

一、知识概要

前面介绍了向量与矩阵之间的乘法，这一节我们要介绍两个矩阵之间的乘法。并讨论逆矩阵存在的条件。最后又介绍了求解逆矩阵的方法。

二. 矩阵乘法

2.1 矩阵乘法最常见求解方式

首先来了解矩阵之间进行乘法运算时，我们是如何求解单个元素的呢？

$$\begin{bmatrix} & ? & \end{bmatrix} \begin{bmatrix} & ? & \end{bmatrix} = \begin{bmatrix} & ? & \end{bmatrix}$$

A * B = C

现在我们取 C 中一个元素：就比如取 C 中一个元素 C_{34} ： $\begin{bmatrix} ? & ? & ? & ? \\ ? & ? & ? & ? \\ ? & ? & ? & C_{34} \end{bmatrix}$ 。

这个下标的意义就是 C 中第 3 行第 4 列。它的计算也和其密切相关。即 A 中第 3 行对 C 中第 3 行第 4 列元素 C_{34} 的计算也和其密切相关。即 A 中第 3 行对 C 中第 3 行第 4 列元素 C_{34} 的计算也和其密切相关。即 A 中第 3 行对 C 中第 3 行第 4 列元素 C_{34} 的计算也和其密切相关。

如图：

A B C

$$\begin{bmatrix} ? & ? & ? & ? \\ ? & ? & ? & ? \\ ? & ? & ? & ? \end{bmatrix} \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix} = \begin{bmatrix} ? & ? & ? & ? \\ ? & ? & ? & ? \\ ? & ? & ? & C_{34} \end{bmatrix}$$

规格： 3x4 x 4x4 = 3x4

$$C_{34} = \sum_{k=1}^n a_{3k}b_{k4} = a_{31}b_{14} + a_{32}b_{24} + a_{33}b_{34} + \dots$$

推广：

$$C_{ij} = (A \text{ 中第 } i \text{ 行向量}) (B \text{ 中第 } j \text{ 列向量}) = \sum_{k=1}^n a_{ik}b_{kj}$$

另外这里还要注意一下矩阵的规格问题，因为我们矩阵乘法运算是使用前一个矩阵的行向量点乘后一个矩阵的列向量，所以从向量角度看，它们必须有相同的分量。故，A 的列数必须与 B 的行数相同，结果 C 矩阵的规格为 A 的行数，B 的列数。

$$\begin{matrix} A & \times & B & = & C \\ m \times n & & n \times p & & m \times p \end{matrix}$$

2.2 列组合与行组合方式

2.2.1 列组合:

还记得我们之前学习过矩阵与列向量的乘积，得到一个列向量：

$$\begin{bmatrix} ? & ? & ? \\ ? & ? & ? \\ ? & ? & ? \end{bmatrix} \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix} = \text{矩阵列的线性组合} = \begin{bmatrix} ? & ? & ? \\ ? & ? & ? \\ ? & ? & ? \end{bmatrix} = 3 \begin{bmatrix} ? \\ ? \\ ? \end{bmatrix} + 4 \begin{bmatrix} ? \\ ? \\ ? \end{bmatrix} + 5 \begin{bmatrix} ? \\ ? \\ ? \end{bmatrix}$$

矩阵 向量

那我们在计算矩阵之间的乘法时，可以把后面的矩阵 B 看做列向量的组合：

$$\begin{bmatrix} ? & ? & ? & ? \\ ? & ? & ? & ? \\ ? & ? & ? & ? \end{bmatrix} \begin{bmatrix} a & b & c & d \\ a & b & c & d \\ a & b & c & d \\ a & b & c & d \end{bmatrix} = \left[A \begin{bmatrix} a \\ a \\ a \\ a \end{bmatrix} \quad A \begin{bmatrix} b \\ b \\ b \\ b \end{bmatrix} \quad A \begin{bmatrix} c \\ c \\ c \\ c \end{bmatrix} \quad A \begin{bmatrix} d \\ d \\ d \\ d \end{bmatrix} \right] = C$$

A B

B 的各列，用来线性组合 A 矩阵的各列向量

之后，每一个类似于 $A \begin{bmatrix} a \\ a \\ a \\ a \end{bmatrix}$ 这样的 A 矩阵与向量乘积，都可转化为矩阵 A 的列

向量线性组合。最后再将得到的列向量组合在 $[C_1 \ C_2 \ C_3 \ \dots]$ 中，即为结果 C

这种方法的关键就是将右侧矩阵 B 看做列向量组合，将问题转化为矩阵与向量的乘法问题。也表明了矩阵 C 就是矩阵 A 中各列向量的线性组合，而 B 其实是在告诉我们，要以什么样的方式组合 A 中的列向量。

2.2.2 行组合:

同理，我们还学习过行向量与矩阵的乘法，得到一个行向量

$$\begin{bmatrix} 1 & 2 & 7 \end{bmatrix} \begin{bmatrix} ? & ? & ? \\ ? & ? & ? \\ ? & ? & ? \end{bmatrix} = \text{矩阵行的线性组合} = \begin{bmatrix} ? & ? & ? \\ ? & ? & ? \\ ? & ? & ? \end{bmatrix} \begin{matrix} *1 \\ *2 \\ *7 \end{matrix}$$

$$= 1 \begin{bmatrix} ? & ? & ? \end{bmatrix} + 2 \begin{bmatrix} ? & ? & ? \end{bmatrix} + 7 \begin{bmatrix} ? & ? & ? \end{bmatrix}$$

同样，按照形式，这次将矩阵 A 看做行向量组合就行了：

$$\begin{matrix} \begin{bmatrix} a & a & a & a \\ b & b & b & b \\ c & c & c & c \end{bmatrix} & \begin{bmatrix} ? & ? & ? & ? \\ ? & ? & ? & ? \\ ? & ? & ? & ? \\ ? & ? & ? & ? \end{bmatrix} & = & \begin{bmatrix} [a & a & a & a]B \\ [b & b & b & b]B \\ [c & c & c & c]B \end{bmatrix} & = & C \\ A & B & & & \end{matrix}$$

之后每一个类似于 $[a \ a \ a \ a]B$ 这样的行向量与矩阵之积，会得到一个行向量，各行向量最后构成了 C 中的各行。

这里要注意， C 中各行都来自于 $[a \ a \ a \ a]B$ 形式求得的行向量，所以是矩阵 B 各行的线性组合组成了 C 的各行。

2.3 列乘以行

常规方法中，计算 $A*B=C$ 矩阵乘法时，使用 A 的行向量乘上 B 的列向量得到 C 中各个位置的元素。

而我们这次介绍的方法，是用 A 的列向量乘上 B 的行向量得到各个矩阵，再将矩阵相加，得到 C 。

我们通过一个例子来讲解。

【例】求解 $\begin{bmatrix} 2 & 7 \\ 3 & 8 \\ 4 & 9 \end{bmatrix} \begin{bmatrix} 1 & 6 \\ 0 & 0 \end{bmatrix}$

列乘行方法：

$$\begin{bmatrix} 2 & 7 \\ 3 & 8 \\ 4 & 9 \end{bmatrix} \begin{bmatrix} 1 & 6 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \begin{bmatrix} 1 & 6 \end{bmatrix} + \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} \begin{bmatrix} 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 12 \\ 3 & 18 \\ 4 & 24 \end{bmatrix}$$

注意这里每一次都是用列向量与行向量相乘得到一个矩阵，而每次得到的矩阵都是有特点的，比如 $\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \begin{bmatrix} 1 & 6 \end{bmatrix} = \begin{bmatrix} 2 & 12 \\ 3 & 18 \\ 4 & 24 \end{bmatrix}$ ，这其中得到的矩阵 $\begin{bmatrix} 2 & 12 \\ 3 & 18 \\ 4 & 24 \end{bmatrix}$ 每一列都和向量 $\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$ 同向，可以说列向量都在 $\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$ 这条直线上，列空间是一条直线。同理，行向量都在 $\begin{bmatrix} 1 & 6 \end{bmatrix}$ 这条直线上，行空间(矩阵行所有可能的线性组合)是一条直线。

2.4. 分块做乘法

分块乘法就是宏观上的矩阵乘法，比如现在有一个 $50*50$ 的矩阵与 $50*50$ 矩阵相乘，一个一个进行运算很麻烦，尤其是如果矩阵在某一区域上有一定的性质，那么我们可以将其分块，如：

$$\begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix} = \begin{bmatrix} C_1 & C_2 \\ C_3 & C_4 \end{bmatrix}$$

其中的 $A_{1,2,3,4}$ 与 $B_{1,2,3,4}$ 都是划分之后的一小块块矩阵，如：

?	?	?
?	?	?
?	?	?

得到

的各个部分。而 $C_1 = A_1B_1 + A_2B_3$ ，和矩阵乘法的计算步骤一样，只是这里的 A_1B_1 ， A_2B_3 都是矩阵之间的乘法而已。只要 A 与 B 分块相互匹配，就可以用这样的分块乘法求解。

三. 逆矩阵

3.1 逆矩阵介绍

之前介绍过一点逆矩阵的定义，对于一个方阵 A，如果 A 可逆，就有这样一个 A^{-1} 使

$$AA^{-1} = I = A^{-1}A$$

如果 A 是非方阵，左侧的 A^{-1} 与右侧的 A^{-1} 不可能相同，因为这时左右侧 A^{-1} 形状一定不相同，那么这就违背了我们说的“一个” A^{-1} 。

再举一个没有逆的矩阵： $\begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$ 。学习过行列式的同学一下子就可以算出来，这个矩阵对应行列式为 0，所以矩阵不可逆。但是本门课程并不是从行列式开始介绍，那么我们换一种看法。我们看到这个矩阵中含两个列向量 $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ ， $\begin{bmatrix} 3 \\ 6 \end{bmatrix}$ 这两个向量，它们之间互成倍数，也就是说这两个向量之一对其线性组合无意义，那么这个 A 不可能有逆。换句话说：

若存在非零向量 x，使得 $Ax = 0$ ，那么 A 就不可能有逆矩阵

为什么呢？这样，如果 A 有逆，在 $Ax = 0$ 这个等式两端同时乘上 A^{-1} ，就有：

$$A^{-1}Ax = Ix = \text{零向量}$$

而 Ix 不可能是零向量，自相矛盾。所以此时 A 没有逆矩阵。

再看看矩阵 $\begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$ ，由于两个列向量线性相关，一定有一个 x，使得 $Ax = 0$ ，如 $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$ 。所以此矩阵不可逆。

3.2 逆矩阵求解

其实求逆矩阵就是解方程组的过程，举例说明：

【例】求解 $\begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix}$ 的逆矩阵

$$\underset{A}{\begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix}} \underset{A^{-1}}{\begin{bmatrix} a & b \\ c & d \end{bmatrix}} = I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

从列向量的角度来看，得到两个方程：

$$\begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix} \begin{bmatrix} a \\ c \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix} \begin{bmatrix} b \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

解这个方程就行了，但是这样做低阶矩阵还好，高阶矩阵计算量未免太大了。所以这里介绍，高斯-若尔当方法：

3.2.1 高斯-若尔当方法

还是上面的例子，两个方程：

$$\begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix} \begin{bmatrix} a \\ c \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$
$$\begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix} \begin{bmatrix} b \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

这个方法就是可以同时处理两个方程组，即使用增广矩阵联系两个方程增广矩阵：

$$\left[\begin{array}{cc|c} 1 & 3 & 1 \\ 2 & 7 & 0 \end{array} \right]$$

接下来进行行变换，将虚线左侧消为单位矩阵 I，此时右侧矩阵即为逆矩阵。

$$\left[\begin{array}{cc|c} 1 & 3 & 1 \\ 2 & 7 & 0 \end{array} \right] \rightarrow \left[\begin{array}{cc|c} 1 & 3 & 1 \\ 0 & 1 & -2 \end{array} \right] \rightarrow \left[\begin{array}{cc|c} 1 & 0 & 7 \\ 0 & 1 & -2 \end{array} \right]$$

逆矩阵即为： $\begin{bmatrix} 7 & -3 \\ -2 & 1 \end{bmatrix}$

检验一下， $AA^{-1} = I$ ，正确。

接下来论证它的合理性：

我们知道，上面这个过程，对 $\begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix}$ 进行消元处理使它变为单位矩阵 I，就相当于左乘一大堆消元矩阵，假设为 E，就有： $E \begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix} = I$ ，那么 E 肯定就是逆矩阵 A^{-1} ，而再看虚线右边，单位矩阵 $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ 经历了与 $\begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix}$ 同样的消元过程，最后

的结果相当于： $EI = A^{-1}$ 。那这虚线右侧得到的结果就是 A^{-1} 没错了。

四. 学习感悟

这节介绍了认识矩阵乘法的不同角度，并介绍了逆矩阵的相关知识以及如何即求解逆矩阵。这节内容很好的体现了我自己认为的这门课的优点之一：少有繁琐的证明，更多的理解与类比。多从向量，空间，线性组合的角度去认识矩阵之间的运算，这是这门课的核心之一。

-----by Dkb

机器学习算法与自然语言处理
丁坤博