

线性代数5

矩阵分解

从整数分解谈矩阵特征分解

- 分解：分解为多个组成部分，找到表达方式不变的通用属性，帮助更好理解。

从整数分解谈矩阵特征分解

- 分解：分解为多个组成部分，找到表达方式不变的通用属性，帮助更好理解。

$$12 = 2 \times 2 \times 3$$

从整数分解谈矩阵分解

- 分解：分解为多个组成部分，找到表达方式不变的通用属性，帮助更好理解。

$$12 = 2 \times 2 \times 3$$

- 12不能被5整除
- 12的倍数可以被3整除

➤ 特征值分解

➤ 奇异值分解

特征值分解

- 如果说一个非零向量 v 是方阵 A 的特征向量，则一定可以表示成下面的形式：

$$Av = \lambda v$$

- λ 被称为特征向量 v 对应的特征值

特征值分解

- 如果说一个非零向量 v 是方阵 A 的特征向量，则一定可以表示成下面的形式：

$$Av = \lambda v$$

- λ 被称为特征向量 v 对应的特征值
- 性质：
 - 非零性
 - v 是特征向量，则 sv 也是特征向量

特征值分解

- 特征值分解是将一个矩阵分解成下面的形式：

$$A = V \text{diag}(\lambda) V^{-1}$$

- V 是矩阵 A 的线性无关的特征向量组成的矩阵

$$V = \{v^{(1)}, v^{(2)}, \dots, v^{(n)}\}$$

- Σ 是一个对角阵，对角线上元素就是一个特征值 $\{v^{(1)}, v^{(2)}, \dots, v^{(n)}\}$ 对应的特征向量

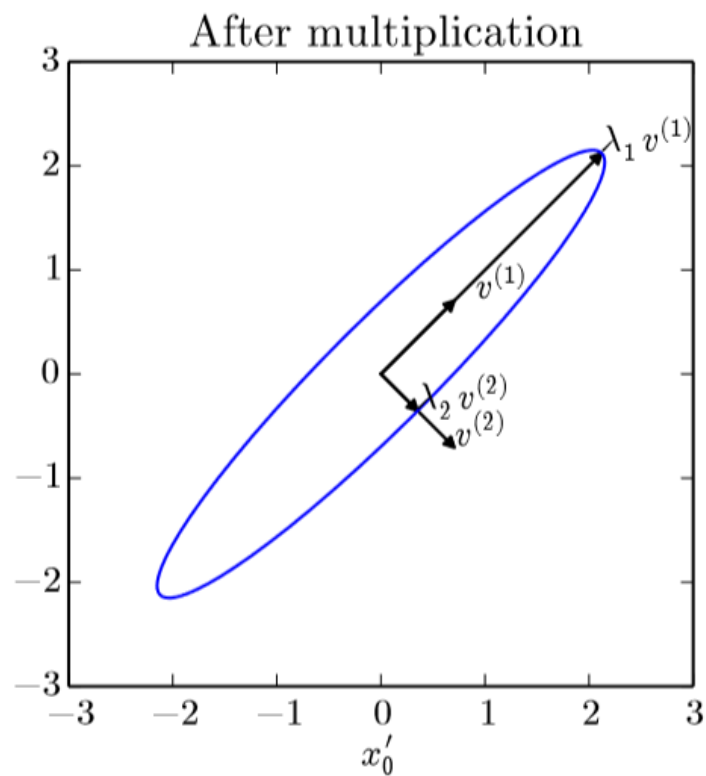
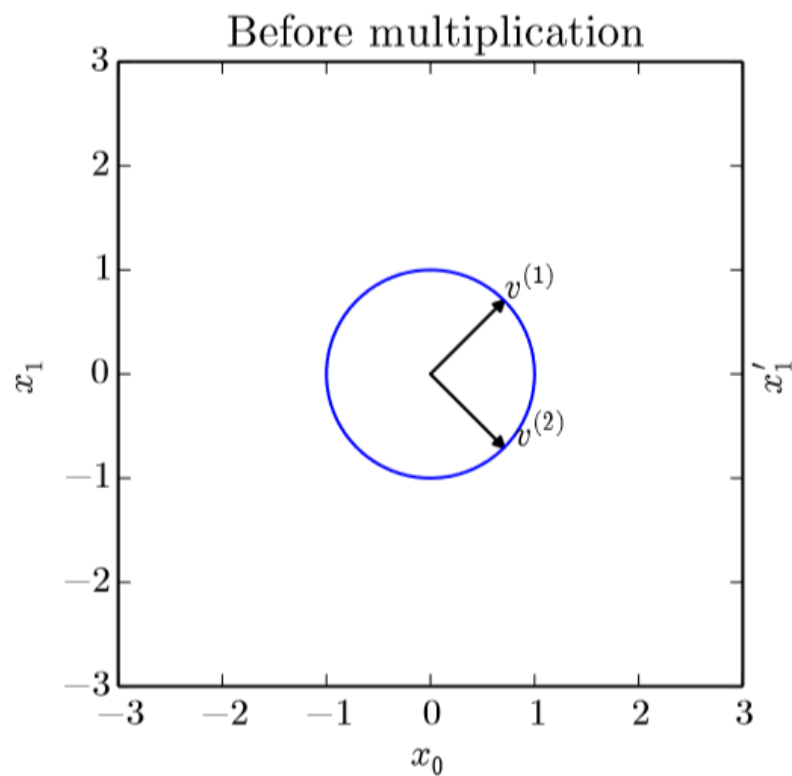
$$\{\lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(n)}\}$$

特征值分解

特征值分解

- 特征值分解的存在性
 - 不是每一个矩阵都可以特征分解
 - 实对称矩阵都可以特征分解
 - $A = Q\Lambda Q^T$
 - Q 是 A 的特征向量组成的正交矩阵， Λ 是对角矩阵
- 特征值分解的唯一性
 - 可能并不唯一

特征值分解



特征值分解

- 正定矩阵
- 半正定矩阵
- 负定矩阵
- 半负定矩阵

$$\forall \mathbf{x}, \mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0$$

$$\mathbf{x}^\top \mathbf{A} \mathbf{x} = 0 \Rightarrow \mathbf{x} = \mathbf{0}$$

特征值分解

- 特征值分解是一个提取矩阵特征很不错的方法
- 但是它只是对**方阵**而言的，在现实的世界中，我们看到的大部分矩阵都不是方阵
- 比如说有N个学生，每个学生有M科成绩，这样形成的一个 **$N * M$ 的矩阵**就不可能是方阵

我们怎样才能描述这样**普通的矩阵**呢的重要特征呢？

奇异值分解可以用来干这个事情，奇异值分解是一个能适用于**任意矩阵**的一种分解的方法

特征值分解

- 应用广泛
 - 每个实数矩阵都有一个奇异值分解，但不一定都有特征分解
 - 非方阵的矩阵没有特征分解，这时只能使用奇异值分解

奇异值分解

分解形式： $A = UDV^T$.

- A 是一个 $N \times M$ 的矩阵
- U 是一个 $M \times M$ 的方阵
 - U 的列向量称为左奇异向量
- Σ 是一个 $N \times M$ 的矩阵
 - 除了对角线的元素都是0，对角线上的元素称为奇异值
- V^T 是一个 $N \times N$ 的矩阵
 - V^T 列向量称为右奇异向量)

A diagram illustrating the SVD equation $A = U \Sigma V^T$. The matrix A is represented by a blue rectangle with dimensions $m \times n$ written below it. This is followed by an equals sign. Then, matrix U is shown in a green rectangle with dimensions $m \times m$ below it. This is followed by a multiplication sign \times . Then, matrix Σ is shown in a blue rectangle with dimensions $m \times n$ below it. This is followed by another multiplication sign \times . Finally, matrix V^T is shown in an orange rectangle with dimensions $n \times n$ below it.

奇异值分解

- 那么奇异值和特征值是怎么对应起来的呢？
- 将一个矩阵A的转置 乘以 A，并将会得到一个方程：

$$(A^T A)v_i = \lambda_i v_i$$

- 我们利用这个方阵求特征值 λ_i 以及特征向量组 V，就是我们上面的右奇异向量。此外我们还可以得到：

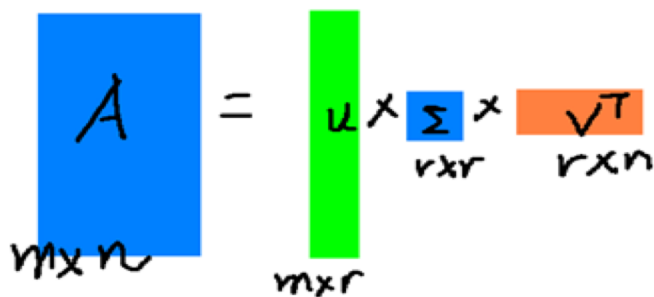
$$\sigma_i = \sqrt{\lambda_i}$$
$$u_i = \frac{1}{\sigma_i} A v_i$$

- 这里的 σ 就是奇异值，u就是上面说的左奇异向量。

奇异值分解

奇异值 σ 跟特征值类似，在矩阵 Σ 中也是从大到小排列，而且 σ 的减少特别的快，在很多情况下，前10%甚至1%的奇异值的和就占了全部的奇异值之和的99%以上了。也就是说，我们也可以用前 r （ r 远小于 m 、 n ）个的奇异值来近似描述矩阵，即部分奇异值分解：

$$A_{m \times n} = U_{m \times m} \Sigma_{m \times n} V_{n \times n}^T$$



The diagram illustrates the partial singular value decomposition (SVD) of a matrix A . It shows a blue square representing matrix A with dimensions $m \times n$ written below it. This is followed by an equals sign and three matrices multiplied together. The first matrix is a green vertical rectangle representing U with dimensions $m \times r$ written below it. This is followed by a blue square representing Σ with dimensions $r \times r$ written below it. The final matrix is an orange rectangle representing V^T with dimensions $r \times n$ written below it. The matrices are connected by multiplication symbols (\times).

右边的三个矩阵相乘的结果将会是一个接近于 A 的矩阵，在这儿， r 越接近于 n ，则相乘的结果越接近于 A 。

奇异值分解

$$A_{m \times n} = U_{m \times m} \Sigma_{m \times n} V_{n \times n}^T$$



其实整个求解 SVD 的过程就是求解这 3 个矩阵的过程，而求解这 3 个矩阵的过程就是求解特征值和特征向量的过程，问题就在于 **求谁的特征值和特征向量**。

- U 的列由 AA^T 的单位化过的特征向量构成
- V 的列由 $A^T A$ 的单位化过的特征向量构成
- Σ 的对角元素来源于 AA^T 或 $A^T A$ 的特征值的平方根，并且是按从大到小的顺序排列的



1. 求 AA^T 的特征值和特征向量，用单位化的特征向量构成 U
2. 求 $A^T A$ 的特征值和特征向量，用单位化的特征向量构成 V
3. 将 AA^T 或者 $A^T A$ 的特征值求平方根，然后构成 Σ

奇异值分解

$$A = \begin{pmatrix} 2 & 4 \\ 1 & 3 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \Rightarrow AA^T = \begin{pmatrix} 20 & 14 & 0 & 0 \\ 14 & 10 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \Rightarrow \begin{aligned} AA^T x &= \lambda x \\ (AA^T - \lambda E)x &= 0 \end{aligned}$$

$$\Rightarrow \begin{vmatrix} 20 - \lambda & 14 & 0 & 0 \\ 14 & 10 - \lambda & 0 & 0 \\ 0 & 0 & -\lambda & 0 \\ 0 & 0 & 0 & -\lambda \end{vmatrix} = 0 \Rightarrow \lambda_1 \approx 29.86606875, \lambda_2 \approx 0.13393125, \lambda_3 = \lambda_4 = 0$$

因为有4个特征，特征向量为4*4的矩阵

$$\Rightarrow x = \begin{pmatrix} 0.8174 & -0.5766 & 0 & 0 \\ 0.576 & 0.8174 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$A_{m \times n} = U_{m \times m} \Sigma_{m \times n} V_{n \times n}^T$$

这个就是U

奇异值分解

同样的过程求解 $A^T A$ 的特征值和特征向量，求得 $\lambda_1 \approx 0.13393125$, $\lambda_2 \approx 29.86606875$ ，将特征值降序排列后对应的单位化过的特征向量为

$$\begin{pmatrix} 0.40455358 & -0.9145143 \\ 0.9145143 & 0.40455358 \end{pmatrix}$$

$$A_{m \times n} = U_{m \times m} \Sigma_{m \times n} V_{n \times n}^T$$

这个就是V

奇异值分解

而矩阵 Σ 根据上面说的为特征值的平方根构成的对角矩阵

$$\begin{pmatrix} 5.4649857 & 0 \\ 0 & 0.36596619 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

奇异值分解

到此，SVD 分解就结束了，原来的矩阵 A 就被分解成了 3 个矩阵的乘积。

$$A_{4 \times 2} = U_{4 \times 4} \Sigma_{4 \times 2} V_{2 \times 2}^T$$
$$= \begin{pmatrix} 0.81741556 & -0.57604844 & 0 & 0 \\ 0.57604844 & 0.81741556 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 4 \\ 1 & 3 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 5.4649857 & 0 \\ 0 & 0.36596619 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0.40455358 & -0.9145143 \\ 0.9145143 & 0.40455358 \end{pmatrix}^T$$