

线性代数笔记

M.

From Oct. 12 2023
To October 20, 2023

Contents

1	Vector, Vector space, and Linear	2
1.1	Vector and Vector space	2
1.2	Linear, Span and Subspace	4
1.3	Matrix and Matrix computation	6

Chapter 1

Vector, Vector space, and Linear

欢迎来到线性代数课程的第一章, 本章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

在高中的学习过程中, 已经接触到了一种特殊的量——向量 (或矢量)。在过去的学习中认为这种量不只有数值大小, 还有方向。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, 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, 在引入更多坐标轴变量的时候也会更费墨水, 因此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)。

首先严谨的定义一下什么是 list 和 list 的 length。

Definition 1.1. list

定义 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成一个元素。一般将这种有序列写作 $(x_1, x_2, x_3, \dots, x_n)$ 。

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

Definition 1.2. \mathbb{R}^n

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

称其中的 x_i 为 the i -th coordinate。

例如:

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

现在非常神奇地发现, 之前所学过的所谓 vector 都是属于 \mathbb{R}^2 或 \mathbb{R}^3 的 list。回忆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, 就可以定义任意 n 下 \mathbb{R}^n 中元素的加法:

Definition 1.3. 加法

在 \mathbb{R}^n 中符号 “+” 可以由以下公式定义, 将这种运算称为**加法**:

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

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

Definition 1.4. 0

定义 \mathbb{R}^n 中的 **0** 为:

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

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

当然, 数学家不愿意止步于此。希望将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。便有了更加抽象和一般化的对 vector space 和 vector 的定义:

Definition 1.5. vector space

定义一个集合 \mathcal{V} 是一个 **vector space** 当且仅当它满足以下所有性质:

1. Commutativity: $\forall \mathbf{u}, \mathbf{v} \in \mathcal{V} \quad \mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
2. Associativity: $\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{V} \quad \mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{v} + \mathbf{u}) + \mathbf{w}$ and $\forall a, b \in \mathbb{R} \quad a(b\mathbf{v}) = (ab)\mathbf{v}$
3. Additive identity: $\exists \mathbf{u} \in \mathcal{V} \quad \forall \mathbf{v} \in \mathcal{V} \quad \mathbf{v} + \mathbf{u} = \mathbf{v}$ 将 \mathbf{u} 记作0
4. Additive inverse: $\forall \mathbf{v} \in \mathcal{V} \quad \exists \mathbf{u} \in \mathcal{V} \quad \mathbf{v} + \mathbf{u} = 0$
5. Multiplicative identity: $\forall \mathbf{v} \in \mathcal{V} \quad 1\mathbf{v} = \mathbf{v}$
6. Distributive properties: $\forall \mathbf{u}, \mathbf{v} \in \mathcal{V} \quad \forall a, b \in \mathbb{R} \quad a(\mathbf{v} + \mathbf{u}) = a\mathbf{v} + a\mathbf{u} \quad (a + b)\mathbf{v} = a\mathbf{v} + b\mathbf{v}$

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

例如:

\mathbb{R}^n 之中, $\forall a \in \mathbb{R} \quad \forall \mathbf{v} = (v_1, v_2, \dots, v_n) \in \mathbb{R}^n$ 定义 $a\mathbf{v} = (av_1, av_2, \dots, av_n)$ 为 \mathbb{R} 中的数乘。这和在高阶段所学的向量的数乘也是一致的。

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

而也可以举出一些更加奇怪的 vector space 的例子:

如果定义 $P(n) = \{\sum_{i=0}^n p_i t^i : \forall p_i \in \mathbb{R}\}$ 那么可以发现 $P(n)$ 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, \dots, p_n) 唯一确定的表示，将这种关系称为 **isomorphism**。对于这一关系的定义完善先按下不表，继续对于 Vector space 本身的探讨。

1.2 Linear, Span and Subspace

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

当提及 linear 也就线性时候，第一时间想到的一定会是线性函数 (linear function)，也就是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，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。而在上一章节中已经定义了 vector space，也就是说，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. *linear combination*

$\forall \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k \in \mathcal{V} \quad \forall a_1, a_2, \dots, a_k \in \mathbb{R}$ 称形如 $\sum_{i=1}^k a_i \mathbf{v}_i$ 的表达式为 $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ 的一个 **linear combination**。

如果 $\mathbf{u} \in \mathcal{V}$ 是由通过关系式 $\mathbf{u} = a_1 \mathbf{v}_1 + a_2 \mathbf{v}_2 + \dots + a_n \mathbf{v}_n$ 获得的，那么说 \mathbf{u} 可以被 $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ 的 **linear combination** 表示。

例如：

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 $(6, 1, 2) = 1(1, 0, 0) + 1(0, 1, 0) + 2(0, 0, 1)$ 。

这样的，我们对数学中 linearity (线性) 这一概念感知也就不言自明：当某一种元素（可以是 vector, number 也可以是 function 或 map）在其作用方式下对加法具有交换和结合率、并且不在意数乘的先后顺序，我们便可以认为元素是线性的。这种感知在 vector 中就表现为 vector space 中的 Commutativity、Associativity 和 Distributive properties。那么更进一步，对于一个 function 或 map，该如何定义线性呢？在这里，做如下符合这一线性感知的定义：

Definition 1.7. *linear*

定义一个 function 或 map $f : \mathcal{A} \rightarrow \mathcal{B}$ 是 **linear** 的当且仅当它满足以下所有性质：

1. $\forall \mathbf{u}, \mathbf{v} \in \mathcal{A} \quad f(\mathbf{u} + \mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v}) \in \mathcal{B}$
2. $\forall \mathbf{v} \in \mathcal{A} \quad \forall a \in \mathbb{R} \quad f(a\mathbf{v}) = af(\mathbf{v}) \in \mathcal{B}$

例如：

$f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ 定义为 $f((x, y)) = (x + y, x - y)$ 是 linear 的，因为 $f((x_1, y_1) + (x_2, y_2)) = f((x_1 + x_2, y_1 + y_2)) = (x_1 + x_2 + y_1 + y_2, x_1 + x_2 - y_1 - y_2) = (x_1 + y_1, x_1 - y_1) + (x_2 + y_2, x_2 - y_2) = f((x_1, y_1)) + f((x_2, y_2))$ 并且 $f(a(x, y)) = f((ax, ay)) = (ax + ay, ax - ay) = a(x + y, x - y) = af((x, y))$ 。

如果将 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（这里必须指出，虽然在定义什么是 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，是通过对一组感兴趣的元素的线性叠加来构建出所希望研究的 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。但在形式上依旧可以说“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。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，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 1.8. *span*

定义一个集合 \mathcal{U} 是由一组 vector $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \in \mathcal{V}$ “**span**” 的当且仅当集合 \mathcal{U} 中的所有元素恰好是 $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ 的全部 linear combination。将这样的 \mathcal{U} 记做 $\text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$ 。用一种更加数学的表示方法，可以将 $\text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$ 表示为：

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

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。

在这个例子中，不难发现， $\text{span}((6, 0, 0), (0, 12, 0))$ 的结果也是一个 vector space。这个 vector space 是属于 \mathbb{R}^3 的子集。而 $\text{span}(f_1, f_2)$ 的结果也是一个 vector space。这个 vector space 是属于 $P(3)$ 的一个子集。那么更一般化的，我们希望将这样的子空间定义化——称为 **subspace**。

Definition 1.9. *subspace*

定义一个集合 \mathcal{U} 是 \mathcal{V} 的一个 **subspace** 当且仅当它满足以下所有性质：

1. \mathcal{U} 是一个 vector space
2. $\mathcal{U} \subseteq \mathcal{V}$

在获得了 subspace 的定义后，可以发现任意一个 vector space 都拥有大于等于一个的 subspace。基于集合的性质，我们自认会考虑不同 subspace 之间加和。在这里我们引入两种加和方式——**sum** 和 **direct sum**。

Definition 1.10. *sum*

定义两个 subspace \mathcal{U} 和 \mathcal{V} 的 **sum** (记做 $+$) 为：

$$\mathcal{U} + \mathcal{V} := \{ \mathbf{u} + \mathbf{v} : \forall \mathbf{u} \in \mathcal{U} \ \forall \mathbf{v} \in \mathcal{V} \}$$

Definition 1.11. *direct sum*

定义两个 subspace \mathcal{U} 和 \mathcal{V} 的 sum 是 **direct sum** (记做 \oplus) 为:

$$\mathcal{U} \oplus \mathcal{V} := \{\mathbf{u} + \mathbf{v} : \forall \mathbf{u} \in \mathcal{U} \ \forall \mathbf{v} \in \mathcal{V} \ \mathbf{u} + \mathbf{v} \text{ 是唯一的}\}$$

例如:

- $\text{span}(6, 0, 0, (0, 12, 0))$ 和 $\text{span}((12, 0, 0), (0, 19, 0))$ 的 sum, 即 $\text{span}((6, 0, 0), (0, 12, 0)) + \text{span}((12, 0, 0), (0, 19, 0))$, 是在 \mathbb{R}^3 中 $z = 0$ 的平面。由于 $(18, 31, 0)$ 可以表示为 $(18, 31, 0) + 0$ (where $(18, 31, 0) \in \text{span}((6, 0, 0), (0, 12, 0))$), $0 \in \text{span}((12, 0, 0), (0, 19, 0))$ 或 $(6, 12, 0) + (12, 19, 0)$ (where $(6, 12, 0) \in \text{span}((6, 0, 0), (0, 12, 0))$, $(12, 19, 0) \in \text{span}((12, 0, 0), (0, 19, 0))$)。这说明 $\text{span}(6, 0, 0, (0, 12, 0))$ 和 $\text{span}((12, 0, 0), (0, 19, 0))$ 的 sum 并不少 direct sum。
- $\text{span}(6, 0, 0, (0, 12, 0))$ 和 $\text{span}((0, 0, 1))$ 的 sum, 我们可以容易验证这二者的 sum 是一个 direct sum。在此不加赘述, 留待读者自证。

1.3 Matrix and Matrix computation

回顾在讲解 span 的过程中距离使用的的线性组合。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。可以将任意的 $\mathbf{v} \in \text{span}((6, 0, 0), (0, 12, 0), (0, 0, 1))$ 表示为:

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

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

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, 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。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 1.12. matrix

定义一个 $m \times n$ 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**。一般使用大写字母的粗体字母 **A, B, C...** 来表示矩阵。

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

例如:

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

进一步, 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。由于需要将 x_1, x_2, x_3 的这种有序关系保留下来, 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 $\begin{bmatrix} 6 & 0 & 0 \\ 0 & 12 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 6x_1 + 0x_2 + 0x_3 \\ 0x_1 + 12x_2 + 0x_3 \\ 0x_1 + 0x_2 + 1x_3 \end{bmatrix}$ 。将这种运算称为 **matrix-vector multiplication**。

Definition 1.13. matrix-vector multiplication

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

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

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 \mathbf{x} 的第 j 列的“点乘”(对于这一高中知识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, 一般会将这种运算写为 $\langle \mathbf{v}, \mathbf{u} \rangle$ 。在此后如果继续提及点乘, 可能会将两种写法混用)。

更进一步的, 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。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, 就引入了一个新的概念——**matrix multiplication**。

Definition 1.14. *matrix multiplication*

定义一个 $m \times n$ 的 matrix **A** 和一个 $n \times k$ 的 matrix **B** 的 **matrix multiplication**, **AB**, 为

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

不难发现, 其中 **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 inner product。

例如:

$$\begin{aligned} \bullet \quad & \begin{bmatrix} 6 & 0 & 0 \\ 0 & 12 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 12 \\ 1 \\ 9 \end{bmatrix} = \begin{bmatrix} 6 \cdot 12 + 0 \cdot 1 + 0 \cdot 9 \\ 0 \cdot 12 + 12 \cdot 1 + 0 \cdot 9 \\ 0 \cdot 12 + 0 \cdot 1 + 1 \cdot 9 \end{bmatrix} = \begin{bmatrix} 72 \\ 12 \\ 9 \end{bmatrix} \\ \bullet \quad & \begin{bmatrix} 6 & 0 & 0 \\ 0 & 12 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 12 & 1 & 6 \\ 1 & 0 & 2 \\ 9 & 7 & 6 \end{bmatrix} = \begin{bmatrix} 72 & 6 & 36 \\ 12 & 0 & 24 \\ 9 & 7 & 6 \end{bmatrix} \end{aligned}$$