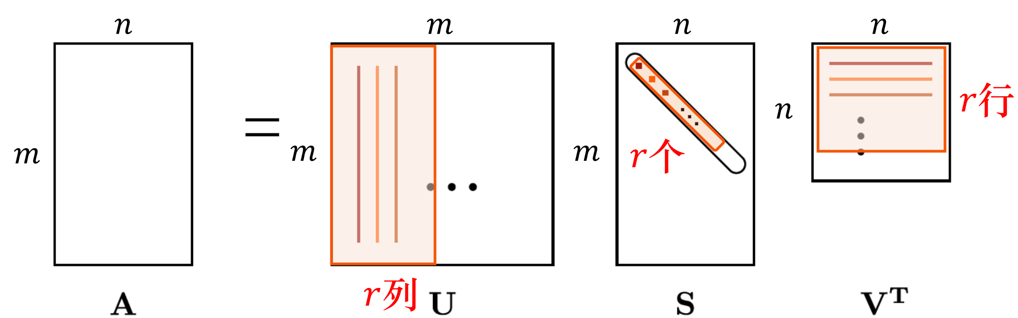
其余列/行向量是
为了凑成正交矩阵

奇异值分解的另一种定义

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

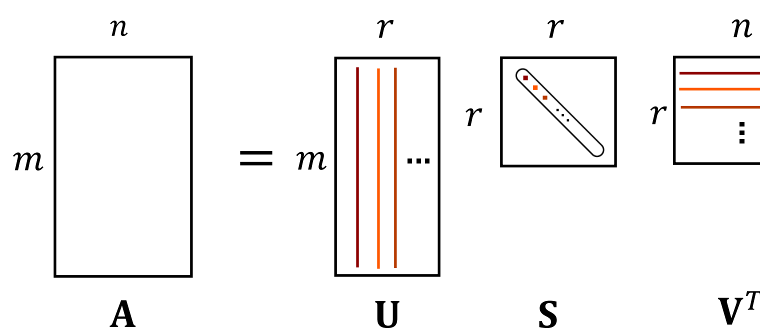
回顾：$m\geq n$时，$s_1,\ldots,s_n$为$\sqrt{\lambda_1},\ldots,\sqrt{\lambda_n}$，即根号下$\mathbf{A}^T\mathbf{A}$特征值

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



其余列/行向量是为了凑成正交矩阵

SVD的另一种定义

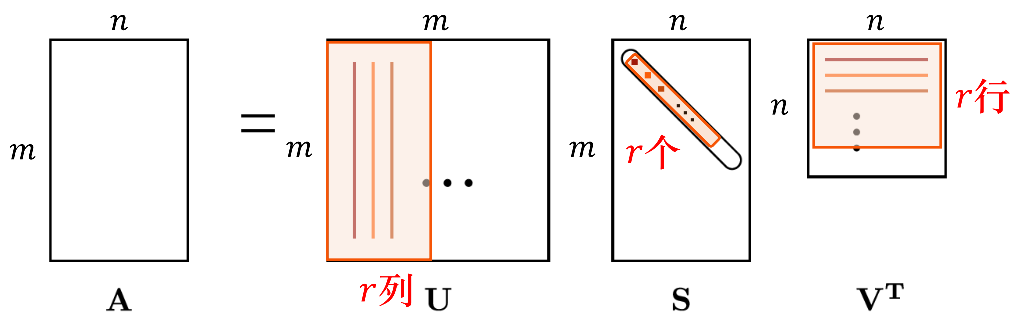


奇异值分解的另一种定义

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

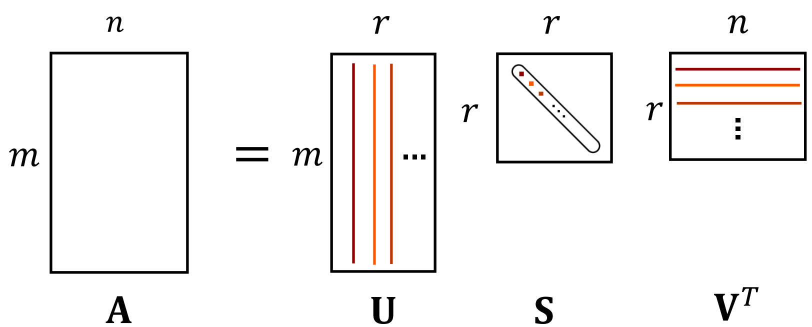
回顾：$m\geq n$时，$s_1,\ldots,s_n$为$\sqrt{\lambda_1},\ldots,\sqrt{\lambda_n}$，即根号下$\mathbf{A}^T\mathbf{A}$特征值

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



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

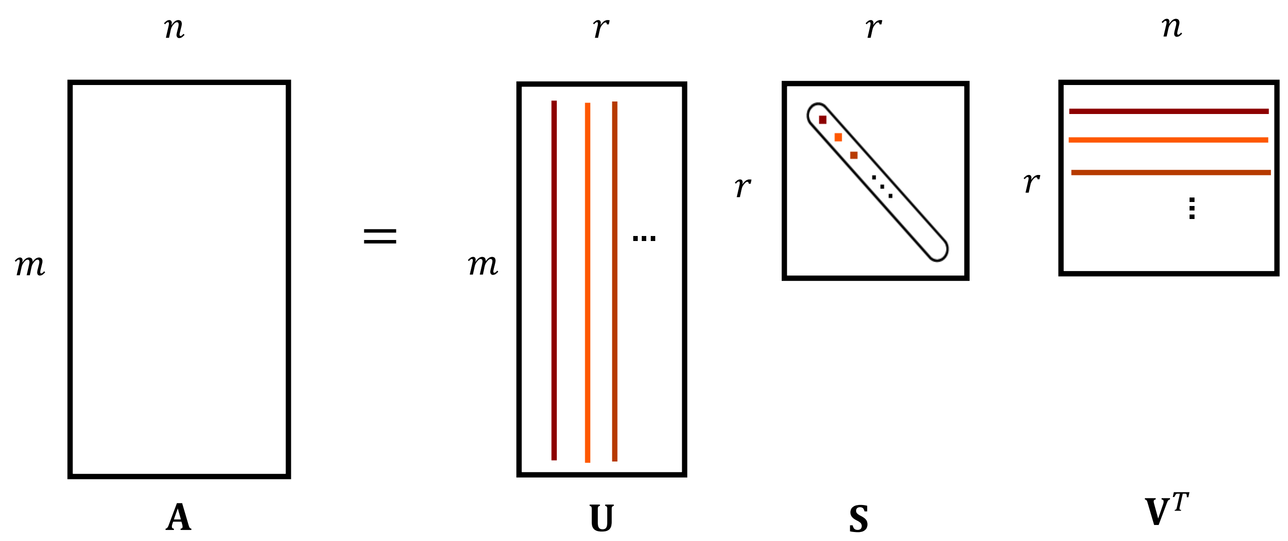
奇异值分解（Singular Value Decomposition）

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

（1）$\mathbf{U}$是$m\times r$矩阵，r个列向量构成标准正交向量组；

（2）$\mathbf{V}$是$n\times r$矩阵，r个列向量构成标准正交向量组；

（3）$\mathbf{S}$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：$\mathbf{A} = \mathbf{USV}^T$，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。

方阵　矩形矩阵　方阵

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：$\mathbf{A} = \mathbf{USV}^T$，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。

矩形矩阵　方阵　矩形矩阵

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！