线性代数习题课 1

王睿

2023 年 3 月 26 日

目录

一点说明

1. 作业讲义部分题过程可能有省略。如对作业仍有疑问可以在群里或答疑课上讨论。
2. 习题课不会讲所有作业的证明。若对没讲的部分有疑问，可以在群里或答疑课上讨论。习题课讲义也不能替代作业讲义。
3. 请及时核对自己在 BB 系统里的分数，如有问题请向对应的助教反馈。
4. 附录里的内容仅供有兴趣的同学参考，有可能涉及之后才会学习或课外的知识，不要求在现阶段掌握。
5. 讲义最好用电脑打开，文档内置了链接功能，复习或查看指定的作业很方便。

1 第一周作业

本周主要讲了将线性方程组转化为增广矩阵，并使用初等行变换求解线性方程组。需要特别注意计算的准确性。

1.1 Cauchy 不等式的证明

证明如下形式的 Cauchy 不等式：若 $a_1,\cdots,a_n,b_1,\cdots,b_n\in\mathbb{R}$，则必有

$$(a_1b_1+a_2b_2+\cdots+a_nb_n)^2\le\left(a_1^2+a_2^2+\cdots+a_n^2\right)\left(b_1^2+b_2^2+\cdots+b_n^2\right)$$

解答：Cauchy 不等式有许多证明方法，这里只给出一种。

$$\begin{aligned}
&\left(\sum_{i=1}^n a_i^2\right)\left(\sum_{i=1}^n b_i^2\right)-\left(\sum_{i=1}^n a_ib_i\right)^2\\
=&\sum_{i=1}^n a_i^2\sum_{j=1}^n b_j^2-\sum_{i=1}^n a_ib_i\sum_{j=1}^n a_jb_j\\
=&\sum_{i=1}^n\sum_{j=1}^n a_i^2b_j^2-\sum_{i=1}^n\sum_{j=1}^n a_ib_ia_jb_j\\
=&\sum_{i=1}^n\sum_{j=1}^n\left(a_i^2b_j^2-a_ib_ia_jb_j\right)\\
=&\frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n\left(a_i^2b_j^2-2a_ib_ia_jb_j+a_j^2b_i^2\right)\\
=&\frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n(a_ib_j-a_jb_i)^2\ge 0
\end{aligned}$$

其中 $\sum$ 符号为求和符号，$\sum\limits_{i=1}^n a_i^2$ 代表从 i 从 1 取到 n，对 a_i^2 的求和，即

$$\sum_{i=1}^n a_i^2=a_1^2+a_2^2+\cdots+a_n^2$$
$$\sum_{i=1}^n a_ib_i=a_1b_1+a_2b_2+\cdots+a_nb_n$$

求和符号参考教材 1.9 节第 27 页。

有一些同学使用的是构造函数的方法：构造

$$\left(\sum_{i=1}^n a_i^2\right)x^2+2\sum_{i=1}^n a_ib_ix+\sum_{i=1}^n b_i^2=\sum_{i=1}^n(a_ix+b_i)^2\ge 0$$

对任意 $x\in\mathbb{R}$ 有

$$f(x)=\sum_{i=1}^n(a_ix+b_i)^2\ge 0$$

因此有关 x 的二次方程判别式小于或等于 0，可得

$$4\left(\sum_{i=1}^n a_ib_i\right)^2-4\sum_{i=1}^n a_i^2\sum_{i=1}^n b_i^2\le 0$$

总体来讲，上述证明没有什么大问题；但应注意的是要分 $\sum_{i=1}^{n} a_i^2 = 0 \text{ or } \neq 0$ 的两种情况。不等于 0 的情况上面已经讨论过；$\sum_{i=1}^{n} a_i^2 = 0$ 时，有

$$\sum_{i=1}^{n} a_i^2 = 0 \Rightarrow a_i = 0 \Rightarrow \left(\sum_{i=1}^{n} a_i b_i\right)^2 = 0 \leq \left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{i=1}^{n} b_i^2\right) = 0$$

可知 Cauchy 不等式成立。这样证明才是比较严谨的。

Cauchy 不等式的其他证明方法可以参考https://www.zhihu.com/question/449533105。

个人而言，不是很推荐使用向量的夹角作为证明方法，主要基于以下原因：**在一般的线性空间理论中，内积不是必要的定义**。事实上，在教材中直到第七章欧几里得空间才定义了线性空间中的内积和“夹角”。夹角的性质

$$-1 \leq \frac{\boldsymbol{\alpha} \cdot \boldsymbol{\beta}}{|\boldsymbol{\alpha}||\boldsymbol{\beta}|} \leq 1$$

需要通过 Cauchy-Schwarz 不等式来证明，而 Cauchy-Schwarz 不等式的证明依赖线性空间上内积的定义。而迄今为止，还不能说严格地定义了线性空间中的内积。没有严格内积的定义、却使用依赖内积定义的“夹角”性质来证明 Cauchy 不等式，有点本末倒置的意味。当然，对几何空间的直观理解——两个向量之间有夹角，余弦值在 $[-1, 1]$ 之间也不能说错；但应注意，在很多线性空间中“夹角”是相当难理解（且没啥用）的概念，比如傅里叶级数也是一种定义内积的线性空间（例 7.1.3）。有关 Cauchy-Schwarz 不等式的证明详见附录，傅里叶级数的线性空间理解会在讲到线性空间时再介绍。

习题 2 设三维空间中有向量 $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$, 它们的长度相同, 两两之间的夹角相等。若 $\boldsymbol{a} = (1, 1, 0), \boldsymbol{b} = (0, 1, 1)$, 求出 $\boldsymbol{c}$.

注意多解，请认真计算。

1.2 增广矩阵计算

所有人的方法都是对的，但不是所有人都能算对，一定要**注意计算**，并且多做检查。

对实际解方程而言，一般化为标准形式就可以了，并不要求化为**约化标准形**，但大家还是有必要了解约化标准形的定义。约化标准形要求所有主元为 1，且在主元所在列，在主元上面的元素为 0。约化标准形的存在唯一性参考老师讲义中给的链接，或见附录助教的翻译版。

2 第二周作业

本周主要讲了从增广矩阵的标准型形式判断解的存在性、如何写出解；并引入矩阵运算规则，以及矩阵计算的基本练习，要特别注意运算的准确性。

2.1 含未知量的增广矩阵

2. 当 a 为何值时，下列线性方程组有解？有解时求出它的通解：

$$(1)\begin{cases} 3x_1 & +2x_2 & +x_3 & =2 \\ x_1 & -x_2 & -2x_3 & =-3 \\ ax_1 & -2x_2 & +2x_3 & =6 \end{cases} \qquad (2)\begin{cases} x_1 & -4x_2 & +2x_3 & =-1 \\ -x_1 & +11x_2 & -x_3 & =3 \\ 3x_1 & -5x_2 & +7x_3 & =a \end{cases}$$

$$(1)\begin{cases} 3x_1 & +2x_2 & +x_3 & =2 \\ x_1 & -x_2 & -2x_3 & =-3 \\ ax_1 & -2x_2 & +2x_3 & =6 \end{cases} \Leftrightarrow \begin{pmatrix} 3 & 2 & 1 & 2 \\ 1 & -1 & -2 & -3 \\ a & -2 & 2 & 6 \end{pmatrix} \xrightarrow{r_1\leftrightarrow r_2} \begin{pmatrix} 1 & -1 & -2 & -3 \\ 3 & 2 & 1 & 2 \\ a & -2 & 2 & 6 \end{pmatrix}$$

$$\xrightarrow[-ar_1\to r_3]{-3r_1\to r_2} \begin{pmatrix} 1 & -1 & -2 & -3 \\ 0 & 5 & 7 & 11 \\ 0 & a-2 & 2+2a & 3a+6 \end{pmatrix} \xrightarrow{-(a-2)/5\ r_2\to r_3} \begin{pmatrix} 1 & -1 & -2 & -3 \\ 0 & 5 & 7 & 11 \\ 0 & 0 & \frac{3}{5}a+\frac{24}{5} & \frac{4}{5}a+\frac{52}{5} \end{pmatrix}$$

$$\xrightarrow{5r_3} \begin{pmatrix} 1 & -1 & -2 & -3 \\ 0 & 5 & 7 & 11 \\ 0 & 0 & 3a+24 & 4a+52 \end{pmatrix}$$

故方程组无解的充要条件是 $3a+24=0, 4a+22\neq 0 \Rightarrow a=-8$。$a\neq -8$ 时，方程组的有效方程个数 $=$ 未知元个数, 从而方程组有唯一解。利用回代可以解出

$$x_1=\frac{4}{a+8}, x_2=\frac{a-20}{3(a+8)}, x_3=\frac{4(a+13)}{3(a+8)}$$

有一点要特别注意，在初等行变换中有未知量被涉及，尽量不要出现形如 $1/a$ 形式的变换，是不够严谨的行为。

补充习题 1 假定 3×4 的方程组的增广矩阵可以化简为如下的（约化）标准形

$$\left(\begin{array}{cccc:c} 1 & 2 & 0 & 3 & 4 \\ 0 & 0 & 1 & 5 & 6 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

在求解该方程组时, 试问可以选取哪些未知元来组成自由元? 找出所有的可能性.

可以这样考虑：有 4 个未知量和 2 个线性无关的方程，需要引入 2 个自由元才能给出通解的一般形式，同时引入的自由元不能彼此制约（或线性无关。意味着不存在某个方程在引入自由元后成为一个仅包含被引入自由元的方程，此时其中一个自由元之间可以用其他自由元表示）。若增广矩阵的前四列对应 x_1,x_2,x_3,x_4，则从中选取 2 个的组合为 $C_4^2=6$，同时根据第二行的方程，应排除选择 x_3,x_4 的情况。即任选 x_1,x_2,x_3,x_4 中的 2 个作为自由元，且不是 x_3,x_4，共 5 种。

补充习题 2 (1) 给出一个由 3 个方程构成的关于未知元 x_1, x_2, x_3 的齐次线性方程组, 使得 $x_1 = 3, x_2 = -2, x_3 = 1$ 是它的一个解, 并且方程组的增广矩阵的约化标准形是以前两列为主元所在的列。

约化标准形的定义：所有主元为 1，且在主元所在列，在主元上面的元素为 0，对给定的矩阵而言，其通过一系列初等行变换所能得到的符合条件的标准形式存在且唯一，这样的标准形式称为原矩阵的约化标准形。

齐次线性方程组一定有零解 $x_1 = 0, \cdots x_n = 0$，而 $x_1 = 3, x_2 = -2, x_3 = 1$ 是它的一个解，说明这个方程组存在多解，说明增广矩阵的标准型有不止一行是全部为 0 的形式。同时由于方程组的增广矩阵的约化标准形以前两列为主元，若前两列是 x_1, x_2，说明只有一个自由元 x_3。同时对于齐次线性方程组，增广矩阵的最后一列一定全为 0。因此可知增广矩阵的约化标准形一定为

$$\begin{pmatrix} 1 & 0 & c_{13} & 0 \\ 0 & 1 & c_{23} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

此时还有 $x_1 = 3, x_2 = -2, x_3 = 1$ 是它的一个解的条件没有用到，我们可以直接构造

$$\begin{pmatrix} 1 & 0 & -3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

即可使特解满足条件。由于题目要求 3 个方程构成，最后一行用前两个方程的线性组合就可以生成了，最简单的：

$$\begin{cases} x_1 & & -3x_3 & = 0 \\ & x_2 & +2x_3 & = 0 \\ & x_2 & +2x_3 & = 0 \end{cases}$$

(2) 是否存在一个由 3 个方程构成的关于未知元 x_1, x_2, x_3 的齐次线性方程组, 使得 $x_1 = 3, x_2 = -2, x_3 = 1$ 仍然是它的一个解, 但是方程组的增广矩阵的约化标准形是以第一列和第三列为主元所在的列?

与上一问类似的分析，我们知道增广矩阵的标准型有不止一行是全部为 0 的形式。若方程组的增广矩阵的约化标准形是以第一列和第三列为主元所在的列，则增广矩阵的标准型形式为

$$\begin{pmatrix} 1 & c_{12} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

在这种条件下，一定有 $x_3 = 0$，与 $x_1 = 3, x_2 = -2, x_3 = 1$ 是它的一个解矛盾。因此不存在这样的齐次线性方程组满足上述条件。

2.2 3 月 16 日布置的作业

补充习题 4 验证: 矩阵乘法 $\boldsymbol{AB}$ 中的每个行向量都是 $\boldsymbol{B}$ 的行向量组的一个线性组合。

记 $\boldsymbol{B}$ 矩阵的按行可以写成

$$\boldsymbol{B} = \begin{pmatrix} \boldsymbol{\xi}_1 \\ \boldsymbol{\xi}_2 \\ \vdots \\ \boldsymbol{\xi}_n \end{pmatrix}_{n\times p} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{pmatrix}_{n\times p}$$

设 $\boldsymbol{C}_{m\times p}=\boldsymbol{A}_{m\times n}\boldsymbol{B}_{n\times p}$ 的第 i 行行向量为 $\boldsymbol{\eta}_i$，即

$$\boldsymbol{C}=\begin{pmatrix}\boldsymbol{\eta}_1\\ \boldsymbol{\eta}_2\\ \vdots\\ \boldsymbol{\eta}_m\end{pmatrix}_{m\times p}=\begin{pmatrix}c_{11}&c_{12}&\cdots&c_{1p}\\ c_{21}&c_{22}&\cdots&c_{2p}\\ \vdots&\vdots&&\vdots\\ c_{n1}&c_{n2}&\cdots&c_{np}\end{pmatrix}_{n\times p}$$

现在我们只考察 $\boldsymbol{C}$ 矩阵的一行，将运算显式地写出

$$\begin{cases}c_{i1}=a_{i1}b_{11}+a_{i2}b_{21}+\cdots a_{in}b_{n1}\\ c_{i2}=a_{i1}b_{12}+a_{i2}b_{22}+\cdots a_{in}b_{n2}\\ \qquad\qquad\qquad\qquad\qquad\vdots\\ c_{ip}=a_{i1}b_{1p}+a_{i2}b_{2p}+\cdots a_{in}b_{np}\end{cases}$$

根据我们之前的考虑，$\boldsymbol{\xi}_j=\{b_{j1},b_{j2},\cdots,b_{jp}\},\boldsymbol{\eta}_i=\{c_{i1},c_{i2},\cdots,c_{ip}\}$，因此上式看成是向量的 p 个分量，可重写成

$$\boldsymbol{\eta}_i=a_{i1}\boldsymbol{\xi}_1+a_{i2}\boldsymbol{\xi}_2+\cdots a_{in}\boldsymbol{\xi}_n$$

即 $\boldsymbol{C}=\boldsymbol{AB}$ 中的每个行向量都是 $\boldsymbol{B}$ 的行向量组的一个线性组合。

有一些同学用到了矩阵分块的性质，数学上来说很多人用的没问题，但还是希望你们确实是理解了矩阵分块而不是随意使用。矩阵分块的规则参考老师讲义。

3 第三周作业

本周主要讲了矩阵多项式、矩阵的逆、矩阵的转置、迹等运算性质，以及矩阵分块运算。

这里只提醒一下有关矩阵多项式的运算。矩阵与多项式运算规则的类比有一个很重要的区别，数域中的运算可交换，但矩阵乘法一般不可交换。要想把矩阵类比到多项式的乘除需要进行运算的单元之间满足可交换的性质（$\boldsymbol{A}$ 与 $\boldsymbol{B}$ 可交换意思是 $\boldsymbol{AB}=\boldsymbol{BA}$）。一般而言，对一个未知的方阵 $\boldsymbol{A}$，与之可交换的方阵至少有 $\boldsymbol{I},\boldsymbol{O},\boldsymbol{A}^k,k\in\mathbb{Z}^+$，若 A 可逆，则还有 $\boldsymbol{A}^k,k\in\mathbb{Z}^-,f(\boldsymbol{A})$，$f(\boldsymbol{A})$ 是 $\boldsymbol{A}$ 的函数。

4 线性空间中的内积：Cauchy-Schwarz 不等式的证明

第一次作业第一题比较合适的方法（第七章内容，自行学习不要求掌握）是：先证明

$$(\boldsymbol{\alpha},\boldsymbol{\beta})=\boldsymbol{\alpha}\cdot\boldsymbol{\beta}=\sum_{i=1}^{n}a_ib_i=a_1b_1+a_2b_2+\cdots+a_nb_n \tag{4.1}$$

是实线性空间 V 中的内积，为此需要满足

1. 对称性：$\forall\boldsymbol{\alpha},\boldsymbol{\beta}\in V,(\boldsymbol{\alpha},\boldsymbol{\beta})=\sum_{i=1}^{n}a_ib_i=(\boldsymbol{\beta},\boldsymbol{\alpha})$
2. 恒正性：$\forall\boldsymbol{\alpha}\in V,\boldsymbol{\alpha}\neq\mathbf{0},(\boldsymbol{\alpha},\boldsymbol{\alpha})=\sum_{i=1}^{n}a_i^2>0$
3. 双线性：$\forall\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha},\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}\in V,\lambda,\mu\in\mathbb{R}$，均有

$$\begin{aligned}(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\boldsymbol{\beta})&=(\boldsymbol{\alpha}_1,\boldsymbol{\beta})+(\boldsymbol{\alpha}_2,\boldsymbol{\beta})\\(\lambda\boldsymbol{\alpha},\boldsymbol{\beta})&=\lambda(\boldsymbol{\alpha},\boldsymbol{\beta})\\(\boldsymbol{\alpha},\boldsymbol{\beta}_1+\boldsymbol{\beta}_2)&=(\boldsymbol{\alpha},\boldsymbol{\beta}_1)+(\boldsymbol{\alpha},\boldsymbol{\beta}_2)\\(\boldsymbol{\alpha},\mu\boldsymbol{\beta})&=\mu(\boldsymbol{\alpha},\boldsymbol{\beta})\end{aligned} \tag{4.2}$$

将方程 (4.1) 代入方程 (4.2) 验证即知满足。

因此方程 (4.1) 定义的是内积。从而有

$$(\boldsymbol{\alpha},\boldsymbol{\beta})^2\leq(\boldsymbol{\alpha},\boldsymbol{\alpha})(\boldsymbol{\beta},\boldsymbol{\beta}) \tag{4.3}$$

这里展示一种与课本略微不同的证明，仅作为课本的补充。当 $\boldsymbol{\alpha}=\mathbf{0}$ 时，$(\mathbf{0},\boldsymbol{\beta})=(0\boldsymbol{\alpha},\boldsymbol{\beta})=0(\boldsymbol{\alpha},\boldsymbol{\beta})=0,(\mathbf{0},\mathbf{0})=0$，因此式 (4.3) 成等式，即当 $\boldsymbol{\alpha}=\mathbf{0}$ 时，式 (4.3) 成立。

设 $\boldsymbol{\alpha}\neq\mathbf{0}$，记 $\boldsymbol{\beta}_1=\boldsymbol{\beta}-\frac{(\boldsymbol{\alpha},\boldsymbol{\beta})}{(\boldsymbol{\alpha},\boldsymbol{\alpha})}\boldsymbol{\alpha}$，由内积的恒正性 $0\leq(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)=(\boldsymbol{\beta},\boldsymbol{\beta})-\frac{(\boldsymbol{\alpha},\boldsymbol{\beta})^2}{(\boldsymbol{\alpha},\boldsymbol{\alpha})}$ 得 $(\boldsymbol{\alpha},\boldsymbol{\beta})^2\leq(\boldsymbol{\alpha},\boldsymbol{\alpha})(\boldsymbol{\beta},\boldsymbol{\beta})$，等号成立当且仅当 $\boldsymbol{\beta}_1=\mathbf{0}$，即 $\boldsymbol{\beta}=\frac{(\boldsymbol{\alpha},\boldsymbol{\beta})}{(\boldsymbol{\alpha},\boldsymbol{\alpha})}\boldsymbol{\alpha}$ 时成立，此时 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 线性相关。

将方程 (4.1) 代入式 (4.3)，有

$$(a_1b_1+a_2b_2+\cdots+a_nb_n)^2\leq\left(a_1^2+a_2^2+\cdots+a_n^2\right)\left(b_1^2+b_2^2+\cdots+b_n^2\right). \tag{4.4}$$

一个例子： 设 $\mathbb{R}^2$ 是所有 2 维实向量 (x_1,x_2) 构成的 2 维实向量空间，$\boldsymbol{\alpha}=(x_1,x_2),\boldsymbol{\beta}=(y_1,y_2)\in\mathbb{R}^2$. 定义 2 维实向量空间 $\mathbb{R}^2$ 上二元实函数 $(\boldsymbol{\alpha},\boldsymbol{\beta})$ 为

$$(\boldsymbol{\alpha},\boldsymbol{\beta})=x_1y_1-x_2y_1-x_1y_2+4x_2y_2 \tag{4.5}$$

容易验证，二元实函数 $(\boldsymbol{\alpha},\boldsymbol{\beta})$ 满足：对称性、恒正性和双线性。所以二元实函数 $(\boldsymbol{\alpha},\boldsymbol{\beta})$ 是 2 维实向量空间 $\mathbb{R}^2$ 上的一个内积。2 维实向量空间 $\mathbb{R}^2$ 连同内积 $(\boldsymbol{\alpha},\boldsymbol{\beta})$ 便构成一个 Euclid 空间。代入式 (4.3) 一样可以得到一个不等式：

$$(x_1y_1-x_2y_1-x_1y_2+4x_2y_2)^2\leq(x_1^2-2x_1x_2+4x_2^2)(y_1^2-2y_1y_2+4y_2^2) \tag{4.6}$$

我这里使用 mathematica 验证上式的正确性,如图4.1,并且我们还可以发现等号成立当且仅当 $x_2y_1 = x_1y_2$，即 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 线性相关。

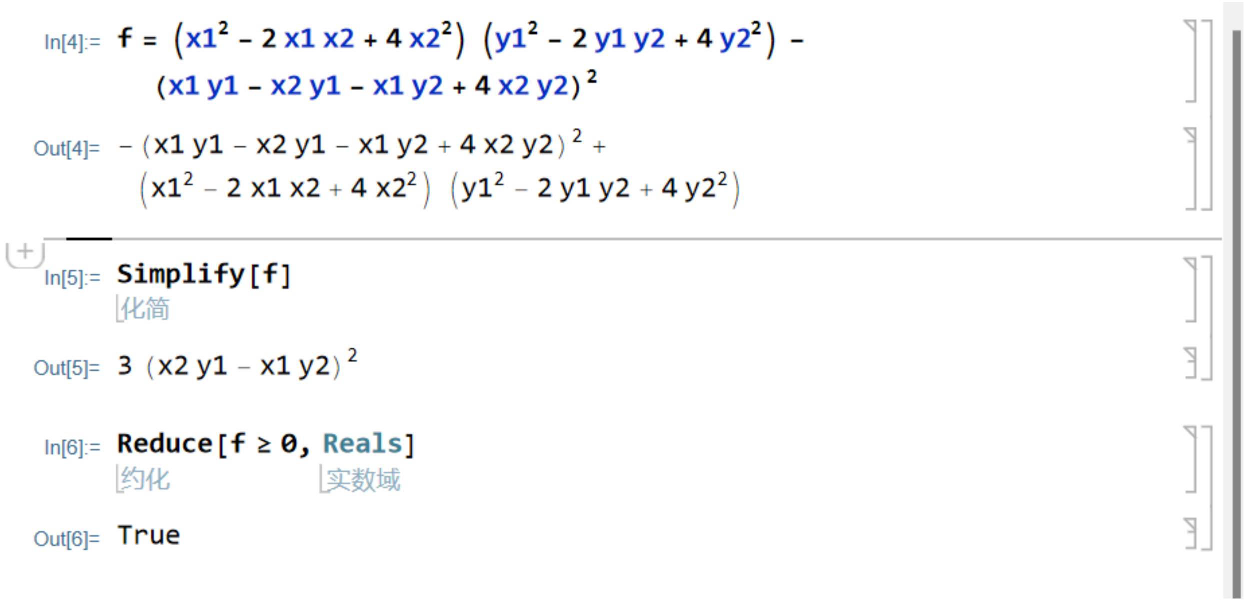

图 4.1: 使用软件可以辅助自己的计算。

当然我们也可以形式地定义“夹角”为

$$\begin{aligned}\theta &= \arccos \frac{(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\sqrt{(\boldsymbol{\alpha}, \boldsymbol{\alpha})(\boldsymbol{\beta}, \boldsymbol{\beta})}} \\ &= \arccos \frac{(x_1y_1 - x_2y_1 - x_1y_2 + 4x_2y_2)}{\sqrt{(x_1^2 - 2x_1x_2 + 4x_2^2)(y_1^2 - 2y_1y_2 + 4y_2^2)}}\end{aligned} \tag{4.7}$$

上式定义的夹角可以直接用来证明 Cauchy-Schwarz 不等式吗？显然是不行的，所以用夹角的性质直接推到 Cauchy 不等式也是有问题的。

5 求和符号的理解

求和符号在线性代数的应用相当重要，有利于大家抽象地理解线性代数。矩阵运算、行列式的计算都需要用到求和符号。下面简单介绍一下我对求和符号的理解。

对求和符号 $\sum_{i=1}^{n} f(i,j)$，其中 $f(i,j)$ 是任意关于 i,j 的函数，这个式子是对被求和部分 i 的求和，求和后 i 作为求和指标应消失，而 j 应保留，即最终结果一定有

$$\sum_{i=1}^{n} f(i,j) = g(j) \tag{5.1}$$

同样地，也可以通过观察等式中被求和的部分，将多项相加用求和符号简单地表达，如

$$\left(\sum_{i=1}^{m} a_{i1}x_i\right) y_1 + \left(\sum_{i=1}^{m} a_{i2}x_i\right) y_2 + \cdots \left(\sum_{i=1}^{m} a_{in}x_i\right) y_n = \sum_{j=1}^{n}\left(\sum_{i=1}^{m} a_{ij}x_i\right) y_j = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}x_iy_j \tag{5.2}$$

对于多个求和符号，如

$$\sum_{i=1}^{m}\sum_{j=1}^{n} f(i,j) \tag{5.3}$$

可以用数学分析多变量积分进行类似的理解：上述求和相当于在点阵上的所有点相加，如图5.1，图中横轴代表 i，纵轴代表 j。因此我们可以用多种方法求和，既可以先按行、再按列求和，也可以先按列、再按行求和，最终结果一定是相同的，即

$$\sum_{i=1}^{m}\sum_{j=1}^{n} f(i,j) = \sum_{i=1}^{m}\left[\sum_{j=1}^{n} f(i,j)\right] = \sum_{j=1}^{n}\left[\sum_{i=1}^{m} f(i,j)\right] \tag{5.4}$$

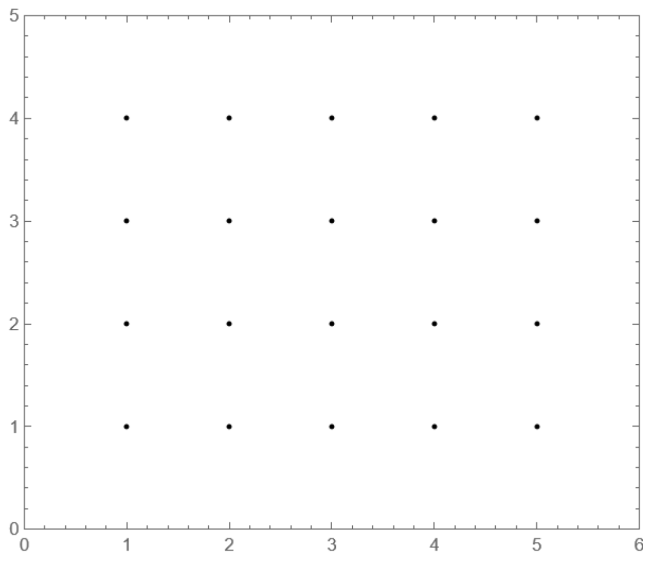

图 5.1: 多重求和与多重积分的类比

类似地可以处理一些略微复杂的问题，如

$$\sum_{i=1}^{m}\sum_{j=i}^{n} f(i,j) = \sum_{i=1}^{m}\left[\sum_{j=i}^{n} f(i,j)\right] = \sum_{j=1}^{n}\left[\sum_{i=1}^{\min\{j,m\}} f(i,j)\right], \text{assuming } n > m \tag{5.5}$$

通过图5.2可以很容易地理解。

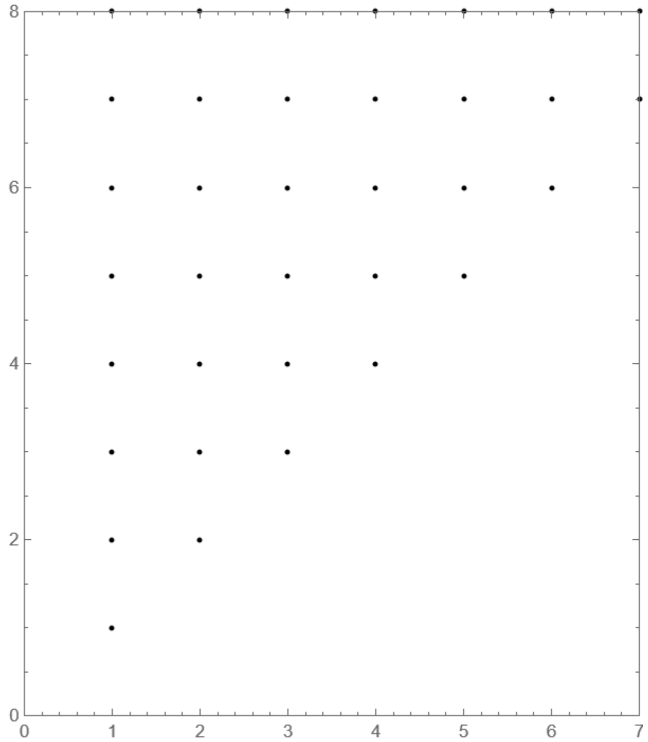

图 5.2: 求和中要求 $j \geq i$

再给出一个例子，如

$$\sum_{i=1}^{m}\sum_{j=i}^{n} f(i)g(j) = \left(\sum_{i=1}^{m} f(i)\right)\left(\sum_{j=i}^{n} g(j)\right) \tag{5.6}$$

可以类比为

$$\int\int f(x)g(y)dxdy = \int f(x)dx \int g(y)dy \tag{5.7}$$

在线性代数这门课，矩阵运算的基本方程 $c_{ij} = \sum\limits_{k} a_{ik}b_{kj}$，可以看作是保持 i, j 不变，消去 k 的过程。

最后，如果大家想知道自己对求和符号的掌握够不够，尝试用符号的形式证明矩阵分块的运算规则，了解矩阵分块的运算规则对线性代数本身也很有帮助。

拓展知识：爱因斯坦求和约定

6 线性代数的奇妙应用

6.1 转移矩阵

生活中很容易举出一些例子：人可能“饿”或者“不饿”，“饿”可能立刻去吃饭变成“不饿”，也可能等会再吃继续“饿”；“不饿”可能消化完变为“饿”，也可能继续“不饿”。

考虑一个系统，在 t 时刻有 $p_1 \geq 0$ 概率处于状态 1，有 $p_2 \geq 0$ 概率处于状态 2，接下来系统经历 A 过程，若系统处于状态 1，有 $a_{11} \geq 0$ 概率保持状态 1 不变，有 $a_{12} \geq 0$ 概率变为状态 2；若系统处于状态 2，有 $a_{21} \geq 0$ 概率变为状态 1，有 $a_{22} \geq 0$ 概率保持状态 1 不变。在下一时刻 $t+\delta t$，系统 $p_1' \geq 0$ 概率处于状态 1，有 $p_2' \geq 0$ 概率处于状态 2，很容易写出

$$\begin{cases} p_1' = p_1 a_{11} + p_2 a_{12} \\ p_2' = p_1 a_{21} + p_2 a_{22} \end{cases}$$

如果用线性代数的形式来理解，很容易写出

$$\begin{pmatrix} p_1' & p_2' \end{pmatrix} = \begin{pmatrix} p_1 & p_2 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

拓展到系统有 n 种可能的状态，系统从状态 i 演化至状态 j 的概率为 a_{ij}，**其核心在于在下一时刻系统处于状态 i 的概率可以用前一时刻处于状态 j 的概率通过线性组合表示**，即

$$p_i' = \sum_{j=1}^{n} p_j a_{ji}$$

则演化过程可以用转移矩阵 A 来表示，转移矩阵 A 满足 $\sum_{j=1}^{n} a_{ij} = 1$，即每行和为 1。

作为拓展知识，如果某个系统不断地经历 A 过程，若转移矩阵 A 的某个幂次的全部元素都是正的（正则矩阵），则系统的末态将趋向唯一确定的、与初态无关的定态。

拓展知识：Markov chain

6.2 有向图的邻接矩阵

这里不打算引入过多概念，仅仅介绍核心的思想和简单的定义，有兴趣的同学可以自行学习相关知识（图论）。

对于给定有向图 $D=(V,E)$，设 $V=v_1,v_2,\cdots,v_n$，定义 D 的邻接矩阵为

$$A(D) = (a_{ij})_{\nu\times\nu}$$

其中 a_{ij} 为图 D 中以 v_i 为尾、以 v_j 为头的边数（从 v_i 到 v_j 的边数）。

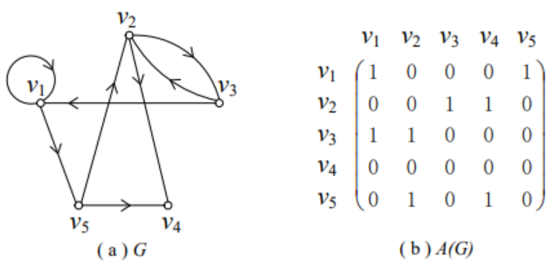

图 6.1: 有向图邻接矩阵

现在已经定义了 $A(D)_{ij}$ 是从 v_i 到 v_j 长为 1 的路径，如果想考虑从 v_i 到 v_j 长为 2 的路径，应该如何计算？

最简单地从 v_i 到 v_j 长为 2 的路径一定是由 $v_i \to v_k \to v_j$ 拼接起来的，因此计数时可以对 v_k 遍历，如果 $v_i \to v_k$ 有 a_{ik} 条路径，$v_k \to v_j$ 有 a_{kj} 条路径，那经过 v_k 的路径会贡献 $a_{ik}a_{kj}$ 种不同的、长为 2 的路径。所有 v_i 到 v_j 长为 2 的路径只需对 v_k 求和，设总路径数量为 c_{ij}，即有

$$c_{ij} = \sum_k a_{ik}a_{kj} \Rightarrow C = A^2 \tag{6.1}$$

可以看到这个问题的核心就在于，我们只关心 v_i 到 v_j 长为 2 路径的结果（c_{ij}），并且计数之后是对所有经历点的求和（$\sum_k$），这恰恰对应了矩阵运算的法则。反过来，也可以直接通过计算 $A^2(D)$ 获得所有我们想知道长为 2 路径的信息，而无需对每两个点分别分析。

作为课后思考，$A^l(D)$ 的含义是什么？如果只关心“是否存在从 v_i 到 v_j 长为 k 的路径”，而不关心“从 v_i 到 v_j 长为 k 的路径数”，矩阵 A 和计算过程应该进行怎样的优化？

作为拓展知识，连通图的关联矩阵的秩等于图的点数减去图的连通片个数，还可以通过连通图的关联矩阵的行列式计算图的生成树数量。

拓展知识：无向图的连通矩阵；Warshall 算法

A 约化标准形唯一性的证明

下面给出链接中证明的简单思路：使用归纳法，假设 A 是个 $m \times n$ 的矩阵，从 $m \times 1$ 开始，通过归纳 n 证明约化标准形唯一。

对 $n = 1$，显然只有一种标准型；设 A' 是 $\boldsymbol{A}$ 删掉第 n 列得到的矩阵，假设 A' 有唯一的约化标准形，现在考虑 A 的约化标准形，通过初等行变换只可能在最后一列有差别（生成约化标准形的过程是逐列生成的），假设 A 有两个约化标准形 $B, C, B \neq C$，考虑 u 是矩阵 A 的通解，$Au = 0$，由于初等行变换不改变解的性质，也有 $Bu = 0, Cu = 0$，则 $(B - C)u = 0$。B 与 C 的前 $n - 1$ 列相同，因此 $B - C$ 的前 $n - 1$ 列全为 0。假设 $B - C \neq 0, (B - C)_{i,m \leq n-1} \equiv= 0$，一定存在 j 使得 $(B - C)_{jn} \neq 0$。考虑 $(B - C)u$ 的第 j 行，有 $\{(B - C)u\}_{j1} = \sum_{k=1}^{n}(B - C)_{jk}u_{k1} = (B - C)_{jn}u_n = 0$，则 $u_n = 0$。由于 u 是 $Bu = 0, Cu = 0$ 的解，且要求 $u_n = 0$，说明 B, C 的第 n 列一定包含一项 "1" 作为主元（否则全为 0，则 u_n 可以有多解）。但 B, C 的前 $n - 1$ 列完全相同，第 n 列的 "1" 一定出现在同一位置。对于约化标准形，主元上下方的元素一定都为 0，说明 B 与 C 的第 n 列也相同，$B = C$ 与假设 $B \neq C$ 产生矛盾，即可说明若 A 删掉最后一列的约化标准形唯一，A 的约化标准形也唯一，通过归纳法即可知对任意的 A，A 的约化标准形也唯一。

致谢

感谢胡铁宁、贺维易同学对本文档的校对工作和内容补充，感谢申伊塃老师以及同学对助教工作的支持。