《数值计算方法》课程

矩阵的特征值和奇异值

奇异值分解定理应用 (非常重要)

胡建芳

（研究方向：计算机视觉）

http://sdcs.sysu.edu.cn/content/5143

计算机学院

课程回顾

奇异值分解定理：

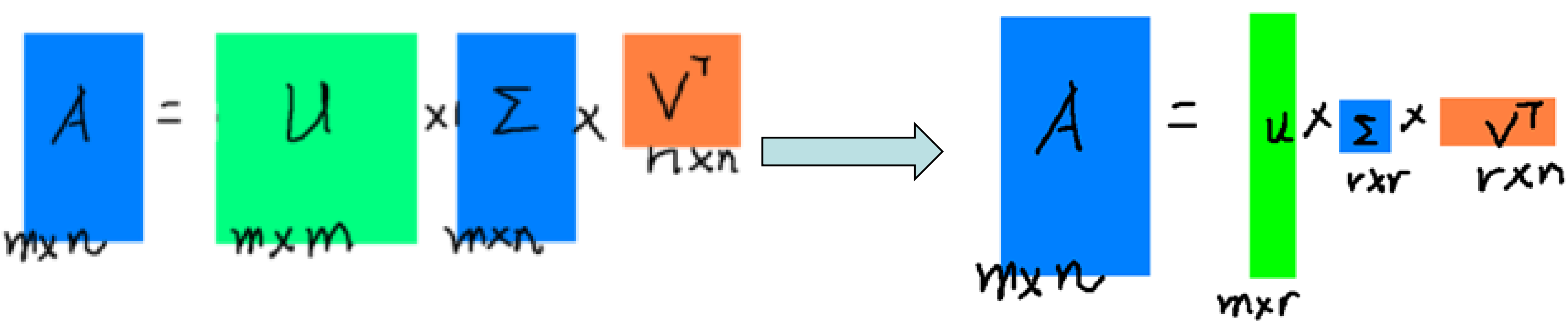
定理 设$A \in R^{m\times n}$，秩$(A)=r$,则存在m阶正交阵U和n阶正交阵V，使得

$$U^T AV = \Sigma = \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}$$

其中$\Sigma_r = diag(\sigma_1,\sigma_2,\ldots,\sigma_r)$,且$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$

称$A = U\Sigma V^T = U\begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}V^T$为矩阵$A$的奇异值分解,

$\Sigma = \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}$为矩阵$A$的奇异值矩阵。



奇异值分解应用

SVD的性质

性质 1　矩阵 $\boldsymbol{A}=\boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^{\mathrm{T}}$ 的秩是 $\boldsymbol{S}$ 中非零元素的个数.

性质 2　如果 $\boldsymbol{A}$ 是 $n\times n$ 矩阵, 那么 $|\det(\boldsymbol{A})|=s_1\cdots s_n$.

性质 3　如果 $\boldsymbol{A}$ 是可逆 $m\times m$ 逆阵, 那么 $\boldsymbol{A}^{-1}=\boldsymbol{V}\boldsymbol{S}^{-1}\boldsymbol{U}^{\mathrm{T}}$.

性质 4　$m\times n$ 矩阵 $\boldsymbol{A}$ 可以写成秩 1 矩阵之和:

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

这里 r 是 $\boldsymbol{A}$ 的秩, 而 $\boldsymbol{u}_i$ 及 $\boldsymbol{v}_i$ 分别是 $\boldsymbol{U}$ 及 $\boldsymbol{V}$ 的第 i 列.

奇异值分解应用

SVD的性质

证

$$
\begin{aligned}
\boldsymbol{A}=\boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^{\mathrm{T}}&=\boldsymbol{U}\begin{bmatrix} s_1 & & \\ & \ddots & \\ & & s_r \end{bmatrix}\boldsymbol{V}^{\mathrm{T}} \\
&=\boldsymbol{U}\left(\begin{bmatrix} s_1 & & \\ & & \\ & & \end{bmatrix}+\begin{bmatrix} & & \\ & s_2 & \\ & & \end{bmatrix}+\cdots+\begin{bmatrix} & & \\ & & \\ & & s_r \end{bmatrix}\right)\boldsymbol{V}^{\mathrm{T}} \\
&=s_1\boldsymbol{u}_1\boldsymbol{v}_1^{\mathrm{T}}+s_2\boldsymbol{u}_2\boldsymbol{v}_2^{\mathrm{T}}+\cdots+s_r\boldsymbol{u}_r\boldsymbol{v}_r^{\mathrm{T}}.
\end{aligned}
$$

性质 4 是 SVD 的低秩近似性质. 通过保留 (12.31) 式的前面 p 项就提供了对秩 $p \leqslant r$ 的 $\boldsymbol{A}$ 的最佳最小二乘近似.

奇异值分解应用

SVD的性质

例 12.8 求矩阵 $\begin{bmatrix} 0 & 1 \\ 1 & \frac{3}{2} \end{bmatrix}$ 的最佳秩 1 近似.

写出 (12.31) 就得到

$$\begin{aligned}
\begin{bmatrix} 0 & 1 \\ 1 & \frac{3}{2} \end{bmatrix} &= \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} \left(\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \right) \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \end{bmatrix} \\
&= 2 \begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix} + \frac{1}{2} \begin{bmatrix} -\frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \end{bmatrix} \\
&= \begin{bmatrix} \frac{2}{5} & \frac{4}{5} \\ \frac{4}{5} & \frac{8}{5} \end{bmatrix} + \begin{bmatrix} -\frac{2}{5} & \frac{1}{5} \\ \frac{1}{5} & -\frac{1}{10} \end{bmatrix}.
\end{aligned}$$

奇异值分解应用

数据降维

SVD 给出了一个直接了当的方法执行降维. 把数据向量考虑为 $m \times n$ 矩阵 $\boldsymbol{A} = (\boldsymbol{a}_1|\cdots|\boldsymbol{a}_n)$ 的列, 并且计算奇异值分解 $\boldsymbol{A} = \boldsymbol{USV}^{\mathrm{T}}$, 设 $\boldsymbol{e}_j$ 表示第 j 个基本基向量 (即第 j 个分量是 1, 其他分量全为 0). 于是 $\boldsymbol{Ae}_j = \boldsymbol{a}_j$. 利用性质 4 的秩 p 近似

$$\boldsymbol{A} \approx \boldsymbol{A}_p = \sum_{i=1}^{p} s_i \boldsymbol{u}_i \boldsymbol{v}_i^{\mathrm{T}},$$

我们可以通过

$$\boldsymbol{a}_j = \boldsymbol{Ae}_j \approx \boldsymbol{A}_p \boldsymbol{e}_j \tag{13.33}$$

把 $\boldsymbol{a}_j$ 投影到由 $\boldsymbol{U}$ 的列 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_p$ 所张成的 p 维空间中, 因为把一个矩阵乘以 $\boldsymbol{e}_j$ 恰好是把它的第 j 列挑选出来, 我们能更有效地叙述求解如下:

奇异值分解应用

数据降维

例 12.9 求拟合数据向量 $[3,2],[2,4],[-2,-1],[-3,-5]$ 的最佳一维子空间.

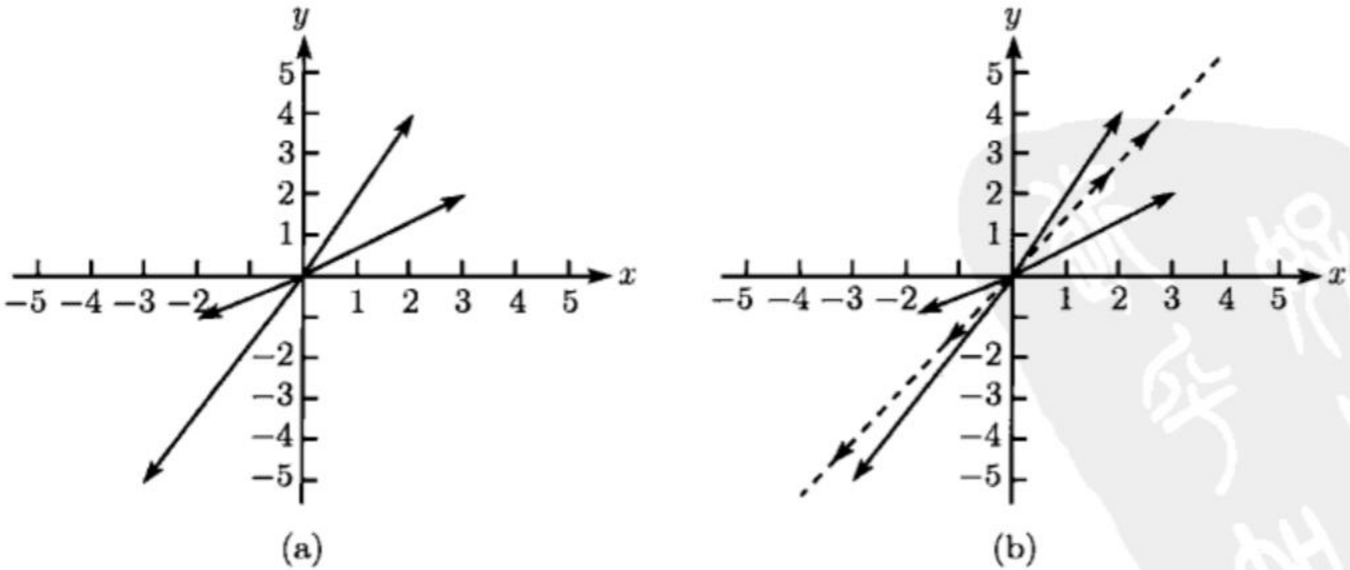


图 12-4 用 SVD 降维. (a) 4 个数据向量被投影到最佳一维子空间. (b) 虚线表示最佳子空间. 箭头表明到子空间的正交投影

奇异值分解应用

数据降维

例 12.9 求拟合数据向量 $[3,2],[2,4],[-2,-1],[-3,-5]$ 的最佳一维子空间.

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

$$\begin{bmatrix}0.588\,6 & -0.808\,4\\ 0.808\,4 & 0.588\,6\end{bmatrix}\begin{bmatrix}8.280\,9 & 0 & 0 & 0\\ 0 & 1.851\,2 & 0 & 0\end{bmatrix}$$
$$\begin{bmatrix}0.408\,5 & 0.532\,7 & -0.239\,8 & -0.701\,4\\ -0.674\,1 & 0.398\,5 & 0.555\,4 & -0.279\,8\\ 0.574\,3 & -0.189\,2 & 0.792\,4 & -0.080\,1\\ 0.221\,2 & 0.722\,3 & 0.078\,0 & 0.650\,7\end{bmatrix}.$$

$$S_1=\begin{bmatrix}8.280\,9 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}$$

$$A_1=\begin{bmatrix}1.991\,2 & 2.596\,4 & -1.168\,9 & -3.418\,8\\ 2.734\,6 & 3.565\,7 & -1.605\,2 & -4.695\end{bmatrix}$$

奇异值分解应用

数据压缩

依旧从奇异值分解的表达式入手， $A = U\Sigma V^T$ ，我们展开成完整的矩阵形式：

$$A = \begin{bmatrix} u_1 & u_2 & \dots & u_m \end{bmatrix} \begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \dots & & \\ & & & \sigma_r & \\ & & & & \\ & & & & \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \\ \dots \\ v_n^T \end{bmatrix}$$

，这里

$r \leq n < m$ 。

这里展开就得到了：

$$A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \sigma_3 u_3 v_3^T + \dots + \sigma_r u_r v_r^T$$

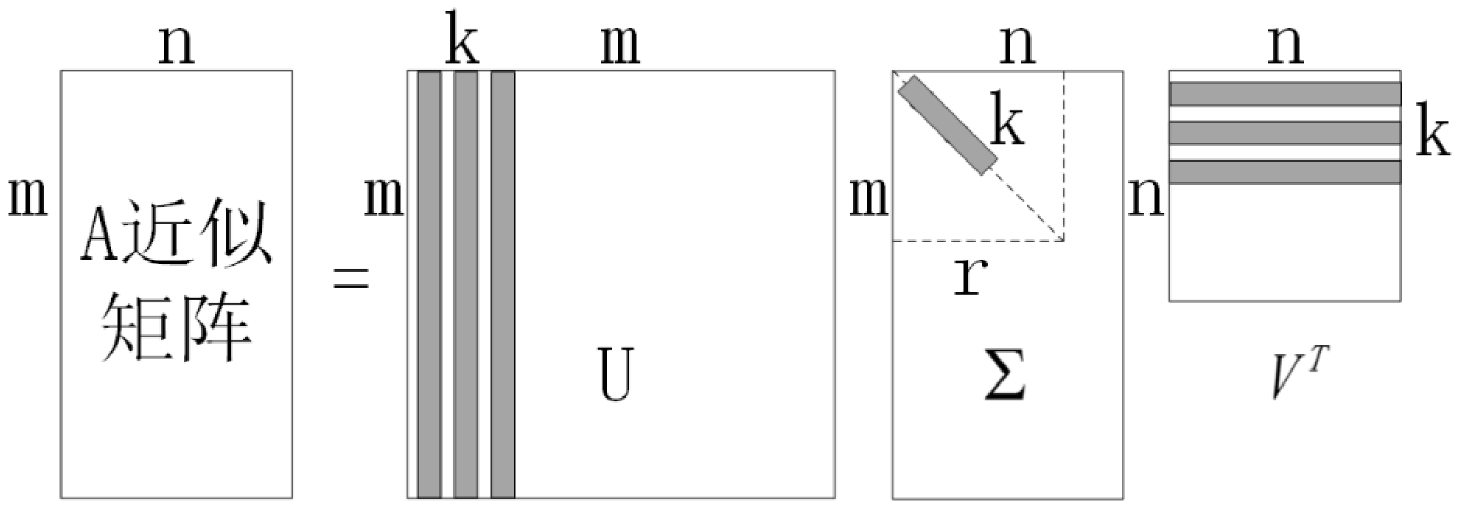
奇异值分解应用

数据压缩

行近似：

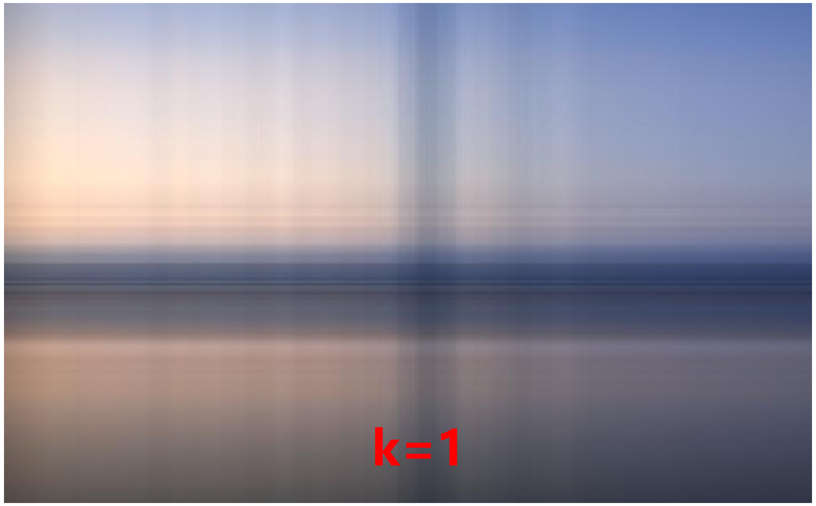
$$A \Rightarrow \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \sigma_3 u_3 v_3^T + \ldots + \sigma_k u_k v_k^T \text{ 。}$$

原理示意图如图1所示：



奇异值分解应用

- **数据压缩** http://www.johnmyleswhite.com/notebook/2009/12/17/image-compression-with-the-svd-in-r/



奇异值分解应用

- **数据压缩**

  压缩是因为样本的相关性较大，冗余成分较多

  对于多张图像，怎么去压缩?

奇异值分解应用

- **PCA 主成分分析**
  自主学习。

奇异值分解大作业

主成分分析（PCA）实验

一）：人脸数据库

每个人拍摄10张照片（**关掉美颜**），不同照片拍摄角度轻微不同，拍摄角度可以参考以下照片。

二）：实验

选择上述一个数据集进行主成分分析（PCA）实验。PCA算法
自行学习。实验结果包含如下内容：

1. 用PCA降维（压缩）后，所得到的人脸图片与原始人脸图片对比。
2. 结合最近邻分类器，测试PCA降维算法的识别率。

2-3人一小组，1个月内完成。

THE END