一文解释 矩阵特征分解

撒旦-cc
已识乾坤大，犹怜草木青

> 特征分解：eigendecomposition 特征向量：eigenvector 特征值：eigenvalue

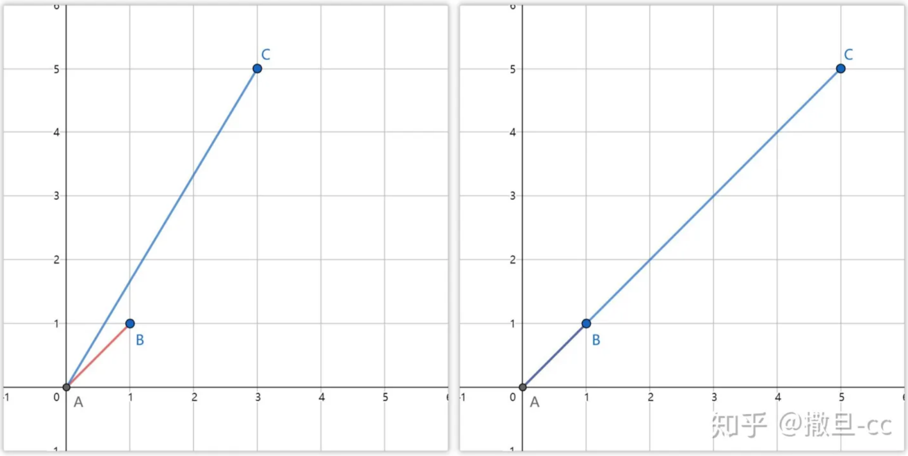
一、理解

当我们在看一个**运动**的时候，我们是如何看的呢？是不是看这个运动的**速度**和**方向**；或者就像物理中的**合力**，我们会拆分成多个**分力**来简化。于是理所当然的会思考，**矩阵**是否也能像这样拆分呢？

1、存在性

我们不得不先说说矩阵的乘法，矩阵乘法本质是一种变换，是把一个向量，通过旋转，拉伸，变成另一个向量的过程

> 举一个例子：给定一个向量$\begin{pmatrix}1\\1\end{pmatrix}$和任意一个矩阵 $A\begin{pmatrix}2&1\\3&2\end{pmatrix}$，他们相乘会得到一个新的向量$\begin{pmatrix}2&1\\3&2\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}3\\5\end{pmatrix}$



通过上面**左图**可以很清楚的看到了这样一个变换的过程，因为矩阵A是随机给出的，这也意味着，我们可以随意变换原来的向量，上面**右图**这样的情况是否也可能出现呢？当然会出现，我们用一个公式来表示描述右图的变换：

$$Av=\lambda v \tag{1}$$

描述的是矩阵A对向量v的变换效果只有拉伸，没有旋转。

也可以看成矩阵 A，向量 v，系数λ这三者建立了一种联系，但显然我们无法通过式(1)来用v和λ表示 A，因为这个式子不是完备的，对于一个秩为m的矩阵 A，应该存在m个这样的式子，完备式子应该是：

$$A\left(\vec{v}_1,\vec{v}_2,\ldots,\vec{v}_m\right)=\left(\lambda_1\vec{v}_1,\lambda_2\vec{v}_2,\ldots,\lambda_m\vec{v}_m\right)=\left(\vec{v}_1,\vec{v}_2,\ldots,\vec{v}_m\right)\begin{bmatrix}\lambda_1&\cdots&0\\\vdots&\ddots&\vdots\\0&\cdots&\lambda_m\end{bmatrix} \tag{2}$$

$$AV=V\Lambda \tag{3}$$

$$A = V\Lambda V^{-1} \tag{4}$$

这种形式是不是就可以看成矩阵A被分解了。

这样我们就得到了一个与运动类似的划分：

- λ 就是运动的速度
- v 就是运动的方向

2、合理性

证明了存在性，那我们为什么要引入这样一种矩阵和向量的表述形式呢？科学家还将满足这样形式的向量v命名为矩阵A的特征向量，系数λ命名为矩阵A的特征值。下面我们就要来探讨这个问题

特征值和特征向量是为了研究向量在经过线性变换后的方向不变性而提出的

一个矩阵和该矩阵的**非特征向量**相乘是对该向量的旋转变换；一个矩阵和该矩阵的**特征向量**相乘是对该向量的伸缩变换，其中伸缩程度取决于特征值大小

矩阵在特征向量所指的方向上具有 **增强（或减弱）特征向量** 的作用。这也就是说，如果矩阵持续地叠代作用于向量，那么特征向量的就会突显出来，我们用Matlab进行计算：

> （在这里用到的实例是与后面计算例子一致，可以先到后面看结果） 我们反复用矩阵 $A = \begin{pmatrix} 4 & 1 & 1 \\ 1 & 2 & 1 \\ 3 & 2 & 3 \end{pmatrix}$ 去乘一个任意向量 $v = \begin{pmatrix} -1 \\ 5 \\ 3 \end{pmatrix}$

```
A = [4 1 1;1 2 1;3 2 3];
V = [-1;5;3];
B = A^1 * V;
C = (B)/sum(B)
```

```
C = 3×1
      0.1250
      0.3750
      0.5000
```



```
A = [4 1 1;1 2 1;3 2 3];
V = [-1;5;3];
B = A^3 * V;
C = (B)/sum(B)
```

```
C = 3×1
      0.3065
      0.2177
      0.4758
```



```
A = [4 1 1;1 2 1;3 2 3];
V = [-1;5;3];
B = A^5 * V;
C = (B)/sum(B)
```

```
C = 3×1
      0.3303
      0.2019
      0.4678
```



```
A = [4 1 1;1 2 1;3 2 3];
V = [-1;5;3];
B = A^10 * V;
C = (B)/sum(B)
```

```
C = 3×1
      0.3333
      0.2000
      0.4667
```

为了提取矩阵这种“不变性”，或者说是为了描述变换（矩阵乘法是一种线性变换）的主要方向是非常有必要的，具体应用会在后面讲到。

这里补充一个理论（来自尹相楠）：

假设矩阵 A 是一个$n \times n$的满秩矩阵，他的单位特征向量为 $v_1, v_2, \ldots, v_n$，对应的特征值为 $\lambda_1, \lambda_2, \ldots, \lambda_n$。那么对一个任意向量$x = x_1 v_1 + x_2 v_2 + \cdots + x_n v_n$，有:

$$\begin{aligned} Ax &= A\left(x_1 \cdot \nu_1 + x_2 \cdot \nu_2 + \ldots + x_n \cdot \nu_n\right) \\ &= x_1\left(A\nu_1\right) + x_2\left(A\nu_2\right) + \ldots + x_n\left(A\nu_n\right) \\ &= x_1\left(\lambda_1\nu_1\right) + x_2\left(\lambda_2\nu_2\right) + \ldots + x_n\left(\lambda_n\nu_n\right) \end{aligned}$$

经过 k 次迭代后:

$$\begin{aligned} A^k x &= x_1\left(\lambda_1^k \nu_1\right) + x_2\left(\lambda_2^k \nu_2\right) + \ldots + x_n\left(\lambda_n^k \nu_n\right) \\ &= \lambda_1^k\left[x_1\nu_1 + x_2\left(\frac{\lambda_2}{\lambda_1}\right)^k \nu_2 + \ldots + x_n\left(\frac{\lambda_n}{\lambda_1}\right)^k \nu_n\right] \end{aligned}$$

由于 $\lambda_1 > \lambda_2 > \ldots > \lambda_n$ (不考虑两个特征值相等的情况，因为太少见了)。所以经过k次迭代后，$\lim_{k\to+\infty}\left(\lambda_i/\lambda_1\right)^k = 0 (i \neq 1)$ 。所以:

$$\lim_{k\to+\infty} A^k x = \lambda_1^k x_1 \nu_1$$

也就是说，经过 k 次迭代后，我们将得到矩阵主特征向量的线性放缩，只要把这个向量归一化，就得到了该矩阵的单位主特征向量，进而可以解出矩阵的主特征值。

二、计算方法

在 (1) 式的基础上，我们进行一些变形：

$$Av = \lambda v \to Av = \lambda I v \to (\lambda I - A)v = 0$$

根据线性方程组理论，为了使这个方程有非零解，矩阵$(\lambda I - A)$的行列式必须是零

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

上式也被称为是A的特征方程，计算出所有 λ 的取值后，再代入$(\lambda I - A)v = 0$求解对应的 v

注意：要注意特征值是重根时的情况！

1、手算

通过一个例子来动手计算一遍

求矩阵 $A = \begin{pmatrix} 4 & 1 & 1 \\ 1 & 2 & 1 \\ 3 & 2 & 3 \end{pmatrix}$ 的特征值和特征向量:

$|\lambda I - A| = \begin{vmatrix} \lambda - 4 & -1 & -1 \\ -1 & \lambda - 2 & -1 \\ -3 & -2 & \lambda - 3 \end{vmatrix} = (\lambda - 6)(\lambda - 2)(\lambda - 1)$，所以

$$\lambda_1 = 6, \lambda_2 = 2, \lambda_3 = 1$$

$$\begin{matrix} -3 & -2 & 3 \end{matrix} \quad v_3$$

$$\begin{cases} 2v_1 - v_2 - v_3 = 0 \\ -v_1 + 4v_2 - v_3 = 0 \\ -3v_1 - 2v_2 + 3v_3 = 0 \end{cases}$$

通过给其中一个变量赋值(通常是1)，得到一个基础解系 $\Rightarrow \begin{cases} v_1 = 5 \\ v_2 = 3 \\ v_3 = 7 \end{cases}$

2.同理得到当 $\lambda = 2$ 时，$\begin{cases} v_1 = 1 \\ v_2 = -1 \\ v_3 = -1 \end{cases}$；

3.当 $\lambda = 1$ 时，$\begin{cases} v_1 = 0 \\ v_2 = 1 \\ v_3 = -1 \end{cases}$

2、Matlab算

我们还可以利用Matlab中的 **eig()** 函数来帮我们求解：

- e = eig(A) 返回一个列向量，其中包含方阵 A 的特征值
- [V,D] = eig(A) 返回特征值的对角矩阵 D 和矩阵 V，其列是对应的右特征向量，使得 $AV = VD$

```
A = [4 1 1;1 2 1;3 2 3];
[V,D] = eig(A)
```

```
A = [4 1 1;1 2 1;3 2 3];
[V,D] = eig(A)

V = 3×3
    0.5488    0.5774    0.0000
    0.3293   -0.5774   -0.7071
    0.7683   -0.5774    0.7071

D = 3×3
    6.0000         0         0
         0    2.0000         0
         0         0    1.0000
```


> 发现Matlab计算的和手算的特征向量值不同，但比例是一样的，这是因为特征向量不是唯一的，特征向量来自齐次线性方程组的解，是齐次线性方程组的**基础解系**的非零线性组合。

应用

1、在图像压缩里，一张图片就可以看作是一个矩阵，当我们提取了这个矩阵的特征值和特征向量后，我们可以只用最大的几个特征值（其他的变为0）就可以基本复原这张图片。

编辑于 2022-10-20 16:50

特征值分解 矩阵

写下你的评论...

默认 最新

请教一下，编程实现特征分解有什么方法，例如用C语言
2021-03-25 回复 1

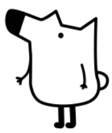
撒旦-cc 作者
很多现成的方法吧，应该一搜一大片
2021-12-09 回复 赞

踢踢
感激👍
01-18 回复 赞

慎ZH
分析的很透彻👍
2022-12-18 回复 赞

list1
把矩阵分解的本质讲透了 nb！！！！😘
2022-11-22 回复 赞

未知
通俗易懂 谢谢



2021-03-25 回复 赞

还好了i
厉害
2022-10-08 回复 赞

取酒酿晚风
太赞了😘😘
2022-09-27 回复 赞

月婵
牛哇！感谢😘😘😘
2022-07-18 回复 赞

发抖的小花猫
棒棒哒👍
2022-06-24 回复 赞

Eleforest
(a-6)(a-2)(a-1)
这个行列式是怎么算出来的？用公式计算后是一个一元三次方程，然后用求根公式算出来么？
2022-02-08 回复 赞

撒旦-cc 作者
很简单的配方法，(a-4)(a-2)(a-3)第一项先别展开，后面剩下展开是-6(a-2)，然后就提取公因式，再就目视配一下就是(a-1)(a-6).
2022-02-08 回复 1

Eleforest ‣ **撒旦-cc**
感谢，确实不应该展开。
2022-02-09 回复 赞

半半小生
感谢楼主
2021-12-09 回复 赞

一文解释系列

用最全的步骤，帮你不走弯路。