计算语言学的数学基础

(1) 线性代数提要

王予沛

2025 年 3 月 3 日

北京师范大学数字人文系

绪论

本课程的主要内容

麻省理工学院数学系荣休教授 Gilbert Strang[1] 所认为的线性代数的基础内容包括:

1. 消元法求解 $Ax = b$
2. 矩阵运算、逆运算和行列式
3. 向量空间和子空间
4. 无关性、维数和矩阵的秩
5. 特征值和特征向量

本课程中，我们将覆盖上述内容，但并不会按照这个顺序讲授。我们将首先复习高中接触过的向量及其运算，然后从直观的线性变换开始，逐步进入矩阵的世界。

[1] https://math.mit.edu/~gs/

什么是线性代数?

代数：用字母代替数。

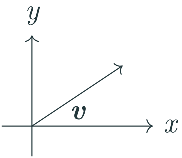
研究算数的时候发现的一些和具体的数无关的性质，例如：

- 加法交换律：$a+b=b+a$
- 加法结合律：$(a+b)+c=a+(b+c)$
- 乘法交换律：$ab=ba$
- 乘法结合律：$(ab)c=a(bc)$

向量和向量空间

向量

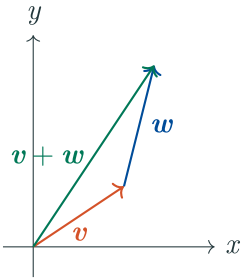
- 向量[2]是具有大小和方向的量。
- 向量可以表示为有序的数构成的元组，例如 $\boldsymbol{v} = (3, 2)$



[2] 向量可以表示为字母在字母上方加箭头，例如 $\vec{v}$，也可以用黑体表示，例如 $\boldsymbol{v}$，本课程中我们会同时使用这两种表示方法，但更多使用后者。

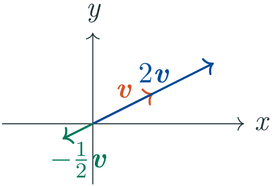
向量加法

- 向量加法可以理解为“位移的叠加”
- 例如：从家到学校的路径 $\boldsymbol{v} = (3, 2)$，从学校到图书馆的路径 $\boldsymbol{w} = (1, 4)$
- 那么从家到图书馆的直接路径就是：
  $\boldsymbol{v} + \boldsymbol{w} = (3 + 1, 2 + 4) = (4, 6)$



数量乘法（标量乘法）

- 向量的数量乘法表示向量在同一直线上的“伸缩变换”
- 正数：保持方向，改变大小
- 负数：改变方向，改变大小
- 例如：速度向量 $\boldsymbol{v} = (2, 1)$
  - $2\boldsymbol{v} = (4, 2)$ 表示“速度加倍”
  - $-\frac{1}{2}\boldsymbol{v} = (-1, -0.5)$ 表示”速度减半且方向相反”



向量的线性组合

向量加法和数量乘法是线性代数的两种核心操作。我们将两向量相加得到 $\boldsymbol{v}+\boldsymbol{w}$，将向量与标量相乘得到 $c\boldsymbol{v}$ 和 $d\boldsymbol{w}$。结合两种操作，我们得到向量的**线性组合**：$c\boldsymbol{v}+d\boldsymbol{w}$。

$$c\boldsymbol{v}+d\boldsymbol{w}=c\begin{pmatrix}1\\1\end{pmatrix}+d\begin{pmatrix}2\\3\end{pmatrix}=\begin{pmatrix}c+2d\\c+3d\end{pmatrix}$$

其中，c 和 d 是标量，称为线性组合的**组合系数**。

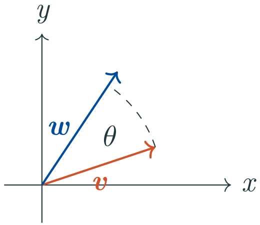
当 c 取不同的值，$c\boldsymbol{v}$ 均在一条直线上，当 $\boldsymbol{w}$ 不在这条直线上时，线性组合 $c\boldsymbol{v}+d\boldsymbol{w}$ 填满了整个二维平面。

向量内积（点乘）

- 两个向量 $\boldsymbol{v}$ 和 $\boldsymbol{w}$ 的**内积**[3]定义为：

$$\boldsymbol{v} \cdot \boldsymbol{w} = a_1 b_1 + a_2 b_2$$

- 几何意义：$\boldsymbol{v} \cdot \boldsymbol{w} = |\boldsymbol{v}||\boldsymbol{w}| \cos \theta$
- 其中 θ 是两个向量之间的夹角
- 重要性质：
  - 内积为零 $\Leftrightarrow$ 两向量垂直
  - $\boldsymbol{v} \cdot \boldsymbol{v} = |\boldsymbol{v}|^2$



[3] 向量的内积也称**数量积、点积、点乘**；相对地，还存在一种称为外积，也称向量积、叉积、叉乘的运算，这里不展开。

内积与度量

- 向量的长度[4]通过内积定义[5]:

$$\|\boldsymbol{v}\| = \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}}$$

- 两点间的距离[6]自然也可以用内积定义:

$$\mathrm{d}(\boldsymbol{v}, \boldsymbol{w}) = \|\boldsymbol{v} - \boldsymbol{w}\| = \sqrt{(\boldsymbol{v} - \boldsymbol{w}) \cdot (\boldsymbol{v} - \boldsymbol{w})}$$

- 内积的重要性:
  - 定义了空间中的度量（距离）
  - 使我们能够谈论长度、距离，进而可以讨论角度

[4] 向量的长度也称为**向量的模**。

[5] 中学时，我们用单竖线 $|$ 表示向量的长度，现在起我们用双竖线 $\|$ 表示向量的长度。这是为了与绝对值区分。

[6] 两点间的距离即两个向量终点间的距离。

向量组是向量构成的集合。

考虑向量 $v=(1,2)$ 和 $w=(2,4)$。它们构成的集合 $\{v,w\}$ 是一个向量组。同时，我们观察到 $w=2v$，即两个向量的终点在一条直线上。此时，我们称向量 v 和 w **线性相关**，向量 w 可以由 v **线性表出**，反过来，v 也可以由 w 线性表出。

考虑向量 $i=(1,0)$ 和 $j=(0,1)$。它们构成的集合 $\{i,j\}$ 也是一个向量组，但这两个向量并不平行，即 i 和 j 不能相互线性表示。此时，我们称向量 i,j **线性无关**。

线性相关

向量组 $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_s\}$，$(s \geqslant 1)$ 是**线性相关**的，如果存在一组不全为零的实数[7] $k_1, k_2, \cdots, k_s$，使得

$$k_1\boldsymbol{v}_1 + k_2\boldsymbol{v}_2 + \cdots + k_s\boldsymbol{v}_s = \mathbf{0} \tag{1}$$

线性无关

向量组 $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_s\}$，$(s \geqslant 1)$ 是**线性无关**的，如果式 (1) 只有零解 $k_1 = k_2 = \cdots = k_s = 0$。

[7] 事实上，这里并不一定是实数，只需要是某一个数域 $\mathcal{F}$ 中的数即可。由于数域需要额外进行定义，这里不展开。在计算语言学中，我们通常也只需考虑实数域。

向量组的最大线性无关组

一个向量组的一个部分向量组称为一个**最大线性无关组**，如果这个部分组本身是线性无关的，并且从这向量组中任意添一个向量（如果还有的话），所得的部分向量组都线性相关。

例 1：向量 $v_1 = (1,1), v_2 = (2,2)$ 构成的向量组 $\{v_1, v_2\}$ 的最大线性无关组为 $\{v_1\}$ 或 $\{v_2\}$。

例 2：向量 $v_1 = (1,1), v_2 = (1,2), v_3 = (2,2)$ 构成的向量组 $\{v_1, v_2, v_3\}$ 的最大线性无关组为 $\{v_1, v_2\}$ 或 $\{v_2, v_3\}$。

例 3：向量 $v_1 = (1,0), v_2 = (0,1)$ 构成的向量组 $\{v_1, v_2\}$ 的最大线性无关组为 $\{v_1, v_2\}$。

例 4：向量 $v_1 = (1,0), v_2 = (0,1), v_3 = (2,3)$ 构成的向量组 $\{v_1, v_2, v_3\}$ 的最大线性无关组为 $\{v_1, v_2\}$ 或 $\{v_2, v_3\}$ 或 $\{v_1, v_3\}$。

一个向量组的最大线性无关组不唯一，但所有最大线性无关组所含向量的个数相同。

向量空间

向量空间 $\mathbb{R}^n$ 是所有包含 n 个实数分量 (components) 的向量构成的向量组。向量空间简称**空间 (space)**。

例如，$\mathbb{R}^2$ 表示常见的 xy 平面，其中每一个向量有两个分量。"空间" 这个词旨在强调所有这样的向量构成的集合，即整个平面。每一个向量给出横坐标 x 和纵坐标 y，表示在平面中的一个点 $\boldsymbol{v} = (x, y)$。

类似地，$\mathbb{R}^3$ 中的向量表示三维空间中的点 (x, y, z)。一维空间则是一条直线（比如 x 轴），表示为 $\mathbb{R}^1$。

向量空间的核心要求

在向量空间中，向量的两种核心操作（加法和数量乘法）的结果仍然在这个向量空间中。这叫做对加法和数量乘法的**封闭性**。

严格来说，向量空间时一个满足 8 条性质且对加法和数量乘法封闭的向量组。

向量空间的定义

我们称向量组 V 是定义在一个域 $\mathcal{F}$ 上的向量空间，当且仅当 V 需要满足以下八条性质[8]：

- 向量加法的交换律：$\forall \boldsymbol{x}, \boldsymbol{y} \in V, \quad \boldsymbol{x}+\boldsymbol{y}=\boldsymbol{y}+\boldsymbol{x}$
- 向量加法的结合律：$\forall \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{w} \in V, \quad (\boldsymbol{x}+\boldsymbol{y})+\boldsymbol{z}=\boldsymbol{x}+(\boldsymbol{y}+\boldsymbol{z})$
- 向量加法的单位元 $\mathbf{0}$ 存在且唯一：

$$\exists \mathbf{0} \in V, \quad \forall \boldsymbol{y} \in V, \quad \boldsymbol{y}+\mathbf{0}=\boldsymbol{y}$$

- 向量加法的逆元素存在且唯一：

$$\forall \boldsymbol{y} \in V, \quad \exists -\boldsymbol{y} \in V \text{ 使得} \boldsymbol{y}+(-\boldsymbol{y})=\mathbf{0}$$

[8]建议将这 8 条性质作为向量计算的参考，因为在计算语言学研究中我们只需要实数域上的实向量空间 $\mathbb{R}^n$。对八条性质是否必要的讨论有很多，参见：`https://math.stackexchange.com/q/1412899`

向量空间的定义

- 数乘（标量乘法）对于向量加法分配律：

$$\forall c \in \mathcal{F}, \forall \boldsymbol{x}, \boldsymbol{y} \in V, \quad c(\boldsymbol{x} + \boldsymbol{y}) = c\boldsymbol{x} + c\boldsymbol{y}$$

- 数乘对于域 $\mathcal{F}$ 上标量加法的分配律：

$$\forall c_1, c_2 \in \mathcal{F}, \forall \boldsymbol{y} \in V, \quad (c_1 + c_2)\boldsymbol{y} = c_1\boldsymbol{y} + c_2\boldsymbol{y}$$

- 数乘对于域 $\mathcal{F}$ 上标量乘法的结合律：

$$\forall c_1, c_2 \in \mathcal{F}, \forall \boldsymbol{y} \in V, \quad (c_1 c_2)\boldsymbol{y} = c_1(c_2\boldsymbol{y})$$

- 数域 $\mathcal{F}$ 中存在单位元 1，使得：

$$\exists 1 \in \mathcal{F}, \quad \forall \boldsymbol{y} \in V, \quad 1\boldsymbol{y} = \boldsymbol{y}$$

向量空间的子空间

设 V 是向量空间，W 是 V 的子集。如果 W 对向量加法和数量乘法封闭，则称 W 是 V 的**子空间**。这时候，8 条性质将自动满足。请回答：

- 为什么向量加法的逆元素一定在子空间中？
- 为什么向量加法的单位元（零向量 $\mathbf{0}$）一定在子空间中？

例子：

- $\mathbb{R}^3$: 整个三维空间
- $\mathbb{L}$: 所有穿过原点的直线
- $\mathbb{P}$: 所有穿过原点的平面
- $\{\mathbf{0}\}$: 零向量

思考： 向量 (x, y) 在 $x, y \geq 0$ 时是否构成 $\mathbb{R}^2$ 的子空间？如果加上 $x, y \leq 0$ 呢？

“张成” 子空间的向量组

如果一向量组的所有线性组合填满了 (fill) 整个空间，则称该向量组**张成 (span)**该空间。

例 1：$\boldsymbol{v}_1 = (1, 0)$ 和 $\boldsymbol{v}_2 = (0, 1)$ 张成整个二维空间 $\mathbb{R}^2$。

例 2：$\boldsymbol{v}_1 = (1, 0), \boldsymbol{v}_2 = (0, 1), \boldsymbol{v}_3 = (4, 7)$ 同样张成整个 $\mathbb{R}^2$。

例 3：$\boldsymbol{w}_1 = (1, 1), \boldsymbol{w}_2 = (-1, -1)$ 仅张成 $\mathbb{R}^2$ 中的一条直线。$\boldsymbol{w}_1 = (1, 1)$ 自己也是如此。

向量空间的基底和维数

如果一向量组张成一个向量空间，并且这个向量组中所有的向量都是线性无关的，则称该向量组为该向量空间的一个**基底**。基底也简称**基**。基底中的向量称为**基向量**。

例：$\boldsymbol{v}_1 = (1, 0)$ 和 $\boldsymbol{v}_2 = (0, 1)$ 是 $\mathbb{R}^2$ 的 “标准基”。

在给定基底的情况下，向量空间中的任意一个向量都可以唯一地表示为基底的线性组合。

思考：为什么这样的线性组合是唯一的？（反证法）

一个向量空间可以有无穷多个不同的基底，但不同基底中基向量的个数相同（请尝试用反证法说明这一点）。

向量空间的各基底中所含基向量的个数称为该向量空间的**维数**。例如，$\mathbb{R}^2$ 的维数为 2。

向量空间是线性代数的核心概念，它不仅适用于几何向量，还可以推广到更一般的对象。下面列举一些常见的非几何向量空间。

- 所有实系数多项式构成的集合 $\mathbb{R}[x]$。
- 所有次数小于等于 n 的实系数多项式的集合 $\mathbb{R}[x]_{\leq n}$。
- 所有实函数 $f\colon \mathbb{R} \to \mathbb{R}$ 构成的集合 $\mathbb{F}(\mathbb{R})$。

从线性变换到矩阵

向量的几何变换

本节讨论向量在空间中的几何变换，比如：

- 旋转：绕定点（通常是原点）转动
- （镜像）反射：按横/纵坐标轴翻转

注意到，向量仅有方向和长度属性，没有位置信息。数学上，平移向量没有意义，因为改变向量的起点不会改变其本质。因此向量没有平移变换。

什么是恒等变换?

恒等变换是最基本的几何变换，它将每个点都映射到自身，因此在坐标系中，点的坐标保持不变，向量的终点也保持不变，故任何向量在恒等变换后都保持不变。

恒等变换的数学表示

设点 $P(x, y)$ 经过恒等变换后变为 $P'(x', y')$:

$$\begin{cases} x' = x, \\ y' = y. \end{cases} \tag{2}$$

式 (2) 可改写为:

$$\begin{cases} x' = 1 \cdot x + 0 \cdot y, \\ y' = 0 \cdot x + 1 \cdot y. \end{cases} \tag{3}$$

恒等变换的矩阵表示

我们可以将式 (3) 的系数取出来写成下面的 “数表”:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

这是一个由数字构成的矩形表格，称为**矩阵**。一般使用大写字母表示矩阵，例如我们可以把上面的矩阵记为 A。矩阵 A 有 2 行 2 列，这时我们称 A 是一个 2×2 的矩阵。用于表示变换的矩阵就称为**变换矩阵**。特别地，恒等变换的变换矩阵称为**单位矩阵** (一般记为 I)，它唯一确定了恒等变换。

可以看出，变换可以写成变换矩阵以某种方式 $\boxed{?}$ 作用在向量上:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \boxed{?} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

什么是旋转变换?

旋转变换是将平面内每个点绕原点旋转固定角度的几何变换，对向量而言，就是将向量的终点绕原点旋转相应角度，保持向量的长度不变但改变其方向。

旋转变换的数学表示

如图1所示，在直角坐标系 Oxy 中，点 $P(x,y)$ 绕原点 O 逆时针旋转 $180°$ 后变为点 $P'(x',y')$:

$$\begin{cases} x' = -x, \\ y' = -y. \end{cases} \tag{4}$$

式 (4) 可改写为:

$$\begin{cases} x' = -1 \cdot x + 0 \cdot y, \\ y' = 0 \cdot x + -1 \cdot y. \end{cases} \tag{5}$$

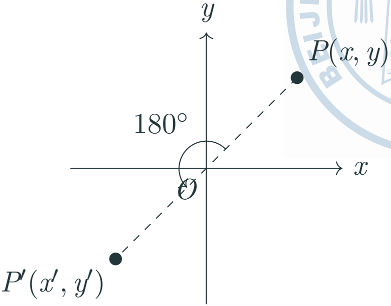


图 1: 旋转变换示意图

我们可以将式 (5) 的系数取出来写成矩阵:

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

这个**旋转矩阵**唯一确定了旋转 180° 这个变换。实际上，对于任意角度 θ 的旋转，都存在唯一的旋转矩阵与之对应。同样地，我们可以将旋转变换写成其变换矩阵以某种方式 $\boxed{?}$ 作用在向量上:

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \boxed{?} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -x \\ -y \end{pmatrix}$$

补充：在直角坐标系 Oxy 内, 将任意一点 $P(x, y)$ 绕原点 O 按逆时针方向旋转 θ 角得到点 $P'(x', y')$ 的旋转变换矩阵是

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

伸缩变换（stretching）是将直角坐标系内每个点的横坐标变为原来的 k_1 倍, 纵坐标变为原来的 k_2 倍的几何变换，其中 k_1, k_2 均为非零常数。对向量而言，就是将向量的各个分量分别伸缩相应倍数。

如图2所示，在直角坐标系 Oxy 中，将点 $P(x, y)$ 的纵坐标变为原来的两倍，横坐标不变，得到点 $P'(x', y')$：

$$\begin{cases} x' = x, \\ y' = 2y. \end{cases} \quad (6)$$

式 (6) 可以改写为：

$$\begin{cases} x' = 1 \cdot x + 0 \cdot y, \\ y' = 0 \cdot x + 2 \cdot y. \end{cases} \quad (7)$$

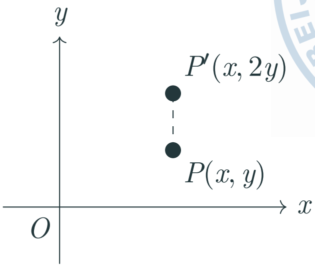


图 2: 伸缩变换示意图

伸缩变换的矩阵表示

我们可以将式 (7) 的系数取出来写成矩阵:

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

这个**伸缩矩阵**唯一确定了纵坐标放大 2 倍这个变换。同样地，我们可以将伸缩变换写成其变换矩阵以某种方式 $\boxed{?}$ 作用在向量上:

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \boxed{?} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 2y \end{pmatrix}$$

思考：将直角坐标系 Oxy 内每个点的横坐标变为原来的 k_1 倍，纵坐标变为原来的 k_2 倍（k_1, k_2 均为非零常数）的伸缩变换的变换矩阵是什么?

切变变换的概念

切变变换（shearing）是一种保持某一坐标轴方向不变，而其他位置的平移量与到该轴的垂直距离成正比的线性变换。

在二维平面上，切变变换主要有两种：

- **水平切变**：保持 x 轴方向不变，点的水平平移量与其 y 坐标成正比；
- **垂直切变**：保持 y 轴方向不变，点的垂直平移量与其 x 坐标成正比。

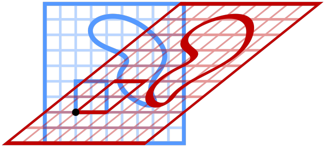


图 3: 平面的水平剪切，将蓝色转变为红色，黑点为原点

水平切变变换的数学表示

如图4所示，在直角坐标系 Oxy 内，水平切变将每一点 $P(x,y)$ 沿着与 x 轴平行的方向平移 ky 个单位变成点 $P'(x',y')$，其中 k 是非零常数：

$$\begin{cases} x' = x + ky, \\ y' = y. \end{cases} \tag{8}$$

式 (8) 可以改写为：

$$\begin{cases} x' = 1 \cdot x + k \cdot y, \\ y' = 0 \cdot x + 1 \cdot y. \end{cases} \tag{9}$$

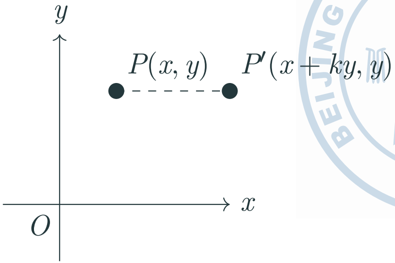


图 4: 水平切变变换示意图

切变变换的矩阵表示

我们可以将式 (9) 的系数取出来写成矩阵:

$$\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$$

这个**切变矩阵**唯一确定了水平切变这个变换，其中 k 表示切变的程度。同样地，我们可以将切变变换写成其变换矩阵以某种方式 $\boxed{?}$ 作用在向量上:

$$\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \boxed{?} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + ky \\ y \end{pmatrix}$$

思考:

- 垂直切变变换的矩阵形式是什么?
- 为什么切变变换的矩阵主对角线上的元素都是 1?

线性变换

线性变换是从一个向量空间到另一个向量空间的映射 $\mathcal{T}: V \to W$，对于所有的 $\boldsymbol{v}, \boldsymbol{w} \in V$ 和任意标量 k_1, k_2，满足以下两个条件：

1. **加法保持性：** $\mathcal{T}(\boldsymbol{v} + \boldsymbol{w}) = \mathcal{T}(\boldsymbol{v}) + \mathcal{T}(\boldsymbol{w})$
2. **数量乘法保持性：** $\mathcal{T}(k\boldsymbol{v}) = k\mathcal{T}(\boldsymbol{v})$

线性变换也称为**线性映射**。

满足上述条件的映射不会改变直线的形状、平行性、与其他直线的交叉点的位置以及直线上点的相对距离。这就是“线性”这一定语的来源。可以看出，之前介绍的旋转、伸缩、切变都是线性变换。

不满足上述条件的映射称为**非线性变换**。例子有很多，如 $f(x) = x^2$，它将直线映射为抛物线。

矩阵与向量的乘法

从前面的例子可见，几种几何变换可以通过矩阵以某种方式作用在向量上来实现。这种作用 $\boxed{?}$ 实际上就是**矩阵与向量的乘法**。具体来说，矩阵 A（从左边）乘以向量 x 的结果是一个向量，这个向量的第 i 个分量等于 A 的第 i 行行向量与列向量x 的内积。

以水平切变为例，矩阵 A 乘以向量 x 的结果是:

$$Ax = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} (1,k)\cdot(x,y) \\ (0,1)\cdot(x,y) \end{pmatrix} = \begin{pmatrix} x+ky \\ y \end{pmatrix}$$

可以看到，结果向量的第一个分量是矩阵 A 的第一行行向量与 x 的内积，第二个分量是矩阵 A 的第二行行向量与 x 的内积。

注意到，如果矩阵 A 的列数与向量 x 无法做内积，则矩阵与向量的乘法无法进行。因此，矩阵与向量的乘法要求矩阵的列数（即行向量的分量数）与向量的行数（即列向量的分量数）相等。

线性变换的复合

我们可以先后在一个向量上施加多个线性变换，得到一个复合变换。取水平剪切矩阵 A 和垂直剪切矩阵 B:

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

考虑一个向量 $x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$，比较两种不同的变换顺序：

- 先施加 B，再施加 A:

$$A(B\boldsymbol{x}) = A\left(\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) = A\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

- 先施加 A，再施加 B:

$$B(A\boldsymbol{x}) = B\left(\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) = B\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

显然，$A(B\boldsymbol{x}) \neq B(A\boldsymbol{x})$。线性变换的顺序不同，结果也不同。

矩阵与矩阵的乘法

线性变换的复合可以表示为矩阵与矩阵的乘法：

$$A(B\boldsymbol{x}) = (AB)\boldsymbol{x}, \quad B(A\boldsymbol{x} = (BA)\boldsymbol{x})$$

这种乘法和向量乘法满足结合律，且可以看作多次矩阵与向量的乘法 $AB = [Ab_1 \cdots Ab_n]$，其中 b_i 是 B 的列向量：

$$AB = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

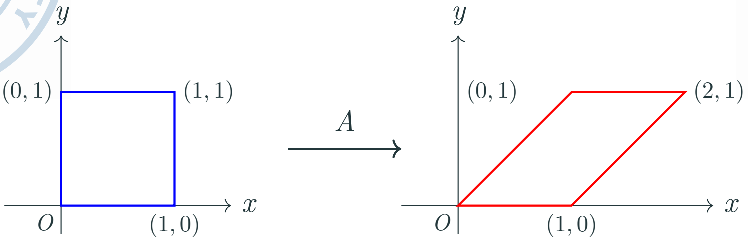
$$BA = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

显然，$AB \neq BA$。

矩阵的行列式

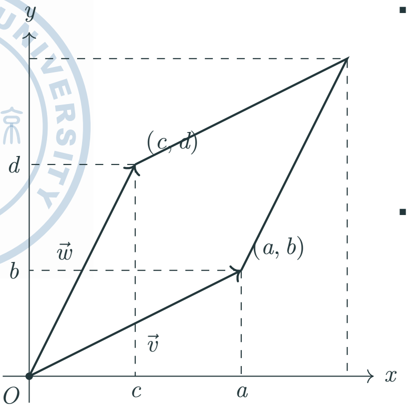
方形矩阵的**行列式**（determinant）表示改矩阵对应的线性变换将单位（超）立方体变换后得到的新（超）立方体的有向体积，其绝对值是体积比率 (volume distortion)，符号表示定向[9]是否翻转。矩阵 A 的行列式记作 $\det(A)$。

以水平切变为例，考虑变换矩阵 $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$，它将单位正方形变换为平行四边形。



[9]定向是坐标轴之间的相对位置，也即手性(chirality)。

考虑一般的 2×2 矩阵 $A=\begin{pmatrix} a & c \\ b & d \end{pmatrix}$，它将单位正方形变换为如下的平行四边形：



- 外框大矩形面积:$(a+c)\times(b+d)$
- 需减去的区域:

$$2\times\frac{1}{2}ab=ab \quad \text{(上下两个三角形)}$$
$$2\times\frac{1}{2}cd=cd \quad \text{(左右两个三角形)}$$
$$2\times ac \quad \text{(左上和右下两个矩形)}$$

- 平行四边形面积:

$$(a+c)(b+d)-(ab+cd+ac+bd)$$
$$=\cancel{ab}+ad+bc+\cancel{cd}-\cancel{ab}-\cancel{cd}-ac-bd$$
$$=ad-bc$$

- 最终结果：$\det(A)=\begin{vmatrix} a & c \\ b & d \end{vmatrix}=ad-bc$

特别地，当 $\det(A) = 0$ 时，矩阵 A 对应的线性变换将单位（超）立方体压缩为零体积的低维几何体（点或线），这时我们称矩阵 A 是**奇异矩阵**。

反之，如果 $\det(A) \neq 0$，我们称矩阵 A 是**非奇异矩阵**。

奇异矩阵意味着线性变换过程中发生了某种信息损失。

线性方程组和矩阵的关系

之前，我们通过大量的例子，将对向量的线性变换抽象为了矩阵与向量的乘法，记作 $A\boldsymbol{x}=\boldsymbol{b}$，其中 $A,\boldsymbol{x}$ 已知，$\boldsymbol{b}$ 作为待求的变换后的结果，是未知的。

现在，我们考虑一个逆问题：已知 A 和 $\boldsymbol{b}$，求 $\boldsymbol{x}$。具体地，

$$A\boldsymbol{x}=\boldsymbol{b} \Leftrightarrow \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \tag{10}$$

利用矩阵乘法，我们可以将上式写为：

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases} \tag{11}$$

由于其中 $a_{11}, a_{12}, a_{21}, a_{22}, b_1, b_2$ 是已知的，x_1, x_2 是未知的，所以式 (11) 是一个线性方程组，A 称为方程组的系数矩阵，式 (10) 则称为线性方程组的矩阵形式，也可简称为线性方程组。

有若干只鸡兔同在一个笼子里，从上面数，有 10 个头；从下面数，有 28 只脚。求笼中各有几只鸡和兔?

设笼中鸡有 x 只，兔有 y 只，则该问题对应的线性方程组为:

$$\begin{cases} x+y=10 \\ 2x+4y=28 \end{cases} \tag{12}$$

将该方程组写成矩阵形式:

$$A\boldsymbol{x}=\boldsymbol{b} \Leftrightarrow \begin{pmatrix} 1 & 1 \\ 2 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}=\begin{pmatrix} 10 \\ 28 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1\cdot x+1\cdot y \\ 2\cdot x+4\cdot y \end{pmatrix}=\begin{pmatrix} 10 \\ 28 \end{pmatrix} \tag{13}$$

式 (13) 可解读为将矩阵 A 作用于向量 $\boldsymbol{x}$，得到向量 $\boldsymbol{b}$。矩阵 A 表示一个线性映射，含义是“数头数脚”。

注意到，矩阵 A 的列向量分别为 $\begin{pmatrix}1\\2\end{pmatrix}$ 和 $\begin{pmatrix}1\\4\end{pmatrix}$，它们分别表示鸡和兔 $\begin{pmatrix}\text{头}\\\text{脚}\end{pmatrix}$ 的数量，可分别称“鸡向量”和“兔向量”。

在我们中学的认识中，线性方程组中的不同方程间是孤立的，不同方程之间 x 的系数是孤立的，y 的系数和等式右边的常数项也是如此。这种认识方程的方式是一行一行地看的，不同方程表示不同的约束条件，最后再把约束条件联立起来。

观察式 (13)，我们发现，鸡 x 和兔 y 对于结果 $\boldsymbol{b}$ 的贡献是独立的。这种独立体现在鸡向量和兔向量可以通过线性组合得到结果向量 $\boldsymbol{b}$:

$$x\begin{pmatrix}1\\2\end{pmatrix}+y\begin{pmatrix}1\\4\end{pmatrix}=\begin{pmatrix}10\\28\end{pmatrix} \tag{14}$$

式 (14) 表明，矩阵 A 左乘向量 $\boldsymbol{x}$ 可以看作将向量 $\boldsymbol{x}$ 的各个分量作为组合系数将矩阵 A 的各个对应列向量进行线性组合，组合的结果向量便是 $\boldsymbol{b}$。

这意味着，如果 $\boldsymbol{b}$ 存在于矩阵 A 的列向量所张成的空间中，则方程组一定有解；反之，如果 $\boldsymbol{b}$ 不在矩阵 A 的列向量所张成的空间中，则方程组无解。

我们将矩阵 A 的列向量所张成的空间，称为矩阵 A 的**列空间**。

因此，方程组有解的充要条件是 $\boldsymbol{b}$ 在矩阵 A 的列空间中。

矩阵的列秩和增广矩阵

矩阵 A 的列向量组的最大线性无关组的向量个数称为矩阵 A 的**列秩**。我们将矩阵 A 和向量 $\boldsymbol{b}$ 并在一起，称为方程组 $A\boldsymbol{x}=\boldsymbol{b}$ 的**增广矩阵**，记作 $[A|\boldsymbol{b}]$。此时，方程组有解的充要条件是增广矩阵的 $[A|\boldsymbol{b}]$ 列秩等于矩阵 A 的列秩。（为什么?）

矩阵 A 的列秩表示未知数向量 $\boldsymbol{x}$ 可以自由取值的分量个数，也就是方程组 $A\boldsymbol{x}=\boldsymbol{b}$ 独立参数的个数。

对照列空间的概念，我们将矩阵 A 的行向量所张成的空间，称为矩阵 A 的**行空间**。

在方程组 (11) 中，每个方程实际上都给出了一个线性约束，要求未知向量 $\boldsymbol{x}$ 与对应的行向量的内积必须等于特定值。因此，方程组的解就是同时满足所有这些行向量给出的线性约束的 $\boldsymbol{x}$。

如果某个方程（即某个行向量）可以由其他方程（行向量）线性组合得到，那么这个方程实际上是冗余的，因为它提供的约束已经被其他方程隐含了。

矩阵 A 的行向量组的最大线性无关组的向量个数称为矩阵 A 的**行秩**，它恰好等于方程组中实际起作用的独立约束的数量。

由于线性方程组中独立约束的数量一定等于独立参数的数量[10]，故矩阵的行秩和列秩相等，合称矩阵的**秩**，记作 $\mathrm{rank}(A)$。

[10]这只是一种理解，证明较为复杂。

线性方程组解的存在性和结构

对于齐次线性方程组$A\boldsymbol{x}=\boldsymbol{0}$（其中 A 为 $m\times n$ 矩阵，$\boldsymbol{x}$ 为 n 维列向量，$\boldsymbol{0}$ 为 m 维零向量）分以下两种情况：

- 如果 $\operatorname{rank}(A)=n$，则方程组有唯一解，即 $\boldsymbol{x}=\boldsymbol{0}$；
- 如果 $\operatorname{rank}(A)<n$，则方程组有无穷多解。

对于非齐次线性方程组$A\boldsymbol{x}=\boldsymbol{b}$（其中 $\boldsymbol{b}\neq\boldsymbol{0}$）分以下三种情况：

- 如果 $\operatorname{rank}(A)=\operatorname{rank}([A|\boldsymbol{b}])=n$，则方程组有唯一解；
- 如果 $\operatorname{rank}(A)=\operatorname{rank}([A|\boldsymbol{b}])<n$，则方程组有无穷多解；
- 如果 $\operatorname{rank}(A)<\operatorname{rank}([A|\boldsymbol{b}])$，则方程组无解。

特别地，非齐次线性方程组 $A\boldsymbol{x}=\boldsymbol{b}$（其中 $\boldsymbol{b}\neq\boldsymbol{0}$）的通解可以表示为齐次线性方程组 $A\boldsymbol{x}=\boldsymbol{0}$ 的通解加上一个特解。（为什么?）

当确定方程组 $A\boldsymbol{x}=\boldsymbol{b}$ 有唯一解时，我们希望找到它的解 $\boldsymbol{x}$。

回忆一元一次方程 $ax=b$，其中 a 和 b 已知，x 未知，且 $a\neq 0$，它的解为 $x=\frac{b}{a}$。这时我们想，如果矩阵也有除法就好了。

换个写法，$x=\frac{b}{a}=a^{-1}b$，这时候我们不把 $\frac{\bullet}{a}=a^{-1}\cdot$ 看成除法，而是视为 $a\cdot$ 的逆运算。

那么，对于矩阵 A 表示的线性变换，是否存在对应的逆变换 B，使得 $B\cdot A=I$ 呢？如果存在，我们称 A 是**可逆矩阵**，并称 B 为 A 的**逆矩阵**，记作 A^{-1}。

考虑矩阵 A 的可逆性实际上是考虑矩阵 A 对应的线性变换的可逆性。我们之前按照特征值 $\det(A)$ 是否为零讨论了线性变换中的信息损失，据此也可以对矩阵的可逆性做下面的分类讨论[11]:

- 若矩阵 A 对应的线性变换可逆，则变换过程不应该存在信息损失，因此 A 的行列式 $\det(A) \neq 0$。
- 若矩阵 A 对应的线性变换存在信息损失，则变换过程存在信息损失，因此 A 的行列式 $\det(A) = 0$。

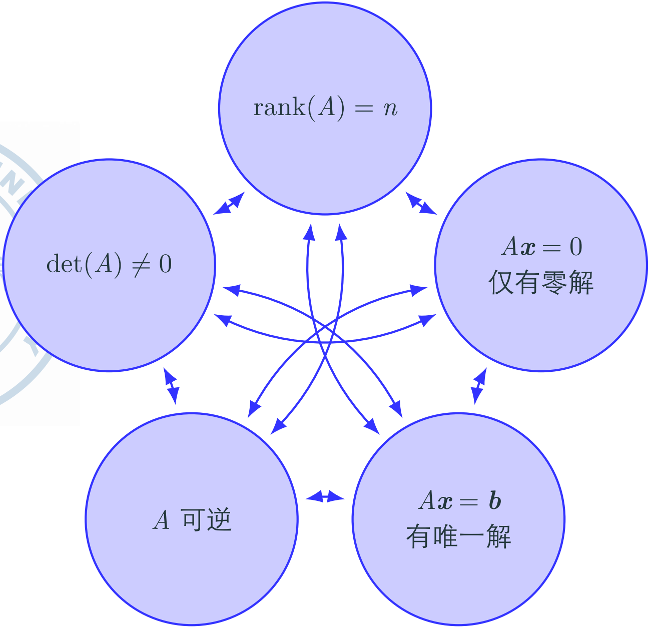
因此，矩阵 A 可逆的充要条件是 $\det(A) \neq 0$。

进一步地，我们可以建立矩阵的可逆性和方程组解的存在性之间的联系:

- 若矩阵 A 可逆，则方程组 $A\boldsymbol{x} = \boldsymbol{b}$ 有唯一解 $\boldsymbol{x} = A^{-1}\boldsymbol{b}$。
- 若矩阵 A 不可逆，则方程组 $A\boldsymbol{x} = \boldsymbol{b}$ 无解或有无穷多解。

[11]这只是一种理解方法，证明较为复杂。

重要结论回顾



$\operatorname{rank}(A) = n$
$\det(A) \neq 0$
$A\boldsymbol{x} = 0$
仅有零解
A 可逆
$A\boldsymbol{x} = \boldsymbol{b}$
有唯一解

特征值与特征向量

变换中的不变量

考虑线性变换 $A = \begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix}$，它改变了二维平面中大多数向量的方向，但将向量 $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ 变换为 $\begin{pmatrix} 3 \\ 3 \end{pmatrix}$；将向量 $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ 变换为 $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$。

事实上，在许多线性变换中，存在一些特殊的向量，它们在变换后只改变了长度，而方向保持不变。我们将这些在变换后方向不变的向量称为**特征向量**(eigenvector)，变换后它们的长度相比原向量长度的倍数称为**特征值**(eigenvalue)。设特征向量为 $\boldsymbol{x}$，特征值为 λ，则满足$A\boldsymbol{x} = \lambda\boldsymbol{x}$。

如果 $A\boldsymbol{x} = \lambda\boldsymbol{x}$ 成立，则：

- $A^2\boldsymbol{x} = \lambda^2\boldsymbol{x} \Longrightarrow A^n\boldsymbol{x} = \lambda^n\boldsymbol{x}$
- $A^{-1}\boldsymbol{x} = \lambda^{-1}\boldsymbol{x}$
- $(A + cI)\boldsymbol{x} = (\lambda + c)\boldsymbol{x}$

$$A\boldsymbol{x} = \lambda\boldsymbol{x} \Leftrightarrow (A - \lambda I)\boldsymbol{x} = \mathbf{0} \tag{15}$$

这是一个齐次线性方程组，称为**特征方程**，为了得到非零的特征向量[12]，需要 $\det(A - \lambda I) = 0$。

以 2 阶矩阵为例，设 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$，则

$$\begin{aligned} \det(A - \lambda I) &= \det \begin{pmatrix} a - \lambda & b \\ c & d - \lambda \end{pmatrix} \\ &= (a - \lambda)(d - \lambda) - bc \\ &= \lambda^2 - (a + d)\lambda + (ad - bc) \end{aligned} \tag{16}$$

式 (16) 称为 A 的**特征多项式**(characteristic polynomial)，其根即为 A 的特征值。

[12]零向量的方向是任意的，作为特征向量没有意义。

求解特征向量示例

设矩阵 $A = \begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix}$，下面求解其特征值和特征向量。

1. 求特征方程 $\det(A - \lambda I) = 0$：

$$(2-\lambda)(3-\lambda) = 0 \quad \Rightarrow \quad \lambda = 2,\ 3$$

2. 对 $\lambda = 2$：
   - 解 $(A - 2I)\boldsymbol{v} = 0$：$\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{0}$
   - 方程组 $y = 0$，得特征向量 $\boldsymbol{v}_1 = k\begin{pmatrix} 1 \\ 0 \end{pmatrix}$，$k \neq 0$
3. 对 $\lambda = 3$：
   - 解 $(A - 3I)\boldsymbol{v} = 0$：$\begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{0}$
   - 方程 $-x + y = 0$，得特征向量 $\boldsymbol{v}_2 = k\begin{pmatrix} 1 \\ 1 \end{pmatrix}$，$k \neq 0$

结论：示例中的 $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ 和 $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ 分别为对应 $\lambda = 2$ 和 $\lambda = 3$ 的特征向量。注意特征向量不唯一，成比例的向量都是特征向量，可以差常数倍。

特征向量是特定变换下最“自然”的基

如果我们不知道矩阵 A 的具体形式，只知道其对特征向量（即特征值），我们是否可以知道任何一个其他点，如 $(2,1)$，经由 A 变换后的结果？

由于特征向量 $\begin{pmatrix}1\\0\end{pmatrix}$ 和 $\begin{pmatrix}1\\1\end{pmatrix}$ 是无关的，因此它们可以线性表出任何二维向量。容易看出，

$$\begin{pmatrix}2\\1\end{pmatrix}=\begin{pmatrix}1\\0\end{pmatrix}+\begin{pmatrix}1\\1\end{pmatrix} \tag{17}$$

在经过 A 变换后，特征向量分别变为 $\begin{pmatrix}2\\0\end{pmatrix}$ 和 $\begin{pmatrix}3\\3\end{pmatrix}$，因此

$$A\begin{pmatrix}2\\1\end{pmatrix}=\begin{pmatrix}2\\0\end{pmatrix}+\begin{pmatrix}3\\3\end{pmatrix}=\begin{pmatrix}5\\3\end{pmatrix} \tag{18}$$

验证一下结果，直接计算 $A\begin{pmatrix}2\\1\end{pmatrix}$，发现结果与上式一致：

$$A\begin{pmatrix}2\\1\end{pmatrix}=\begin{pmatrix}2&1\\0&3\end{pmatrix}\begin{pmatrix}2\\1\end{pmatrix}=\begin{pmatrix}5\\3\end{pmatrix} \tag{19}$$

基变换

在式 (17) 中，我们将向量 $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ 分解为了特征向量的线性组合。虽然这个例子很简单，但对于一般的向量，这种分解往往不直接。这个过程的本质是基的转换，换了一个视角看同一个向量。

事实上，我们能写出任何 2 维向量 (x, y) 的坐标，是因为我们默认选择了标准基，即 $\boldsymbol{w} = x\boldsymbol{i} + y\boldsymbol{j}$。如果我们希望选择特征向量作为基，那么应该寻找使得 $\boldsymbol{w} = x'\boldsymbol{v}_1 + y'\boldsymbol{v}_2$ 成立的 x' 和 y' 作为在特征向量基下 $\boldsymbol{w}$ 的新坐标。写成等式：

$$x\boldsymbol{i} + y\boldsymbol{j} = \boldsymbol{w} = x'\boldsymbol{v}_1 + y'\boldsymbol{v}_2 \Longrightarrow x\begin{pmatrix} 1 \\ 0 \end{pmatrix} + y\begin{pmatrix} 1 \\ 0 \end{pmatrix} = x'\begin{pmatrix} 1 \\ 0 \end{pmatrix} + y'\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

注意，这里两个等式之间的 $\Longrightarrow$ 成立的关键在于用标准基下特征向量的坐标表达了特征向量。这是一种以标准基为中心的表达，目的是将特征向量基的视角中看到的“翻译”为标准基的视角中看到的。进一步写成矩阵形式：

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x' \\ y' \end{pmatrix}$$

我们可以得到：

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{-1}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = P^{-1}\begin{pmatrix} x \\ y \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

基变换

我们将矩阵 P 称为**基变换矩阵**，其列向量是特征向量在标准基下的坐标。矩阵 P 将特征向量基下表示的向量重新表示为标准基下的线性组合（即标准基下的坐标）；矩阵 P^{-1} 将标准基下表示的向量重新表示为特征向量基下的线性组合（即特征向量基下的坐标）[13]。再来看看矩阵 P 的意义：

$$\begin{pmatrix} x \\ y \end{pmatrix} = P \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}$$ 将特征向量基的语言翻译为标准基的语言

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = P^{-1} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} x \\ y \end{pmatrix}$$ 将标准基的语言翻译为特征向量基的语言

[13]更多关于基变换的理解和动画演示，参考 3Blue1Brown 的讲解：
https://youtu.be/P2LTAU01TdA?si=Uugs8Xm3h_5M46LS

特征分解

对于任何一般的向量，我们都可以这样做:

- 将向量表示为特征向量的线性组合
- 将每个特征向量缩放特征值倍
- 按照之前线性组合的系数组合变换后的特征向量

现在我们称使用矩阵表示这个过程，首先利用矩阵 P^{-1} 将任意向量分解为特征向量的线性组合，然后利用对角矩阵 $\Lambda = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ 将每个特征向量缩放特征值倍，最后利用矩阵 P 将变换后的特征向量重新表示为标准基下的线性组合。

因此，可以将 A 分解为:

$$A = P\Lambda P^{-1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{-1} \tag{20}$$

小结

矩阵的转置

矩阵的**转置**是将矩阵的行和列互换。

例如，矩阵 $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ 的转置记作 A^T:

$$A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

沿着从左上角到右下角 $45°$ 的直线翻转矩阵，即可得到矩阵的转置。这条直线叫做矩阵的**主对角线**。

对矩阵 A, B, 我们有 $(AB)^T = B^T A^T$。（提示：想一下矩阵相乘的尺寸对齐关系）

矩阵相乘的逆

对可逆矩阵 A, B, 我们有 $(AB)^{-1} = B^{-1}A^{-1}$。（提示：套定义和原矩阵相乘就可以证明）

回顾

- 我们从向量开始，回顾了向量的加法、数乘、内积运算。
- 之后引入了线性组合的概念，从而引入了线性空间的概念。
- 之后我们讨论了线性变换，并借此引入矩阵的概念，并进一步介绍了矩阵和向量、矩阵和矩阵的乘法，以及行列式的概念。
- 通过线性变换的逆问题，我们引入了方程组的矩阵形式、矩阵的秩以及逆矩阵的概念。
- 之后我们讨论了方程组的解、矩阵的秩、矩阵的可逆性、方阵的行列式之间的关系。
- 最后我们通过“变换中的不变量”引入特征值和特征向量，并指出特征分解的含义是先用特征向量表示标准基中的任意向量，进一步通过特征值缩放特征向量后将结果映射回标准基中。

推荐资源

- Gilbert Strang 教授的教材 *Introduction to Linear Algebra*;
- Gilbert Strang 教授的 MIT 18.06 Linear Algebra 课程视频 (https://youtu.be/7UJ4CFRGd-U?si=5EZDUaXEqaJoYi9G);
- 高中数学选修 4-2 矩阵与变换;
- 3Blue1Brown 线性代数课程 (https://www.3blue1brown.com/topics/linear-algebra);
- UP 主：漫士沉思录 (https://space.bilibili.com/266765166) 的 “无痛线代” 系列视频.