



Makiror 线性代数随笔【1】

引言

说好写计算机的...真的想不到什么好写了，要说我最近比较多看的是手抓饼线代，那就顺手写个线代基础的笔记吧，碍于时间问题本文内容并没有很多，以后有机会会继续完善。

本文有初中数学基础即可，重点在于理清一些概念，对内容的解释和应用不会过度偏到某个领域。

我想从一个比较轻松的例子中引入这个话题，就是，咱小学四年级时有提到『鸡兔同笼』问题，然后好像是六年级有二元一次方程组反正我记不清，这些各个未知变量呈线性关系的都是，反正线性代数这种东西逃不掉的，你小学时的一元一次方程是线代，家楼下的手抓饼也是线代，线代无处不在。

感谢数学大佬过客的锐评和提供的写作经验，这使我我写这篇文章时特别注意了一些问题，以及在此期间尝试调整文章结构，我真的谢谢你

考虑一个简单的二元一次方程组：

$$\begin{cases} 2x + 3y = 10 \\ x - y = 2 \end{cases}$$

这是一个线性方程组，因为它的

严格来说，线性方程组形如：

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

理所应当，线性方程组的解是一组满足所有方程的数值，当然它也能无解或有无穷多的解，我们要考虑的就是如何处理它们。

那现在，正文开始。

向量与矩阵

向量定义和表示

向量 (Vector)，是一个具有大小和方向的『量』，除了数学，它在很多学科也是一个基本概念，在物理学等学科中它会被称为矢量，用于描述物理量等。本文『线性代数』我们需要讨论的就是向量空间（线性空间），而向量是向

量空间的基本元素。

向量通常表示为一个有序的集合，其中包含一组有限个数的实数或复数元素，它可以有不同的维度，维度指的是向量中有多少个分量，例如，二维向量有两个分量 x 和 y ，表示为 $[x, y]$ ，这种水平排列的我们称为横向量，对应的，一个纵向量就是垂直排列的向量，例如 $\begin{bmatrix} x \\ y \end{bmatrix}$ 。

从几何性质和表示来说，因为向量具有方向，所以有时我们用箭头来表示一个向量例如 \vec{v} 。综上所述，我们可以将一个向量写成 $\vec{v} = (x, y, z)$ ，再分成横向量 and 纵向量可以表示为：

$$\vec{v} = [x, y, z]$$

$$\vec{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

消歧义：Wiki将其称为列（column）向量或行（row）向量，但是由于这个表示在大陆与其他地区的用法存在差异，所以此处使用横向量 and 纵向量这样没有歧义的叫法

范数

范数（Norm），又叫模，你可以想象向量是一个空间中的一个箭头（二维、三维甚至更高维），模就是这个箭头的长度，符号为 $\|\vec{v}\|$ ，范数应该满足以下条件

- 范数的值为非负实数
- 对于向量 \vec{v} 和任意实数 α ，满足 $\|\alpha\vec{v}\| \leq |\alpha| * \|\vec{v}\|$ ，这叫齐次性，这意味着向量缩放，也就是乘一个实数，其范数会相应缩放。
- 范数满足 $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$ ，这也叫三角不等式，这意味着两个向量的范数之和不会超过它们的向量之和的范数。

范数有很多种类型，最常用的有以下两种，在 n 维的空间中：

- 曼哈顿范数（L1）：L1范数是向量元素绝对值之和， $\|\vec{v}\|_1 = |a_1| + |a_2| + \dots + |a_n|$ ，也就是 $\|\vec{v}\|_1 = \sum_{i=1}^n |x_i|$
- 欧几里德范数（L2）：L2范数是向量元素平方和的平方根， $\|\vec{v}\|_2 = \sqrt{a_1^2 + a_2^2 + \dots + a_n^2}$ ，也就是 $\|\vec{v}\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$

另外，还有L0范数用来表示该向量中非零元素的个数，但是它不满足齐次性和三角不等式，所以它并不算一个标准的范数。

一般情况，我们这样对向量进行运算：

- 加减法运算：同维度的向量可以进行加减法运算，将两个相同维度的向量逐元素相加/相减即可
- 数乘：向量与一个实数相乘时，将向量的每个元素与该实数相乘即可。

内积和叉积可能需要解释一下，所以此处开个小标题。

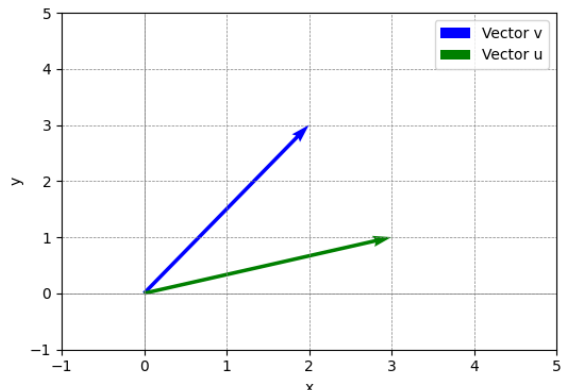
内积

同维度的向量可以相乘，内积就是逐元素相乘，再将结果相加得到一个标量，对于向量 v, u 在 n 维实数向量空间 \mathbb{R}^n ，它们的内积可以表示为 $\langle v, u \rangle$ ，或者：

$$v \cdot u = v_1 u_1 + v_2 u_2 + \cdots + v_n u_n$$

例如

$$\begin{bmatrix} 2 \\ 3 \end{bmatrix} \times \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 2 * 3 + 3 * 1 = 9$$



内积结果的正负可以判断两个向量在同一方向上还是相反方向上，图例中的两个向量内积为正，说明它们在同一方向上，且夹角小于90度，如果内积为0则代表两个向量垂直，为负则表示两个向量在相反方向且夹角大于90度。

内积与向量范数和夹角存在关系，若有向量 v, u ， θ 是这两向量的夹角，则 $v \cdot u = \|v\| * \|u\| * \cos \theta$ ，且内积满足交换律 $v \cdot u = u \cdot v$ 。同理地我们可以用来判断两个向量位置关系和夹角大小。

另外，内积空间内可以理解是一种，具有内积运算的向量空间，例如我们在前面提及的 n 维实数向量空间 \mathbb{R}^n 。

叉积

叉积是一种适用于三维向量的运算方式，结果是一个向量，若有向量 v, u ，它们进行叉积的结果是一个垂直于原始向量 v, u 的向量，因为二维空间中不存在垂直于两个向量的唯一向量，所以叉积仅适用于三维向量。

我们可以使用3阶行列式来表示两个三维向量的叉积，并根据其展开规则展开得到（关于行列式，后面会提到）：

$$v \times u = \begin{vmatrix} i & j & k \\ x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{vmatrix} = (y_1 z_2 - z_1 y_2)i - (x_1 z_2 - z_1 x_2)j + (x_1 y_2 - y_1 x_2)k$$

其中， i, j, k 是三维空间中的单位向量，分别表示 x 、 y 、 z 轴方向，它们等于

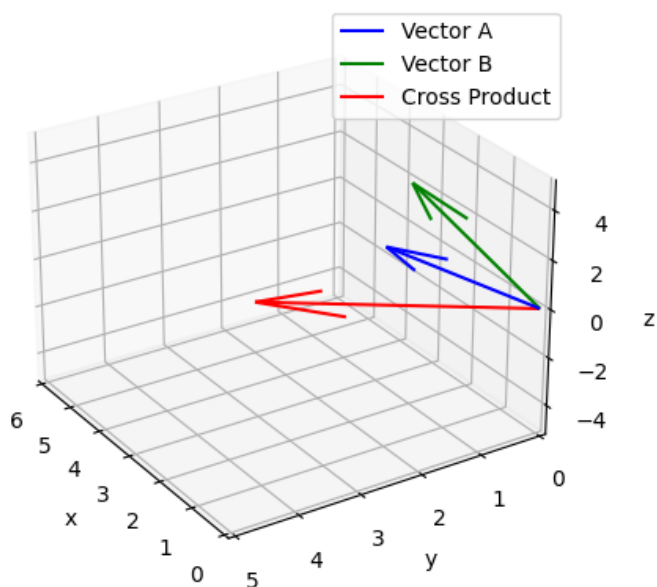
$$i = (1, 0, 0)$$

$$j = (0, 1, 0)$$

$$k = (0, 0, 1)$$

紧接着，上述公式的单位向量被简化后，我们得到叉积的展开式为：

$$v \times u = (y_1 z_2 - z_1 y_2, z_1 x_2 - x_1 z_2, x_1 y_2 - x_2 y_1)$$



我们关注叉积结果向量的方向，垂直于原始向量 v, u 所在的平面，有趣的是，叉积向量的方向遵循右手法则，你可以尝试将右手食指指向 v 的方向，中指指向 u 的方向，则此时右手拇指就是叉积向量的方向（比较难想象的话可以看旁边的图片）

叉积的长度等于原始向量 v, u 所构成的平行四边形的面积，你会发现， $v \times u$ 大小与原始向量 v, u 的夹角以及它们的长度有关，而当它们夹角为0时（即共线），这个平行四边形面积为0，叉积结果也是一个零向量。

此外，叉积满足反交换律，即：

$$v \times u = -(u \times v)$$

这意味着两个向量的叉积的顺序交换会导致结果的方向反向，你还是可以拿你的右手来理解，如果你反转右手的方向，那结果方向就会相反。

矩阵定义和表示

再看矩阵（Matrix），矩阵是一个由元素按照矩形排列的矩形阵列，元素可以是数值、函数等，一个 $m \times n$ （这被称为矩阵的维度）的矩阵，有 m 行， n 列，可以表示为：

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}$$

严格来说，向量是矩阵的一种特殊情况，向量可以被视为只有一行/列的特殊矩阵，或者说是一个元素按照一定顺序排列成的矩阵，此外还有一些特殊的矩阵：

- 方阵：行数和列数相等的矩阵
- 对角矩阵：除了主对角线上的元素外，其它元素都为零的矩阵
- 零矩阵：所有元素都为零的矩阵
- 单位矩阵：主对角线上的元素为1，其它元素都为零的对角矩阵

主对角线是指一个矩阵中从左上角到右下角的对角线，主对角线上的元素具有相同的行号和列号，这个概念对于我们后面的分析而言很重要。

我们可以这样对矩阵进行常见的运算：

- 加减法运算：同维度的矩阵可以进行加减法运算，将两个相同维度的矩阵逐元素相加/相减即可。
- 数乘：矩阵与一个实数相乘时，将向量的每个元素与该实数相乘即可。
- 矩阵乘法：两个矩阵 A, B 相乘，则将矩阵 A 的行与 B 的列进行内积运算，例如 $A = m \times n, B = n \times p$ ，则它们乘积矩阵 $C = m \times p$

行列式

行列式是对一个方阵的计算，得到的是一个标量，一般我们表示为 $\det(A)$ 或 $|A|$ 。二阶行列式对应的是 2×2 的方阵，它的计算顺序为 $ad - bc$ ，形式为：

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

同理，三阶行列式对应的是 3×3 的方阵，计算顺序为 $a(ei - fh) - b(di - fg) + c(dh - eg)$

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$$

这种展开行列式的方式叫做拉普拉斯展开，也叫代数余子式展开。首先，余子式的定义是，在行列式中将 a_{ij} 的 i 行 j 列划去后的部分，叫 a_{ij} 的余子式。阿就例如上面那个三阶行列式， f 的余子式为：

$$M_{23} = \begin{vmatrix} a & b & \square \\ \square & \square & \square \\ g & h & \square \end{vmatrix} = \begin{vmatrix} a & b \\ g & h \end{vmatrix}$$

代数余子式就是，在这个余子式基础上乘以 $(-1)^{i+j}$ ，表示为：

$$C_{ij} = (-1)^{i+j} \cdot M_{ij}$$

例如，之前那个 M_{23} 对应的 $2 + 3 = 5$ ，而 $(-1)^5 = -1$ ，所以 f 的代数余子式为：

$$C_{23} = (-1) \cdot \begin{vmatrix} a & b \\ g & h \end{vmatrix} = - \begin{vmatrix} a & b \\ g & h \end{vmatrix}$$

在拉普拉斯展开中， $n \times n$ 的行列式可以被展开成行各元素和其代数余子式的乘积之和，若该方阵为 A ， i 是基准行/列的索引，对于任意 $i, j \in \{1 \cdots n\}$ ，如果你选择以行进行展开则：

$$|A| = \sum_{j=1}^n a_{ij} \cdot (-1)^{i+j} \cdot M_{ij}$$

要注意的是以 i 行展开的情况，求和索引是 j ，反之同理。

转置和逆矩阵

转置是将矩阵的行和列进行对调得到一个新的矩阵。考虑一个 $m \times n$ 的矩阵 A ，其转置记作 A^\top 或 A' ，转置后

变成 $n \times m$ 。

考虑方阵 A ，它与另一个方阵 B 相乘得到的是单位矩阵，交换顺序相乘得到的也是单位矩阵，也就是 $A * B = B * A = I$ ，则 B 是 A 的逆矩阵，记作 A^{-1} 。要注意的是，一个可逆方阵的逆矩阵是唯一的，且它的逆矩阵的逆矩阵是它本身，也就是 $(A^{-1})^{-1} = A$ 。

强调一下矩阵乘积与行列式的关系，如果有两个矩阵 A, B ，它们的乘积为 $A * B = C$ ，则它们的行列式 $\det(A), \det(B)$ 等于 C 的行列式 $\det(C)$ 。如果一个方阵的行列式为0，则它是不可逆的，这叫奇异矩阵，因为行列式为0意味着它乘任何矩阵的行列式都不会等于1，不可能满足上述条件，所以无法找到一个唯一的解来求解逆矩阵。例如矩阵 $\begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix}$ ，表示为二阶行列式并求值：

$$\begin{vmatrix} 3 & 4 \\ 6 & 8 \end{vmatrix} = 3 * 8 - 4 * 6 = 0$$

所以该矩阵是奇异矩阵，是不可逆的。

线性方程组

矩阵和向量可以用来紧凑地表示线性方程组，考虑一个线性方程组形式如下：

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

则可以以矩阵和向量的形式表示为：

$$Ax = b$$

其中， A 是一个 $m \times n$ 的矩阵， x 是一个 $n \times 1$ 的列向量， b 是一个 $m \times 1$ 的列向量：

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}, x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

最直接地，如果矩阵 A 是可逆的，则我们可以直接通过左乘 A^{-1} 来求解方程，即

$$x = A^{-1}b$$

本文不写求解方程组的做题技巧，所以其他情况，例如什么高斯消元法的就不写了。

线性变换

向量空间

向量空间，也叫线性空间，是一组满足一定条件的向量集合。向量空间要求里面的向量进行加减或缩放后仍然属于它（也就是进行加减和数乘），这被称为封闭性。向量空间中存在一个特殊的零向量，它与任何向量相加不改变其值，每个向量在向量空间中有唯一的反向量，使得它与原向量相加等于零向量。这样描述太抽象了，我们还是先看定义：

向量空间 V 是在一个域 F （通常为实数域 \mathbb{R} 或复数域 \mathbb{C} 等）上的非空集合，对于其中任意两个向量 u, v 以及任意标量 α, β ，满足：

- $\forall u, v \in V, u + v \in V$ ，这意味着向量空间里的向量进行加法运算后仍属于该向量空间，这叫做加法封闭性。
- $\forall u, v, w \in V, (u + v) + w = u + (v + w), \forall u, v \in V, u + v = v + u$ ，向量空间内的向量具有加法结合性和交换性
- $\forall u \in V, \alpha u \in V$ ，这意味着向量空间里的向量进行数乘运算后仍属于该向量空间，这叫做数乘封闭性。
- $\exists 0 \in V, \forall u \in V, u + 0 = u$ ，这是向量空间中特殊的零向量，它与任何向量相加不改变其值。
- $\forall u \in V, \exists -u \in V$ ，使得 $u + (-u) = 0$ ，这意味着每个向量在向量空间中有唯一一个反向量，使其与这个向量相加等于零向量。

我觉得应该解释一下的就是这个域 F ，这叫标量域，定义了向量空间中用于进行数乘运算的元素（即标量），标量乘以向量可以得到一个新的向量（数乘操作），你可以理解为这是给向量空间提供了一个数学结构。

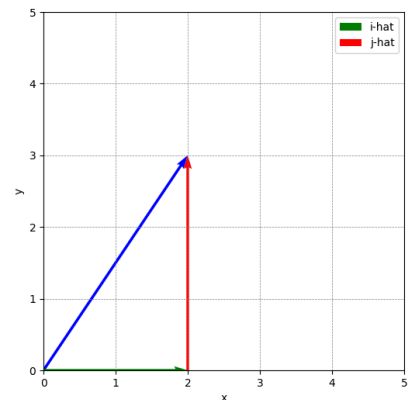
基与维数

基是向量空间中一组线性无关的向量，通过线性组合可以表示出向量空间中的所有向量，它使得我们可以用较少的向量来表示整个向量空间。我们接着之前的向量空间写定义，存在向量 $v_1, v_2, \dots, v_n \in V$ ，且它们线性无关，不存在非零标量 c_1, c_2, \dots, c_n 使得 $c_1 v_1 + c_2 v_2 + \dots + c_n v_n = 0$ ，我们可以通过合适的标量 c_1, c_2, \dots, c_n 使得 $c_1 v_1 + c_2 v_2 + \dots + c_n v_n = v$ 。这意味着对于向量空间 V 中的任意向量 v ，都可以通过合适的标量来表示，你可以理解为基为向量空间提供了一种坐标系使得我们可以用有限的标量来描述整个向量空间中的向量。

图中绿色和红色的向量，就是我们所说的基，分别写为 \hat{i} 和 \hat{j} ，这个蓝色的向量 $\vec{v} = (2)\hat{i} + (3)\hat{j}$ ，无论正负，向量空间的任何向量都能使用基向量来描述。

$$\hat{i} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \hat{j} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\vec{v} = (2)\hat{i} + (3)\hat{j} = 2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$



维数是衡量向量空间的大小，即其所包含的独立方向的数量，也就是基的数量，例如图例中的维数是2，有两个基向

量，如果要用基表示三维的向量，那我们对应的就要使用三个基。

线性变换定义和表示

线性变换在满足加法和数乘性质的前提下，描述了一个向量空间中的向量如何被映射到另一个向量空间中的向量，因此你可以将它理解为一个特殊的函数，输入一个变量得到另一个变量，从几何上看，可能是向量旋转、缩放、投影之类的动态的操作。

我们先考虑矩阵表示线性变换，考虑 V 和 W 都是 n 维向量空间，那么一个线性变换 $T: V \rightarrow W$ 可以用一个 $m \times n$ 的矩阵 A 来表示。对于任意向量 $v \in V$ ，其线性变换结果 Tv 可以表示为

$$T(v) = Av$$

所以现在我们要考虑的就是，这个表示线性变换的矩阵 A 具体来说如何表示。因为向量空间的任何向量都能使用基向量来描述，所以我们要描述向量的变化也可以使用基向量，此处我写了一些常见的，二维平面上的线性变换的矩阵表示。

平移

点 (x, y) 被平移到 (x', y') 上，则矩阵 A 表示变换为：

$$A = \begin{bmatrix} 1 & 0 & dx \\ 0 & 1 & dy \end{bmatrix}$$

其中， dx, dy 分别表示该点在 x, y 方向的位移量，所以，我们要求结果的向量 \vec{w} 即：

$$\vec{w} = \begin{bmatrix} 1 & 0 & dx \\ 0 & 1 & dy \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \cdot x + 0 \cdot y + dx \\ 0 \cdot x + 1 \cdot y + dy \end{bmatrix} = \begin{bmatrix} x + dx \\ y + dy \end{bmatrix}$$

运算方式后面不再重复，仅提供变换的矩阵表示。

缩放

同理，缩放变换的矩阵形式如下，其中 Cx, Cy 分别是 x, y 方向的缩放因子：

$$\begin{bmatrix} Cx & 0 \\ 0 & Cy \end{bmatrix}$$

旋转

旋转的矩阵形式如下，其中 θ 是旋转的角度：

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

剪切

剪切变换分为水平剪切和垂直剪切，水平剪切是通过改变 y 轴上的坐标值，使得所有点在 y 轴方向上平移，垂直同理，是 x 轴，其中 a 是物体在坐标上的偏移量，它的正负值也会影响到物体倾斜的方向正负，分别表示为：

$$\begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 \\ a & 1 \end{bmatrix}$$

镜像

反射变换是将向量关于某个轴对称反转，水平和垂直分别表示为

$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

特征方程

此处应该先明确的是特征值与特征向量的概念，我们可以将特征值理解为将向量进行缩放操作的因子，是一个标量。在定义上，对于一个 $n \times n$ 的方阵 A ，一个非零向量 v ，以及特征值 λ ，满足：

$$Av = \lambda v$$

那 λ 就是特征值， v 是特征向量。

特征方程无非就是求解特征值的方程，对于一个 $n \times n$ 的方阵 A ，一个 n 阶的单位矩阵 I ，我们可以对上式做一个简单的变形，变成 $(A - \lambda I)v = 0$ ，展开后你会得到一个齐次线性方程组，为了使这个方程有非零解， $A - \lambda I$ 必须为零，所以特征方程表示为：

$$|A - \lambda I| = 0$$

对角化

对角化是将一个矩阵转化为对角矩阵（即除了主对角线上的元素外，其它元素都为零的矩阵）的过程。在定义上。对于一个 $n \times n$ 的方阵 A ，存在一个可逆矩阵 P 使得 P^{-1} 与 A 相乘得到一个对角矩阵 D ，也就是：

$$D = P^{-1}AP$$

实现对角化的前提是，原矩阵 A 要有 n 个线性无关的特征向量，并且对于每个特征值 λ 对应的特征向量是构成矩阵 P 的列向量。 D 形如：

$$\begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

为什么要这样做？（又来了，愚蠢的自问自答），通过对角化，我们可以将线性变换表示为对角矩阵，其中特征向量构成了新的基向量，这可以简化变换的理解和分析。

核与值域

特征向量是在线性变换中只发生缩放而不改变方向的向量，而线性变换还有一种特殊情况就是，通过线性变换时的向量被映射为零向量。核是指使得线性变换结果为零向量的所有输入向量的集合，核是一个子空间，这意味着它也满足加法和标量乘法的封闭性。考虑一个线性变换 $T: V \rightarrow W$ ，其中 V 和 W 是向量空间， $\mathbf{0}$ 是零向量，核是针对线性变换 T 的概念，表示为：

$$\ker(T) = \{v \in V | T(v) = \mathbf{0}\}$$

对于一个线性变换 $T: V \rightarrow W$ ，其中 V 是向量空间，值域是线性变换的输出空间，其中包含了所有可能的输出向量，表示为：

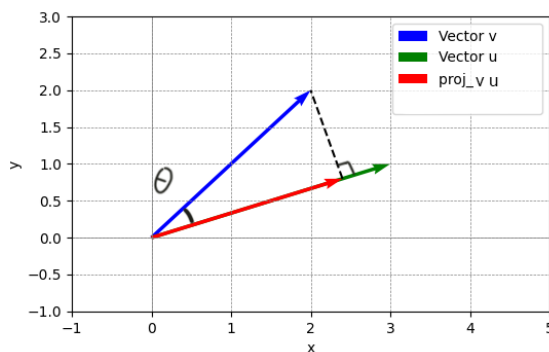
$$\text{Im}(T) = \{T(v) | v \in V\}$$

值域的维度和核的维度的和等于输入向量空间的维度，这叫秩—零化度定理。对于一个线性变换 $T: V \rightarrow W$ ， $\text{rank}(T)$ 是值域的维度（就是秩）， $\text{nullity}(T)$ 是核的维度， $\dim(V)$ 是输入向量空间 V 的维度，即：

$$\text{rank}(T) + \text{nullity}(T) = \dim(V)$$

投影和正交化

向量投影



向量投影是一个向量在另一个向量或子空间上的映射，如图，向量 v 投影到向量 u 上，则表示为 $\text{proj}_v u$ 。我们可以通过向量的内积和范数来计算投影向量：

$$\text{proj}_u \mathbf{v} = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|^2} \mathbf{u}$$

注意，该公式的范数是L2范数。这个公式可以理解为， $\frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|^2}$ 求出了 u 在 v 方向上分量的大小，也就是确定了投影的长度，再与 u 相乘，表示其方向，得到的就是向量 u 在 v 方向

的投影向量。

Gram-Schmidt 正交化

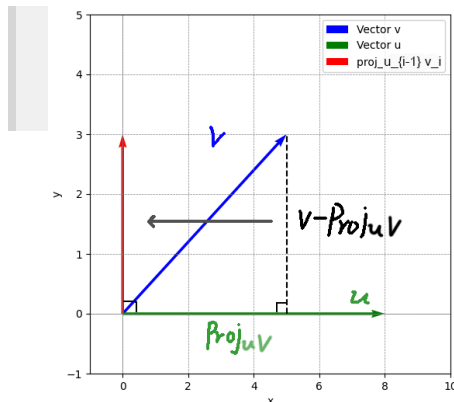
正交向量是指两个向量的内积为0，也就是在空间中垂直。正交化就是将一组向量转为一组正交向量的过程，其目的是使向量之间没有冗余信息。例如，我们可以通过 Gram-Schmidt 正交化从线性无关的向量组构建出一组正交基向量，基向量可以表示任何向量，而我们一般更常用单位直角坐标系，所以我们构建出一个正交基向量就能简化向量的表示。

有一组线性无关的向量 $\{v_1, v_2, \dots, v_n\}$ ，我们初始化一个向量 $u_1 = v_1$ ，作为正交基第一个向量， $\{u_1, u_2, \dots, u_{i-1}\}$ 。然后对于第 i 个向量 v_i ，计算它在它前一个的正交向量 u_{i-1} 的投影：

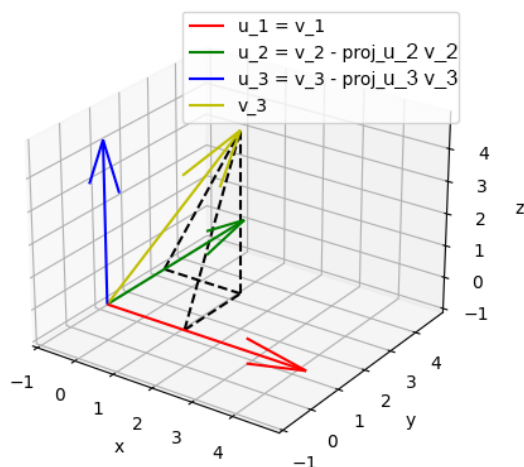
$$\text{proj}_{u_{i-1}} v_i = (v_i \cdot u_1)u_1 + (v_i \cdot u_2)u_2 + \dots + (v_i \cdot u_{i-1})u_{i-1}$$

再将当前向量减去它前一个向量的投影，使得 $w_i = v_i - \text{proj}_{u_{i-1}} v_i$ ，最后将其归一化，即使得 $u_i = \frac{w_i}{\|w_i\|}$ ，最后我们得到的 $\{u_1, u_2, \dots, u_{i-1}\}$ 就是我们需要的正交基向量。

有些时候我不得不承认，线代这种东西的理解是很依赖几何直观的，就上面这段描述，可能很难理解这是在做什么，所以稍微加以解释一下。



为1的正交基向量。



现在是2023年8月11日凌晨一点半，小酌了瓶珠啤睡不着爬起来写文章。此处继续用上面向量投影的图，里面的备注是我懒得打label就直接ps手写的标记，可以用于理解 Gram-Schmidt 正交化

在二维平面上，这两个线性无关的向量可以构建一个平面直角坐标系，也就是我们前面说的一组基。从图上看就是，投影、挪过去，而三维空间同理，只是要分别投影两个，而第二个需要分别在前两个向量上投影并减去两个投影的和（所以前面有涉及求和与迭代操作，前面为了清晰就没有写带求和符的）

归一化的意义就显而易见了，将向量除以其范数，确保我们得到的是长度为1的正交基向量。

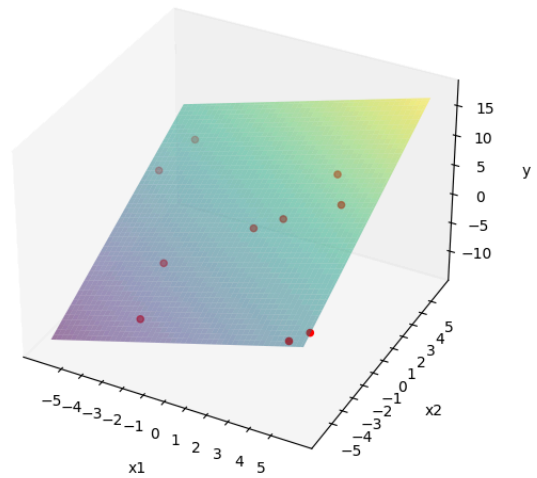
最小二乘法

最小二乘法在 *Machine Learning* 中的应用我在上篇文章有详细解释（包括单变量和多变量的）和推导过程，文末有链接。

我觉得本文的内容不应该依赖于ML系列的博客，所以这里使用那篇博客的一些例图再讲一次。

在现实中，一组线性方程组可能无精确解，存在一定的误差（前提是自变量与因变量存在线性关系），简单的说，最小二乘法的目标就是使得误差的总和最小，找到方程组最合适的解。图是一个最小二乘法的二元线性回归，拟合出一个最佳平面的例子

在ML文章第一篇我有将一元线性回归和多元线性回归的公式推一遍，没有很复杂，只是单纯略微涉及导数知识和矩阵知识可以通过近两篇文章补一下。



在现实中经常会有方程没有精确解的情况，有一个不存在确切解的方程组 $Ax = b$ ，我们求解时最好的方法就是找到一个 b 尽可能的接近 Ax 。如果我们的残差是 $\epsilon = b - Ax$ ，我们的目标就是找到使 $\|\epsilon\|$ 最小（注意这里通常是L2范数）的 x 。

为了防止正负误差互相抵消，最小二乘法的核心是使得残差平方和最小，一般情况下我们将残差 ϵ 视为列向量，则我们先需要对其进行转置操作再相乘，得到的就是残差平方和 S ：

$$\epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \dots \\ \epsilon_n \end{bmatrix} \quad \epsilon^\top = [\epsilon_1, \epsilon_2, \dots, \epsilon_n]$$

$$S = \epsilon^\top \epsilon$$

我们可以通过将 $\epsilon = b - Ax$ 代入 S 进行运算，再将 $\frac{dS}{d\beta}$ 置为0后进行一系列的推导和简化，结论的 x 是：

$$x = (A^\top A)^{-1} A^\top b$$

因为是笔记就不写冗杂的推导过程了，我会将我两个月前写的推导的链接放到文末。

你可以将这理解为一种数据的降维，最小二乘法通过投影从数据中提取主要特征，将数据投影到一个低维子空间，使我们更能够区分不同类别或捕获数据变化的信息。

结尾

其实写文章多少还是会花一点时间，所以我可能会暂时不写知识类的文章了，花时间去学点新东西。

我倒是发现了，要把基础的知识组织好、写好真的不简单，这次虽然撑死只是初高中数学（大概，我对应试教育的进度不是很清楚），但我感到比任何一次都难写，很多时候会感到迷茫和不知所措，真的非常感谢过客大佬在这期间提供的建议。

自己也不知道结尾该写点啥，就这样吧反正。

一些参考链接

因为这篇文章比较即兴，我写的时候并没有像以往一样很勤快地参考资料，所以此处就只放一些相关链接，以及早期我自学线代的书籍

线性代数及其应用: <https://www.zhihu.com/topic/20129721/hot>
Introduction to Linear Algebra（线性代数导论，几个月前浅看过，不知道现在有没有中翻）\

本文相关的推导，不时更新: <https://www.makiror.xyz/pdf/derivation/01.pdf>

MAKIROR gzanan@gmail.com 21:08 15/08/2023

什么事线性代数？	形式纯粹主义	形式中立主义	形式自由主义
	必须是线性的	在一个环上也行	只要能加就行
内容纯粹主义 必须研究代数	 高等代数是线性代数	 文焕代数是线性代数	 椭圆曲线是线性代数
内容中立主义 也可研究其他结构	 泛函分析是线性代数	 代数几何是线性代数	 代数拓扑是线性代数
内容自由主义 不必和数学相关	 在线性环是线性代数	 指环王也是线性代数	 平煎饼难道不是线性代数吗？

什么不事线性代数？：

