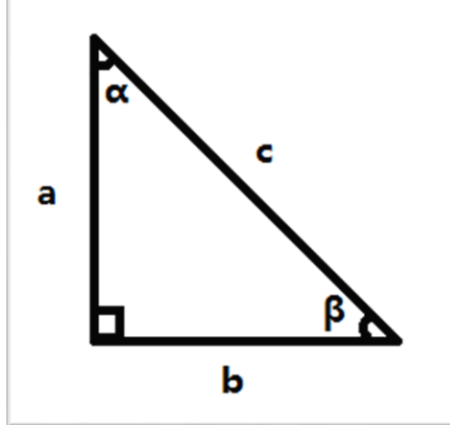
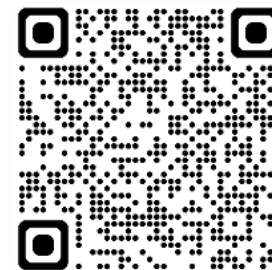




坐标向量长度呼唤二次型

蒋剑剑@宁德师范学院 math3j@qq.com



缘起



Interactive Linear Algebra

Dan Margalit, Joseph Rabinoff

- 从智能时代的应用来说，内积的占比呈上升趋势
- 引进抽象向量空间（大抽象）前充分铺垫小抽象
- 列向量空间之后自然应该充实几何讨论——度量
- 二次型的引入过于突兀，整体内容可以精炼简化

6 Orthogonality

Dot Products and
Orthogonality

Orthogonal Complements

Orthogonal Projection

Orthogonal Sets

The Method of Least Squares

LECTURES ON LINEAR ALGEBRA

L. M. GEL'FAND

Academy of Sciences, Moscow, U.S.S.R.

Translated from the Revised Second Russian Edition

by A. SHENITZER

Adelphi College, Garden City, New York

TABLE OF CONTENTS

	Page
Preface to the second edition	v
Preface to the first edition	vii
I. n-Dimensional Spaces. Linear and Bilinear Forms	1
§ 1. n -Dimensional vector spaces	1
§ 2. Euclidean space	14
§ 3. Orthogonal basis. Isomorphism of Euclidean spaces	21
§ 4. Bilinear and quadratic forms	34



导言	1	
第 1 章 线性方程组	9	16
1.1 什么是数? 域的公理	9	
1.2 消元法解线性方程组	12	
1.3 线性方程组线性换元	18	
1.4 线性方程组解的情况	23	
习题	25	
第 2 章 矩阵论初步	27	16
2.1 矩阵基本运算	27	
2.2 矩阵运算法则	30	
2.3 可逆矩阵和初等矩阵	32	
2.4 矩阵分块运算	35	
2.5 矩阵的秩与相抵关系	38	
2.6 矩阵若干应用	41	
习题	43	
第 3 章 行列式基础	47	17
3.1 低阶的行列式	47	
3.2 排列的逆序数	51	
3.3 行列式的定义	53	
3.4 行列式的性质	55	
3.5 按行或列展开	57	
3.6 矩阵与行列式	61	
习题	64	
第 4 章 有限维空间模型	69	16
4.1 列向量空间模型	69	
4.2 向量的线性关系	72	
4.3 极大线性无关组	75	
4.4 空间的基和维数	78	
4.5 坐标与基的变换	81	
4.6 再论线性方程组	83	
习题	85	

第 5 章 内积和双线性型	89	16
5.1 内积诱导空间度量	89	
5.2 正交化与空间分解	92	
5.3 双线性型与二次型	95	
5.4 二次型的分类问题	98	
5.5 正定二次型的刻画	102	16
习题	105	
第 6 章 一元多项式	107	21
6.1 带余长除法	107	
6.2 最大公因式	111	
6.3 既约多项式	114	
6.4 多项式函数	117	
6.5 整系数多项式	119	
6.6 部分分式分解	121	21
6.7 中国剩余定理	124	
习题	128	
第 7 章 向量空间的态射	131	19
7.1 向量空间公理化定义	131	
7.2 线性映射及矩阵表示	135	
7.3 线性变换的特征向量	139	
7.4 相似关系下的标准形	143	
7.5 代数基本定理的证明	147	19
习题	150	
第 8 章 欧氏空间算子论	153	

第 5 章 内积和双线性型

阅读提示

本章基本动机是要对列向量赋予长度概念，据中学几何知识想到了二次函数 $f(X) := x_1^2 + \cdots + x_n^2$. 一个关键观察是 $f(X)$ 可用矩阵 $X^T X$ 表达，在三角形中就容易演变成两个向量的点积 $X^T Y$. 进而考虑任意子空间当中的点积，择一基将子空间坐标化，转换到坐标空间 $\mathbb{R}^m$ 中考虑，于是向量的点积由形如 $g(X, Y) := X^T A Y$ ($A \in \mathbb{R}^{m \times m}$) 之式表达. 分析其基本性质，主要有**双线性**、**对称性**、**正定性**三方面，尤以双线性最为基本. 三性兼备者就称为内积，结构较为丰富，与经典几何直观相符. 所以本章先探讨内积，然后考察更为宽泛的双线性型，接着缩小到对称双线性型，最后对正定性深入刻画，也即回到内积，认识螺旋上升. 学习本章应明辨概念，注重数形结合，紧跟思路和直观，厘清每一步推理的逻辑，时常揣摩，体会从基本性质出发、严谨演绎的公理化方法.

标准内积的自然引入

$q(X) := x_1^2 + x_2^2 + \cdots + x_n^2$. 显然 $q: \mathbb{R}^n \rightarrow \mathbb{R}$ 可视作 $\mathbb{R}^n$ 上的一个二次函数, 而其式 (n 元二次齐次多项式) 可用矩阵语言写成 $X^T X$. 这是很要紧的一步, 它将问题纳入我们业已熟悉的矩阵基本理论之中. 进一步考虑最基本的几何图形——三角形, 设它由向量 $X, Y, X + Y$ 所围, 则三边长度平方存在关系

$$|X + Y|^2 = (X + Y)^T (X + Y) = X^T X + Y^T Y + X^T Y + Y^T X = |X|^2 + |Y|^2 + 2X^T Y.$$

这是因为一阶方阵 $Y^T X = (Y^T X)^T = X^T Y$. 设 $Y := (y_1 \ y_2 \ \cdots \ y_n)^T$, 于是

$$b(X, Y) := X^T Y = \sum_{i=1}^n x_i y_i = \frac{1}{2} [b(X + Y, X + Y) - b(X, X) - b(Y, Y)] \quad (5.1)$$

衡量了 X, Y 偏离直角的程度 (请回忆勾股定理), 而且它正好是中学接触过的点积 (数量积), 作为行将定义的更一般概念之特例, 也称为标准内积. 上式体现了长度和内积的相互决定, 常被称为极化恒等式.

子空间继承的标准内积

现在考虑 $\mathbb{R}^n$ 的子空间 V ，设其基矩阵为 $B := (\beta_1 \ \beta_2 \ \cdots \ \beta_m)$ 。则 V 中诸向量 α 可由其坐标 $X_\alpha \in \mathbb{R}^m$ 来代表，即 $\alpha = BX_\alpha$ 。于是 $\alpha, \beta \in V$ 的点积为

$$\alpha^T \beta = (BX_\alpha)^T (BX_\beta) = X_\alpha^T (B^T B) X_\beta.$$

这说明借由基“坐标化”，子空间 V 上的度量问题相应地转化为 $\mathbb{R}^m$ 上的度量问题；只不过用来反映度量的标准内积 $X^T Y$ 换成了更一般的内积 $X^T (B^T B) Y$ ；或者说，由标准度量矩阵 I_m 转换为更一般的度量矩阵 $G := B^T B \in \mathbb{R}^{m \times m}$ 。图方便，将两个向量 X 和 Y 的内积 $X^T G Y$ 简记为 (X, Y) 。先熟悉内积的一些基本性质：

以“正交投影”为基本工具

据几何直观，我们名正言顺地称 $P_Y(X) := \frac{(X,Y)}{(Y,Y)}Y$ 为 X 在 Y 上的（正交）投影. 进而称 $X - 2P_Y(X)$ 为 X 沿 Y -镜面的反射. 利用投影，也能获得 X 到直线 $\langle Y \rangle$ 的“距离向量” $X - P_Y(X)$ ：即与 Y 正交（垂直）， $(X - P_Y(X), Y) = 0$ ，且长度 $\|X - P_Y(X)\|$ 代表 X 到 $\langle Y \rangle$ 的距离. 所以 X 可正交分解为 $P_Y(X) + [X - P_Y(X)]$ ，由下一段“勾股定理”得 $\|X\|^2 = \|P_Y(X)\|^2 + \|X - P_Y(X)\|^2 \geq \|P_Y(X)\|^2$ ，化简后回到 Cauchy 不等式，且等号成立条件为 $X = P_Y(X)$ ，即 $X \in \langle Y \rangle$. 这一切暗示正交问题是内积的核心问题，后面会不断加深这个主题.

5.2 正交化与空间分解

上一节接触过一个向量 X 到另一向量 Y (张成的子空间) 的投影 $P_Y(X)$, 而 $X - P_Y(X)$ 即为垂直 (正交) 于 Y 的向量. 能否推广为向一般的子空间投影? 不妨先看看二维子空间 (平面) 情形, 它有很强的立体几何直观. 设二维子空间 W 由正交组 α_1, α_2 张成, 今有向量 β , 如何作 β 到 W 的投影? 或者说如何作 β 到平面 W 的“垂足”? 这其实是中学立体几何中“三垂线定理”之用: 设垂足为 η , 则 η 到 $\langle \alpha_i \rangle$ 的垂足也是 β 到 $\langle \alpha_i \rangle$ 的垂足; 即 $\eta = P_{\alpha_1}(\beta) + P_{\alpha_2}(\beta)$; 而 $\beta - \eta$ 即为 β 到 $\langle \alpha_1, \alpha_2 \rangle$ 的距离向量. 这就很容易推广到高维子空间了: 设 W 的一个正交基为 $\alpha_1, \dots, \alpha_s$, 那么向量 β 到 W 的 (正交) 投影为 $P_W(\beta) := \sum_{i=1}^s P_{\alpha_i}(\beta)$. 投影与距离向量 $\beta - P_W(\beta)$

利用正交化可将内积的度量矩阵 $G = B^T B$ 变得更“内蕴”. 回顾 §5.1 从 $\mathbb{R}^n$ 的点积出发, 对其子空间 V 任选一基 $\beta_1, \dots, \beta_m$ 构成基矩阵 B ; 通过坐标化将 V 上点积转化为 $\mathbb{R}^m$ 上内积 $(X, Y) := X^T G Y$. 由于 m 阶方阵 G 代表 $\mathbb{R}^m$ 的度量结构, 理应与 n 无关; 但现在用来定义 G 的基矩阵 $B \in \mathbb{R}^{n \times m}$ 却与 n 有关. 故有理由猜测 G 能改写成 $M^T M$ 的样子, 其中 $M \in \mathbb{R}^{m \times m}$. 用 Gram-Schmidt 正交化证之:

将 $\beta_1, \dots, \beta_m$ 正交化为 $\alpha_1, \dots, \alpha_m$, 则 $A := (\alpha_1 \ \cdots \ \alpha_m) = B(B \rightsquigarrow A)$. 令 P 为过渡矩阵 $(A \rightsquigarrow B) = (B \rightsquigarrow A)^{-1} \in \mathbb{R}^{m \times m}$, 立即有

$$G = B^T B = (AP)^T (AP) = P^T (A^T A) P =: P^T D P, \quad D = \text{diag}(\|\alpha_1\|^2, \dots, \|\alpha_m\|^2).$$

容易看出 $(B \rightsquigarrow A)$ 是对角元为 1 的上三角型矩阵, 于是其逆 P 也是对角元为 1 的上三角型矩阵 (为何?). 从而 P^T 是对角元为 1 的下三角型矩阵, DP 是上三角型矩阵. 这说明度量矩阵总可表为下三角与上三角型矩阵之积 (LU 分解). 再令 $\sqrt{D} := \text{diag}(\|\alpha_1\|, \dots, \|\alpha_m\|)$, 则 $D = \sqrt{D}^T \sqrt{D}$, 从而 $G = (\sqrt{D} P)^T (\sqrt{D} P)$.

顺便说下, 据“体积 = 底 $\times$ 高”的直观, 我们可定义向量组 $\beta_1, \dots, \beta_m$ 所围的平行体之体积为

$$\text{vol}(\beta_1, \dots, \beta_m) := \|\alpha_1\| \cdots \|\alpha_m\| = \det \sqrt{D} = \sqrt{|D|} = \sqrt{|P^T D P|} = \sqrt{\det G}.$$

重新看待线性方程组

我们用上述几何观点重新看待齐次线性方程组. 设 $A := (a_{ij}) \in \mathbb{R}^{m \times n}$, 而齐次线性方程组 $AX = 0$ 的解空间为 U . 将每个线性方程看成点积的形式, 即设 A^T 的列向量组为 $\alpha_1, \dots, \alpha_m$, 则 $AX = 0$ 等价于 $(\alpha_i, X) = \alpha_i^T X = 0, 1 \leq i \leq m$, 即 $U = \langle \alpha_1, \dots, \alpha_m \rangle^\perp$. 这说明求正交补空间相当于解齐次线性方程组. 特别地, 若 $\text{rank } A = m = n - 1$, 则 U 是超平面 $\langle \alpha_1, \dots, \alpha_m \rangle$ 的法线, 其中任何一个非零向量都可作为法向量. 利用行列式, 一个法向量可形式化地表示为

$$\eta := \begin{vmatrix} \varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_n \\ a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{vmatrix}.$$

5.3 双线性型与二次型

我们已经看到，内积赋予向量空间度量结构，充实了几何性质的讨论. 所以本节回头重新审视内积概念，并做一些推广. 数学史表明，恰当推广重要概念往往能催生重大进展；如将 $\mathbb{R}^4$ 的点积 $x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4$ 换成 $x_1y_1 + x_2y_2 + x_3y_3 - x_4y_4$ ，即 Minkowski¹ 度量，便能反映相对论的四维“时空”，而 Lorentz 变换唾手可得.

回想 $\mathbb{R}^m$ 上的内积 $(X, Y) := X^T G Y$ ，其中度量矩阵 G 形如 $B^T B$. 我们注意到 $G^T = G$ ，这反映了内积的**对称性**；矩阵 G 还满足**正定性**，即 $X^T G X \geq 0$ 且仅当 $X = 0$ 时才有 $X^T G X = 0$. 如果放弃这些特性，用一般的矩阵 $A \in \mathbb{F}^{m \times m}$ 构造从 $\mathbb{F}^m \times \mathbb{F}^m$ 到 $\mathbb{F}$ 的映射（函数） $\phi(X, Y) := X^T A Y$ ，那它仅继承了内积的**双线性**：

▷ **左线性** $\phi(k_1 X_1 + k_2 X_2, Y) = k_1 \phi(X_1, Y) + k_2 \phi(X_2, Y).$

▷ **右线性** $\phi(X, k_1 Y_1 + k_2 Y_2) = k_1 \phi(X, Y_1) + k_2 \phi(X, Y_2).$

¹著名数学家 Minkowski 是相对论创立者 Einstein 的大学老师.

5.4 二次型的分类问题

回想内积给我们带来很多好处和丰富结构，但它不过是由经典的长度平方函数 $f(x_1, \dots, x_n) := x_1^2 + \dots + x_n^2$ 引出的. 这是特殊的二次型，用它对长度概念建模反映经典时空观；相对论却表明，用 Minkowski 度量 $g(x_1, x_2, x_3, x_4) := x_1^2 + x_2^2 + x_3^2 - x_4^2$ 对时空建模才更符合现实. 所以到底该用哪种二次型当由实践（实验）决定，数学的任务则是分辨各种可能性，即要对二次型分类.

另一方面，上节末也指出，二次超曲面方程的要素即为二次型，要想搞清楚二次超曲面的形态，同样要对二次型分类. 在曲线（平面几何）情形，通过简单换元化为标准方程，众所周知，仅有两大类（有心非退化的）二次曲线，即椭圆（包括圆）和双曲线，标准方程分别为 $ax^2 + by^2 = 1$ 和 $ax^2 - by^2 = 1$ ，其中参数 $a, b > 0$. 因正数总可以开方，再通过简单的伸缩变换可将这两类方程化至极简，即 $x^2 \pm y^2 = 1$. 通常称具备 Minkowski 度量的空间为**双曲空间**，并非说这种空间是双曲面，而是其度量

5.5 正定二次型的刻画

我们返回内积这一重要主题. 在经典长度和向量点积的启发下, §5.1 将 $\mathbb{R}^n$ 上的内积定义为形如 $g(X, Y) := X^T(B^T B)Y$ 的对称双线性型, 或者等价地, 形如 $f(X) := X^T(B^T B)X$ 的二次型. 已知内积具**正定性**, 即 $f(X) \geq 0$ 且仅当 $X = 0$ 时 $f(X) = 0$; 为方便阐述, 将**正定二次型**的矩阵称为**正定矩阵**. 照此定义易见正定矩阵之和还是正定矩阵. 而且 §5.4 也提到, 实对称矩阵可描述刚体的惯性张量, 它在刚体力学中的角色类似于质量在质点力学中的角色; 众所周知质量是一个正数, 那么惯性张量应该也有某种“正性”, 这就是矩阵的正定性.

本章前几节仅利用双线性、对称性和正定性就获得了丰富的信息, 我们自然想问: 这三性是否就决定了内积的概念? 换言之: 是否正定的对称双线性型必为内积? 答案在 §5.4 之后就昭然若揭了; 结果是肯定的, 理由如下.

自然开启下一章：一元多项式的算术

设可逆矩阵 B 使 $A = B^T B$, 则 $A^{-1} = B^{-1}(B^{-1})^T = P^T P$, 其中 $P := (B^{-1})^T$, 表明 A^{-1} 确实也正定. 另一方面, 任何一个正定矩阵 $A = B^T B$ 的 $m \in \mathbb{N}$ 次幂 $\underbrace{B^T B B^T B \cdots B^T B}_{\text{共 } m \text{ 对}}$ 可写成 $Q^T Q$ 的样子, 因而也是正定的. 于是 $A^{-m} := (A^{-1})^m$ 都

正定. 此外对任何正数 k , 显然 kA 也正定. 而正定矩阵之和亦正定, 说明由正定阵 A 经过正数乘、加法、乘法、除法（取逆）混合运算所得的矩阵都正定, 实际上这样运算出来的矩阵都形如 $(a_m A^m + \cdots + a_1 A + a_0 I)(b_s A^s + \cdots + b_1 A + b_0 I)^{-1}$, 其中 $a_i, b_j \in \mathbb{R}$ 都是正数. 它们可看成分式在矩阵 A 处的取值, 而分式无非由分子和分母两个多项式组成. 研究多项式在矩阵上的取值将构成后续章节的一个主旋律.

坐标向量长度呼唤二次型

Thanks for Attention.



蒋剑剑@宁德师范学院 math3j@qq.com