线性代数笔记

M.

From Oct. 12 2023
To October 17, 2023

Contents

Chapter 1

Vector, Vector space, and Linear

欢迎来到我们线性代数课程的第一章, 本章我们会从之前所学的对向量的朴素认知出发发展出什么是 vector space，进而延伸到对 linear 的思考和定义。

1.1 Vector and Vector space

在我们高中的学习过程中, 我们已经接触到了一种特殊的量——向量（或矢量)。在过去的学习中我们认为这种量不只有数值大小, 还有方向。我们一般会将他们写成 (a,b,c) 的形式。例如 (6,1,2) 就代表着一个在 x 方向延伸六个单位长度, 在 y 方向延伸一个单位长度, 在 z 方向延伸了两个单位长度的向量。但是这种表述方式并不够规范, 我们也不好表示一组有着某些抽象关系的向量, 例如所有从原点发出的向量。毕竟像 $\{(x, y, z) : x \in \mathbb{R}, y \in \mathbb{R}, z \in \mathbb{R}\}$ 的表达并不算简单, 在我们引入更多坐标轴变量的时候也会更费墨水, 因此我们想要引入一些特殊的符号来更加方便的表述这样的 list (有序列)。

我们首先严谨的定义一下什么是 list 和 list 的 length。

Definition 1.1 我们说 n 个数构成了一个 length 为 n 的 list 当且仅当这 n 个数以一种有序的方式排列在一起组合成一个元素。我们一般将这种有序列写作 $(x_1, x_2, x_3, ..., x_n)$ 。

在此基础上, 我们便可以定义在任意 n 下所有长度为 n 的 list 的集合了我们定义 $\mathbb{R}^n$ 为所有长度为 n 的 list 的集合, 即

Definition 1.2

$$\mathbb{R}^n := \{(x_1, x_2, x_3, ..., x_n) : \forall i \in \mathbb{Z}\ x_i \in \mathbb{R}\}$$

称其中的 x_i 为 the i-th coordinate。

例如:

$$\mathbb{R}^4 := \{(x_1, x_2, x_3, x_4) : \forall i \in \mathbb{Z}\ x_i \in \mathbb{R}\}$$

现在我们非常神奇地发现, 我们之前所学过的所谓 vector 都是属于 $\mathbb{R}^2$ 或 $\mathbb{R}^3$ 的 list。回忆我们曾经学过的向量的加法, 比如 (6,1,2)+(12,1,9)=(6+12,1+1,9+9) 加法的方式就是简单的将不同 coordinate 上的元素分别相加。

那么类比的, 我们就可以定义任意 n 下 $\mathbb{R}^n$ 中元素的加法:

Definition 1.3 在 $\mathbb{R}^n$ 中符号"+"可以由以下公式定义，我们将这种运算称为**加法**:

$$\forall x, y \in \mathbb{R}^n, x + y \ = z \text{ where } z_i = x_i + y_i \forall i$$

我们可以显然的发现, 这里的加法是符合交换律的, 在这里我们就不做证明, 留待读者自证。而根据我们对零的一般感知, 也就是所有数加 0 都是它本身, 我们可以定义在 $\mathbb{R}^n$ 中的 0:

Definition 1.4 我们定义 $\mathbb{R}^n$ 中的 **0** 为:

$$\forall x \in \mathbb{R}^n \ \ x + 0 = x$$

我们可以显然的发现在所有 $\mathbb{R}^n$ 中 **0 都是唯一的并且 0=(0,0,0,···,0)**.

当然, 我们的数学家不愿意止步于此。我们希望将我们熟悉的欧几里得空间也就是 $\mathbb{R}^n$ 中的朴素认知拓展推广到更加抽象的一般集合中。再根据我们对 vector 也就是向量的朴素认知将其推广到一般集合中的一般元素。我们便有了更加抽象和一般化的对 vector space 和 vector 的定义:

Definition 1.5 我们说一个集合 $\mathcal{V}$ 是一个 **vector space** 当且仅当它满足以下所有性质:
1. Commutativity: $\forall \mathbf{u}, \mathbf{v} \in \mathcal{V} \ \mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
2. Associativity: $\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{V} \ \mathbf{u} + (v + w) = (\mathbf{v} + \mathbf{u}) + \mathbf{w}$ and $\forall a, b \in \mathbb{R} \ a(b\mathbf{v}) = (ab)\mathbf{v}$
3. Additive identity: $\exists \mathbf{u} \in \mathcal{V} \ \forall \mathbf{v} \in \mathcal{V} \ \mathbf{v} + \mathbf{u} = \mathbf{v}$ 将**u**记作0
4. Additive inverse: $\forall \mathbf{v} \in \mathcal{V} \ \exists \mathbf{u} \in \mathcal{V} \ \mathbf{v} + \mathbf{u} = 0$
5. Multiplicative identity: $\forall \mathbf{v} \in \mathcal{V} \ 1\mathbf{v} = \mathbf{v}$
6. Distributive properties: $\forall \mathbf{u}, \mathbf{v} \in \mathcal{V} \ \forall a, b \in \mathbb{R} \ a(\mathbf{v} + \mathbf{u}) = a\mathbf{v} + a\mathbf{u} \ (a + b)\mathbf{v} = a\mathbf{v}$ and $+ b\mathbf{v}$

同时定义所有属于 $\mathcal{V}$ 这个 vector space 的元素都是 **vector**。

在这个定义中，我们默认了数乘，也就是 $\forall a \in \mathbb{R} \ \ \forall \mathbf{v} \in \mathcal{V} \ \ a\mathbf{v} \in \mathcal{V}$，的存在性。这也指出当我们说一个集合 $\mathcal{V}$ 是一个 vector space 之前，我们必须先定义这个集合中元素的数乘。

例如:
$\mathbb{R}^n$ 之中，$\forall a \in \mathbb{R} \ \ \forall \mathbf{v} = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$ 我们定义 $a\mathbf{v} = (av_1, av_2, \ldots, av_n)$ 为 $\mathbb{R}$ 中的数乘。
这和我们在高中阶段所学的向量的数乘也是一致的。

我们可以很简单的验证 $\mathbb{R}^n$ 是一个 vector space。

而我们也可以举出一些更加奇怪的 vector space 的例子:
如果我们定义 $P(n) \ = \{\sum_{i=0}^{n} p_i t^i : \forall p_i \in \mathbb{R}\}$ 那么我们可以发现 P(n) 也是一个 vector space。

自然的，我们会思考，像 P(n) 和 $\mathbb{R}^n$ 的这些 vector space 之间会不会有一种内在联系呢？显然我们会发现任意一个 $\sum_{i=0}^{n} p_i t^i$ 的多项式都可以被 $(p_1, p_2, ..., p_n)$ 唯一确定的表示，我们将这种关系称为 **isomorphism**。对于这一关系的定义完善我们先按下不表，继续我们对于 Vector space 本身的探讨。

1.2 Linear Combination and Span

在上一个小节, 我们介绍了什么是 vector space。而在这一章节中, 我们将会从一个重要的概念——linear 出发，探讨我们该如何构建一个 vector space 以及 vector 的一些重要定义。

当我们提及 linear 也就线性的时候，我们第一时间想到的一定会是线性函数（linear function），也就是我们从初中开始接触到的一次函数。根据对线性函数的直观印象，我们可以简单地认为，线

性本质上就是对不同的元素进行加法和数乘的运算。而我们在上一章节中已经定义了 vector space，也就是说，我们可以将线性函数初中学习的最基础的多元一次函数拓展到 vector space 上，看作是对 vector space 中的元素进行加法和数乘的运算。而我们将这种对 vector space 中的多个元素进行加法和数乘的运算称为 **linear combination**。

Definition 1.6 $\forall \mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k \in \mathcal{V}$ $\forall a_1, a_2, ..., a_k \in \mathbb{R}$ 我们称形如 $\sum_{i=1}^{k} a_k \mathbf{v}_k$ 的表达式为 $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ 的一个 **linear combination**。
如果 $\mathbf{u} \in \mathcal{V}$ 是由通过关系式 $\mathbf{u} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + ... + a_n\mathbf{v}_n$ 获得的，那么我们说 $\mathbf{u}$ 可以被 $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ 的 **linear combination** 表示。

例如：
我们可以说 $(6,1,2)$ 可以被 $(1,0,0),(0,1,0),(0,0,1)$ 的 linear combination 表示，因为 $(6,1,2) = 1(1,0,0) + 1(0,1,0) + 2(0,0,1)$。

如果我们将 vector space 的定义视作构建大楼的图纸，将 vector 是为构建大楼的砖块。那么 linear combination 就可以被视为搭建蓝图中大厦所必要的工具（这里我们必须指出，虽然我们在定义什么是 vector 的时候使用了 vector space。但是事实上这并不是必须的，甚至在许多情况下，我们是通过对一组我们感兴趣的元素的线性叠加来构建出我们所希望研究的 vector space 的。值得指出的是，这种循环看起来会产生形式上的问题。但在形式上我们依旧可以说“我们通过某一组‘元素’构建起了一个集合，而这个集合经过验证‘居然神奇的’是一个 vector space。而我们最初用来构建这一 vector space 的元素‘居然’也是这个 vector space 中的元素”这样的过程来避免形式上的问题。我们在之后不会再注意使用这种冗长的避免方式，而是直接将构建的元素称为 vector）。因此，我们引入一个重要的定义——**span**，来严格表示使用一组 vector 来生成一个 vector space 的过程。

Definition 1.7 我们称一个集合 $\mathcal{U}$ 是由一组 vector $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in \mathcal{V}$ **“span”** 的当且仅当集合 $\mathcal{U}$ 中的所有元素恰好是 $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ 的全部 linear combination。我们将这样的 $\mathcal{U}$ 记做 $\text{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$。
用一种更加数学的表示方法，我们可以将 $\text{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ 表示为：

$$\text{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n) = \left\{ \sum_{i=1}^{n} a_k \mathbf{v}_k : \forall a_k \in \mathbb{R} \right\}$$

例如：

- $(6,0,0)$ 和 $(0,12,0)$ 的 span，即 $\text{span}((6,0,0),(0,12,0))$ 是在 $\mathbb{R}^3$ 中的一个平面，即所有 $z = 0$ 的点的集合。
- $f_1(t) = t^2$ 和 $f_2(t) = t^3$ 的 span，即 $\text{span}(f_1, f_2)$ 是在 $P(3)$ 中的一个子空间，即所有次数只有二次项和三次项的多项式的集合。

更进一步，我们观察例子中 span 的线性组合。我们可以发现，这种线性组合可以被表示为一组多元一次方程组的形式。我们可以将任意的 $\mathbf{v} \in \text{span}((6,0,0),(0,12,0),(0,0,1))$ 表示为：

$$\begin{cases} 6x_1 + 0x_2 + 0x_3 &= v_1 \\ 0x_1 + 12x_2 + 0x_3 &= v_2 \\ 0x_1 + 0x_2 + 1x_3 &= v_3 \end{cases}$$

让我们在进一步观察这个线性方程组的形式，我们可以发现如果我们放弃将 vector 写成横向，而是以竖向表示，那么 $\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 6x_1 + 0x_2 + 0x_3 \\ 0x_1 + 12x_2 + 0x_3 \\ 0x_1 + 0x_2 + 1x_3 \end{bmatrix}$。

我们可以发现，x_1, x_2, x_3 的出现是存在规律的，在每一行上第一个出现的都是 x_1 第二个都是 x_2... 以此类推。而且不难验证，这种关系对于任意的 n 都是普遍存在的。作为懒惰且不想多费墨水的我们，自然会想要寻找一种更加简单的方式来反应这样的组合。由于 x_i 的出现位置是一定的，因而只有每个 x_i 前面的系数是重要的。我们需要做的只是将这些系数按照一定的顺序保存下来，而这种保存的方式就是 **matrix**。

Definition 1.8 我们称一个 $m \times n$ 的形如 $\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$ 的二维自然数阵列 $(\forall i \in [1, m] j \in [1, n]\ a_{ij} \in \mathbb{R})$ 称为一个 **matrix**。我们一般使用大写字母的粗体字母 **A,B,C**...来表示矩阵。

我们将矩阵的第 i 行第 j 列的元素称为矩阵的第 i 行第 j 列的 **entry**。我们将矩阵的第 i 行的所有 entry 组成的 list 称为矩阵的第 i 行的 **row**。我们将矩阵的第 j 行的所有 entry 组成的 list 称为矩阵的第 j 列的 **column**。

特别的，我们将这种反应线性方程系数关系的 matrix 称为 **coefficient matrix**。

例如：

对于我们刚刚提到的例子，$\begin{bmatrix} 6 & 0 & 0 \\ 0 & 12 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ 就是这个线性方程组的 coefficient matrix。

那么再进一步，我们有没有可能定义出一种运算，来直接反应 x_1, x_2, x_3 是如何做用到这个 coefficient matrix 上的呢？答案是肯定的。由于我们需要将 x_1, x_2, x_3 的这种有序关系保留下来，我们会自然的想到将 x_1, x_2, x_3 表示为一个 vector，即 $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$。定义运算 $\begin{bmatrix} 6 & 0 & 0 \\ 0 & 12 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 6x_1 + 0x_2 + 0x_3 \\ 0x_1 + 12x_2 + 0x_3 \\ 0x_1 + 0x_2 + 1x_3 \end{bmatrix}$。我们将这种运算称为 **matrix-vector multiplication**。

Definition 1.9 我们定义 $m \times n$ 的 matrix $\mathbf{A}$ 和一个 length 为 n vector $\mathbf{x}$ 的 **matrix-vector multiplication**，$\mathbf{Ax}$，为

$$\mathbf{Ax} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{bmatrix} =: \mathbf{b}$$

我们不难发现，其中 $\mathbf{b}$ 是一个 length 为 m 的 vector，它的第 i 个元素 b_i 为 A 的第 i 行和 x 的第 j 列的“点乘”（对于这一高中知识我们将会在后面进行更加全面和数学化的介绍，其名称为 **inner product**。不同于高中 $\mathbf{v} \cdot \mathbf{u}$ 的记法，我们一般会将这种运算写为 $\langle \mathbf{v}, \mathbf{u} \rangle$。在此后如果我们继续提及点乘，我们可能会将两种写法混用）。

那么进一步的，我们是否可以同时计算多个用同一个 coefficient matrix 表示的线性方程组呢？答案是肯定的。我们只需要将多个 vector 按照一定的顺序排列在一起，形成一个 matrix，然后将这个 matrix 和 coefficient matrix 进行 matrix-vector multiplication 即可。这样，我们就引入了一个新的概念——**matrix multiplication**。

Definition 1.10 我们定义一个 $m \times n$ 的 matrix $\mathbf{A}$ 和一个 $n \times k$ 的 matrix $\mathbf{B}$ 的 **matrix multiplication**，$\mathbf{AB}$，为

$$\mathbf{AB} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1k} \\ b_{21} & b_{22} & \dots & b_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nk} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + \dots + a_{1n}b_{n1} & \dots & a_{11}b_{1k} + a_{12}b_{2k} + \dots + a_{1n}b_{nk} \\ a_{21}b_{11} + a_{22}b_{21} + \dots + a_{2n}b_{n1} & \dots & a_{21}b_{1k} + a_{22}b_{2k} + \dots + a_{2n}b_{nk} \\ \vdots & \ddots & \vdots \\ a_{m1}b_{11} + a_{m2}b_{21} + \dots + a_{mn}b_{n1} & \dots & a_{m1}b_{1k} + a_{m2}b_{2k} + \dots + a_{mn}b_{nk} \end{bmatrix} =: \mathbf{C}$$

我们不难发现，其中 $\mathbf{C}$ 是一个 $m \times k$ 的 matrix，它的第 i 行和第 j 列的 entry 为 A 的第 i 行和 B 的第 j 列的“点乘”。