

线性代数第一次习题课

李弢

2019 年 9 月 15 日

目录

1 向量复习与思考	1
2 平面与空间向量的特殊运算	2
3 抽象到一般的向量	3
4 线性方程组	5
5 线性方程组与向量的关系	7

1 向量复习与思考

平面上的向量在中学应该已经学习过：

- 设 \vec{a}, \vec{b} 为两个向量，我们可以定义它们之间的加法 $\vec{a} + \vec{b}$ 。
- 给定一个实数 λ ，我们还可以考虑实数 (标量) 与向量之间的数乘 $\lambda\vec{a}$ 。

在有了坐标系 (不一定是笛卡尔坐标系) 之后，设基底为 $\{\vec{i}, \vec{j}\}$ ，任何一个向量 \vec{a} 都可以用一个二元组来表示： $\vec{a} = x\vec{i} + y\vec{j}$ 。这样，我们可以把向量与二元组等同起来，向量的加法与数乘也就变成了二元组的运算。这两种运算还满足一定的性质，如加法交换律、结合律、数乘与加法的分配律等等。注意，基底的选取不同，同一向量对应的二元组也不同。这里说的“等同”是指给定基底的情况下。

我们刚刚说 $\{\vec{i}, \vec{j}\}$ 是一组基底，基底的意思是说任何其他的向量都可以由它们唯一地 (线性) 表达出来。

- 是不是随便找两个向量都可以构成一个坐标系？显然不是，比如我们考虑 $\{(1, 0), (-1, 0)\}$ ，向量 $(0, 1)$ 不能写成 $x(1, 0) + y(-1, 0)$ 的形式，因此它们没有资格称为一组基底。
- 给定两个向量，我们怎么知道它们有没有资格构成一组基底？判断它们是否共线！
- 为什么我们局限于考虑两个向量，而不是更多的向量呢？比如我们考虑 $\{(1, 0), (0, 1), (1, 1)\}$ ，显然任何向量都可以写成 $x(1, 0) + y(0, 1) + z(1, 1)$ 的形式。但现在的问题在于，这样“自由度”太高了，写法不唯一。
- 我们先验地知道平面是一个“2 维”空间，所以自然而然地认为基底里应该有且仅有两个向量。问题是，“维数”这件事情如何精确地定义出来？

平面向量的定义可以自然地推广到空间向量，此时上面这些问题仍然存在。

- 给定三个向量，我们怎么知道它们是否可以构成一组基底？
- 为什么我们只考虑三个向量构成基底，而不是更多或者更少？

这些问题涉及到线性代数里面的一些核心概念：线性相关、线性无关、基底、维数、坐标等等。

2 平面与空间向量的特殊运算

两个向量可以作内积： $\vec{a} \cdot \vec{b} = |\vec{a}| |\vec{b}| \cos(\theta)$ ，其中 θ 是 \vec{a}, \vec{b} 的夹角。这一定义没有依赖于基底和坐标，只要我们能清晰地定义“模长”和“夹角”的概念就行了，而在平面与空间中这两个概念都是熟知的。反之，如果我们知道了内积，也可以求模长和夹角：

- $|\vec{a}| = \sqrt{\vec{a} \cdot \vec{a}}$
- $\theta = \arccos\left(\frac{\vec{a} \cdot \vec{b}}{|\vec{a}| |\vec{b}|}\right)$

此外，我们还知道“垂直”的两个向量夹角为 $\frac{\pi}{2}$ ，因此内积必然是 0。如果在平面或者空间中我们选取两两内积为 0 的一组向量作为基底，就称为笛卡尔坐标系。在这种坐标系下，内积运算可以方便地由坐标写出：设 $\vec{a} = (x_1, x_2)$ ， $\vec{b} = (y_1, y_2)$ ，那么 $\vec{a} \cdot \vec{b} = x_1 y_1 + x_2 y_2$ ，空间向量也是类似的。注意在这个表达式里面，内积的结果只与坐标有关了。

我们还可以求向量 \vec{a} 在向量 \vec{b} 上的投影，这里以平面向量为例，空间向量的情况是非常类似的。设 $\vec{a} = x\vec{b} + y\vec{c}$ ，其中 \vec{c} 与 \vec{b} 垂直。那么两边同时与 \vec{b} 作内积，可以得到 $\vec{a} \cdot \vec{b} = x|\vec{b}|^2$ ，因而 $x = |\vec{b}|^{-2} \vec{a} \cdot \vec{b}$ 。不过这里有几个小问题： \vec{c} 是否一定存在， $|\vec{b}|$ 会不会为 0？

对于两平面向量，我们还有叉乘运算， $\vec{a} \times \vec{b}$ 是一个向量，它的方向由右手螺旋定则决定（垂直于平面），大小等于以 \vec{a}, \vec{b} 为边的平行四边形的面积。对于三个空间向量，我们有混合积运算， $(\vec{a} \times \vec{b}) \cdot \vec{c}$ ，它等于以 $\vec{a}, \vec{b}, \vec{c}$ 为边的平行六面体的体积。

3 抽象到一般的向量

我们之前只考虑了平面向量与空间向量，并且有一些运算和性质。在前面的过程中，我们对于某一性质或者运算经常只对平面向量进行推导，空间向量的情况也是类似的，就略去不写了。如果二者都验证，则稍显冗余。这里还好，只有平面向量与空间向量两种情况。如果要考虑另一个新品种的向量，是否之前做的那些都得再从头推导一遍？由归纳推理，我们知道结果很可能是成立的，但必须经过演绎推理才能从逻辑上确保正确。像之前那样具体问题具体分析有好处，可以利用一些具体的先验知识，比如“长度”“夹角”“面积”“体积”的定义。但是这种做法不具有“普适性”，想要“一劳永逸”，需要从更抽象的层面出发去考虑问题。我们把平面向量和空间向量的运算和性质中，提取出最本质的几条，作为一般向量的定义。由这个一般向量的定义出发，进行推演，得到的结果就具有普适性了。

- 比如，我们可以考虑 n 维向量。由于平面向量等价于二元组，空间向量等价于三元组，那我们直观地认为， n 维向量等同于 n 元组好了。对于 n 元组，我们类似地去定义加法、数乘，毫无困难。但是在定义内积时就遇到了困难，因为这时我们不再有“长度”“夹角”这种直观的概念了。为了与平面和空间向量的情况相符合，我们就直接定义 n 维向量的内积为 $x_1y_1 + \cdots + x_ny_n$ 。有了内积的定义之后，我们反而可以去考虑“长度”“夹角”这些不太直观的概念了。基于 n 维向量推出的一切东西，对于一切 n 维向量都成立，这就增加了普适性。
- 上面的抽象过程是从平面、立体向量走到了 n 维向量，那么我们能不能更进一步抽象？一个向量一定要是数组的形式吗？我们考虑两个函

数 $f(x), g(x)$, 它们可以作加法 $f(x) + g(x)$, 也可以定义数乘 $\lambda f(x)$, 而且加法和数乘之间具有与平面向量类似的性质。是不是可以把函数也看成是 (无穷维) 向量? 如果我们能合适地定义两个函数的内积, 甚至就可以讨论两个函数是否是“垂直”的了, 这似乎很反直觉。

- 更进一步, 我们可以毫不关心考虑的对象到底是什么、长什么样, 只要它们能定义加法、数乘, 并且加法数乘满足一些性质。由这些性质出发, 我们能推出一些结论, 这样的结论具有某种“一劳永逸”的感觉, 对任何满足这个“模板”的对象都适用。
- 直到刚才, 我们说的数乘一直是实数和向量的运算。我们一定要考虑实数吗? 能不能考虑有理数、整数、自然数? 能不能“不是数”?

我们从具体到抽象, 引发了许多的问题。从具体走向抽象, 得出的结论越来越普适, 但能用的条件与直观也越来越少。该抽象到什么层次为止? 这是个很难回答的问题。不过, 在线性代数这门课里, 至少上面说的这么几层抽象我们都会有所涉及。在这里我们引用李尚志《线性代数》课本前言中的一段话:

抽象确实是学习线性代数的一个拦路虎。一提起抽象, 给人的印象就是莫名其妙、晦涩难懂、脱离实际、没有用处, 总之是一个令人害怕的贬义词。然而, 抽象并不是线性代数特有的, 也不是从大学开始的。比如, 幼儿园的小孩就要学 $3 + 2 = 5$, 这是抽象还是具体? 怎样教小孩 $3 + 2 = 5$? 是先教加法的定义, 然后再按照定义来做 $3 + 2 = 5$ 吗? 加法的定义, 幼儿园没教过, 小学和中学也没教过。然而小孩们却学会了加法, 不是靠定义学会加法, 而是通过例子学会了加法。比如, 可以教小孩数自己的手指来学 $3 + 2 = 5$, 3 根手指加 2 根手指就是 5 根手指。也可以数铅笔, 3 枝铅笔加 2 枝铅笔就是 5 枝铅笔。假如已经数过了手指, 又数过了铅笔, 一个细心而胆大的小孩发现手指是肉做的, 铅笔是木头做的, 举手问老师: “5 是肉做的还是木头做的?” 老师怎样回答? 假如又数了 5 个乒乓球, 发现手指和铅笔是长的, 乒乓球是圆的, 再问: “5 是长的还是圆的?” 老师又怎样回答? 也许老师会斥责这个不听话的调皮小孩: “好好听课, 不要胡说八道!” 然而, 这样的小孩才是聪明的小孩, 会思考的小孩。他注意到了 5 根手指、5 枝铅笔、5 个乒乓球的差别, 这确实是聪明的表现。但只是注意到差别还不够, 还要让他学会忽略这种差别, 将肉做的 5 根手指、木头做的 5 根铅笔“混为一谈”, 将 5 个长的物体 (手指和铅笔) 和 5 个圆的物体 (乒乓球) “混为一谈”, 忽略它们的差别而只关心它们的共同点: 数量的多少, 这才学会了 $3 + 2 = 5$, 这才能够将 $3 + 2 = 5$ 用到千千万万其他例子, 如 3 本书加 2 本书, 3 张桌

子加 2 张桌子等。要让小孩学会忽略这些差别，不是一件容易的事情。这正如郑板桥说的，聪明难，糊涂亦难，由聪明而糊涂尤其难。这种忽略差别的过程，就是“由聪明而糊涂”的过程，也就是数学的抽象的过程。抽象不是从天而降，而是来自于实际，来自于具体的例子。然而，抽象又没有停留于实际，而是“脱离”了实际：它脱离了具体的例子，舍弃了不同例子的不同点而提取了它们的共同点，这样才能应用到更多更广泛的实际例子中。有一个电视节目的时事评论员常说：“许多看似不相干的事情，其实都是相互关联的。”我们可以说：“许多看似不相同的事情，其实都有共同点。”从不同的事情中发现共同点，研究共同点，得到放之四海而皆准的真理，用到更多的不同事物中去，这就是抽象。这样的抽象不是没有用处，反而是神通更广大。数学由低级到高级的过程，就是抽象的程度由低到高的过程，也是应用的范围由狭窄到广泛的过程。幼儿园的 $3 + 2 = 5$ 忽略掉了大小、长短、原料的差别，只关心数量的多少。初中的 $(a - b)^2 = a^2 - 2ab + b^2$ 将字母 a, b 所代表的数是多少也忽略掉了，只关心它们的共同的运算规律。更进一步的“糊涂”是：公式 $(a - b)^2 = a^2 - 2ab + b^2$ 中的字母 a, b 可以不代表数而代表几何向量，将其中的乘法理解为向量的内积，公式照样成立。这就是余弦定理！当夹角为直角时就是勾股定理！只不过一念之差，在乘法公式 $(a - b)^2 = a^2 - 2ab + b^2$ 中“难得糊涂”，将数与向量“混为一谈”，就立即得到了余弦定理和勾股定理，数学的抽象的威力由此可见一斑

4 线性方程组

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

这一方程组里面有 m 个方程， n 个未知数。我们关心两个问题：

- 方程组是否有解？
- 方程组如果有解，解是否唯一？

一个直观的感觉是，每个未知数代表一个“自由度”，每个方程代表一个“约束”，若约束多于自由度，则不存在解；若自由度多于约束，则解不唯一。这个感觉不对，或者说，它是“naive”的，它是一个基本的出发点，但需要再加很多限定条件才能对。我们下面通过一个具体例子来考察线性方程组相关的一些问题。

假设这门课程的给分规则是 平时分 $\times x +$ 期中 $\times y +$ 期末 $\times z =$ 总评。假设你现在不知道给分规则，但你知道自己的平时分、期中成绩、期末成绩、总评，你想推算老师的给分比例。

- 仅由你自己的成绩，原则上是无法算出给分规则的。你至少需要找两个同学，来建立一个 3 个未知数，3 个方程的方程组：

$$\begin{cases} 90x + 90y + 90z = 90 \\ 95x + 98y + 96z = 96.4 \\ 80x + 85y + 79z = 81 \end{cases} \quad (2)$$

由此，你可以求解出 $x = 0.2, y = 0.3, z = 0.5$ 。

```
A = rbind(c(90,90,90),c(95,98,96),c(80,85,79))
b = c(90,96.4,81)
solve(A,b)
```

```
## [1] 0.2 0.3 0.5
```

- 但是也有可能出现以下几种情况：

$$\begin{cases} 90x + 90y + 90z = 90 \\ 90x + 90y + 90z = 90 \\ 80x + 85y + 79z = 81 \end{cases} \quad \begin{cases} 90x + 90y + 90z = 90 \\ 100x + 100y + 100z = 100 \\ 80x + 85y + 79z = 81 \end{cases} \quad \begin{cases} 90x + 90y + 90z = 90 \\ 95x + 98y + 96z = 96.4 \\ 92.5x + 94y + 93z = 93.2 \end{cases} \quad (3)$$

这时，虽然有 3 个方程组，但你没法得出唯一解，这是因为这里面有些方程是“假”的，看似存在但其实没有提供额外的信息。

- 也有可能出现这种情况：

$$\begin{cases} 90x + 90y + 90z = 90 \\ 90x + 90y + 90z = 91 \\ 80x + 85y + 79z = 81 \end{cases} \quad (4)$$

这个方程组不存在解，因为前两个方程矛盾了。第二位同学与第一位同学各项成绩都一样，总评却不一样，这可能是有一些额外因素（如小论文加分）没有考虑进来。

- 由此可见，并不是说方程组个数和未知数个数一样，就一定存在唯一解。我们先不考虑刚刚说的矛盾的情况，考虑有“假”方程组的情况。为了解决这个问题，假设我们全班同学都坦诚相待，制作了一张包含全班同学各项成绩的表格，那么可以得到一个具有 60 个方程，3 个未知数的方程组。如果不是特别凑巧，这个方程组肯定可以求解。但是，实际中给分不太可能出现小数点后面的数字，这就意味着你们的分数可能是经过四舍五入的，也就是说方程未必严格成立。

$$\begin{cases} 47 \times 0.2 + 48 \times 0.3 + 49 \times 0.5 = 48.3 \rightarrow 48 \\ 94 \times 0.2 + 96 \times 0.3 + 98 \times 0.5 = 96.6 \rightarrow 97 \end{cases} \quad (5)$$

这样，这 60 个方程可能就会导出一些矛盾来。在这种不可能存在解的情况下，如何寻找近似解？我们知道肯定存在这样的近似解，因为老师确实按照这一方式来计算分数，只是数据中存在误差而已。一种解决方式就是最小二乘法，这构成了“计量经济学”这门课一半以上的内容。

- 也许有些同学会有疑问，为什么有些时候仅靠自己的成绩也能大概估算出给分规则？比如平时成绩 100，期中 95，期末 85，总评 91，则大概能估算出 0.2, 0.3, 0.5。事实上，0.1, 0.1, 0.841176 也是解，但是在这个语境下我们不考虑这种解，因为我们知道成绩的权重不太可能出现这种奇怪的形式，相当于我们对与解有“先验”。如果将具体问题的先验知识引入模型中，则不在这门课的讨论范围之内，但也许是经济学的重要课题之一。

5 线性方程组与向量的关系

我个人习惯用两种观点来看线性方程组，一种基于行向量的观点，一种基于列向量的观点。对于一个一般的线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \quad (6)$$

如果右端项全为 0，我们称之为齐次方程组，否则称为非齐次方程组。

- 对于非齐次方程组,我们采用列向量观点,记 $\vec{c}_i = (a_{11}; a_{21}; \cdots; a_{m1}), i = 1, 2, \cdots, n$ 以及 $\vec{b} = (b_1; b_2; \cdots; b_m)$, 那么原方程组即等价于寻找 $\sum_{i=1}^n x_i \vec{c}_i = \vec{b}$ 。如果 $\{\vec{c}_i\}$ 能构成 m 元数组向量空间的一组基底,那么就等价于求 \vec{b} 在这组基底下的坐标了。当然一般情况下这不太可能成立。我们自然要问什么情况下能有解、且解唯一? 考虑 $m = 3, n = 2$ 的情况, 容易发现方程组有解实际上就等价于 \vec{b} 落在 \vec{c}_1, \vec{c}_2 所决定的平面(共线时可能是一条直线)内。若 \vec{c}_1, \vec{c}_2 共线, 且 \vec{b} 也在这条线上, 则有无穷多组解。对于 m, n 更大的一般情况, 如何类似地去考虑? 这需要线性代数中“秩”的概念。
- 对于齐次方程组,我们采用行向量观点,记 $\vec{r}_i = (a_{11}, a_{12}, \cdots, a_{1n}), \vec{x} = (x_1, x_2, \cdots, x_n)$, 则原方程组就等于寻找 \vec{x} 使得 $\vec{r}_i \cdot \vec{x} = 0, \forall i = 1, 2, \cdots, m$ 。这就是说, 求解原齐次方程组等价于寻找一个与系数矩阵行向量都“垂直”的向量。这一观点很重要, 特别是在考虑正交投影与最小二乘法正则方程组时可以很大程度上帮助理解。显然, 如果这样的向量存在则不唯一, 因为乘以常数倍后仍然是解。我们至少可以断言, 对于非齐次线性方程组而言, 要么无(非平凡)解, 要么有无穷多解, 不可能存在唯一非零解。此外, 在这种情况下, 方程组个数小于未知数个数时存在非平凡解。