

肖书奇

考研数学一

2022 年

目录

关于数学学习

知识的三种属性与其理想载体

- 理解性:故事书
- 程序性:流程图
- 记忆性:工具书

如何讲好故事

定义(结构)是逻辑上的起点,却是学科发展的终点。明确各个定义的逻辑关系,阐明其背景与动机。

定理(性质)分清主次,区分基本性质与重要定理。¹

同一个对象,追求不同的理解角度;不同的对象,看出联系。

¹ 体系的核心与逻辑脚手架,自然的推论与惊人的发现
自然语言与形式语言默写定义和定理,指出关键细节(如限制条件)。定理需要掌握证明思路与使用方法,收集“触发器”。

待填

第一章

- 区分公理与自然推论

第二章

- 连通性

第三章

- 上下极限与根、比值审敛法的应用
- 柯西稠密审敛法的应用
- 非主干知识
- 上下极限性质的证明

第四章

- 洛必达法则的特例
- 所有的不定式
- 柯西余项的泰勒定理
- 所有的证明

Part I

高等数学

实数、复数、Euclidean 空间

实数

实数的公理化定义^{2 3}: 有序域、最小上界性质

² 戴德金分割提供了具体构造方法
³ 预备知识: 域、有序集、有序域、界、最小上界、最小上界性质

复数

定义复数及其加法、乘法。定义虚数单位⁴, 定义实部、虚部与共轭。

⁴ 提供了另一种等价的复数刻画, 使复数乘法与实数乘法相协调。

Euclidean 空间

定义 Euclidean 向量及其加法、数乘、内积、范数。

拓扑基础

可数集与不可数集

根据与自然数建立等价映射的可行性,为“无限”分类。⁵

度量空间

距离函数的公理化定义:半正定性、对称性、次可加性。

点集拓扑的 12 个基本概念:邻域、内点、聚点、孤立点、开集、闭集、有界集、完美集、稠密集、凸集、紧集、连通集。⁶

紧集

利用开覆盖定义紧集。⁷ 核心是实数/复数/Euclidean 空间两个常用的紧性定理,Borel-Heine 定理和 Weierstrass 定理。



图 1: 集合的分类

⁵ 经典命题 (Cantor 对角线法)

- 可数个可数集的并集可数
- 有理数可数
- 实数不可数

⁶ 关于邻域、内点、聚点、开集、闭集的基本性质

- 邻域是开集
- 某集合的聚点的任意邻域包含该集合的无穷元素
- 有穷集合不可能有聚点
- 某集合是开(闭)集等价于其补集为(开)集
- 任意个开集的并集是开集;任意个闭集的交集是闭集;任意有限个开集的交集是开集;任意有限个闭集的并集是闭集
- 集合的封闭是包含它的最小闭集;集合是闭集等价于与自身的封闭相等
- 集合的开/闭具有相对性;集合是相对开集(相度量空间的某个子集是开集)等价于其等于某个绝对开集(相整个度量空间的开集)与背景集合的交集
- 实数上的有界集若是闭集,则其确界在集合内部

⁷ 关于紧集的基本性质

- 紧性与嵌入的背景集合无关(绝对紧性等价于相对紧性)
- 紧集是闭集
- 紧集的闭子集是紧集
- 任意紧簇中的任意有限紧集的交集非空,则该紧簇的交集非空;单调缩小的非空紧集序列的交集非空;
- 实数域的闭区间是紧集;Euclidean 空间的 k-cell 是紧集(Bolzano 二分区间法)

序列与级数

序列

定义收敛序列⁸,重要结论是在实数、复数、Euclidean 空间上的收敛序列的极限运算与算术运算的关系。

定义子列⁹,重要结论是任意有界序列必存在收敛子列。(Bolzano-Weierstrass 致密性定理)

定义 Cauchy 列¹⁰,重要结论是实数、复数、Euclidean 空间的柯西列必然收敛¹¹。定义完备性。

定义实数序列的单调性,重要结论是单调且有界的序列必然收敛。

定义实数序列的上、下极限。¹²

经典实序列极限如下:

参数	收敛序列	极限
$p > 0$	$\lim_{n \rightarrow \infty} \frac{1}{n^p}$	0
$p > 0$	$\lim_{n \rightarrow \infty} \sqrt[n]{p}$	1
$p > 0 \wedge \alpha \in \mathbb{R}$	$\lim_{n \rightarrow \infty} \frac{n^\alpha}{(1+p)^n}$	0
$ x < 1$	$\lim_{n \rightarrow \infty} x^n$	0
\searrow	$\lim_{n \rightarrow \infty} \sqrt[n]{n}$	1

级数

在实、复数中定义部分和序列,定义数项级数¹³,重要结论是基于柯西收敛准则刻画级数敛散性的等价条件。

数项级数的敛散判定是重点。比较审敛法¹⁴、柯西稠密审敛法¹⁵、根审敛法¹⁶、比值审敛法¹⁷

经典数项级数如下(比较审敛法和柯西稠密判定法可判定):

⁸ 收敛序列的基本性质

- 极限唯一
- 收敛序列有界
- $\{p_n\}$ 收敛于 p 等价于 p 的任意邻域包含除了 $\{p_n\}$ 有限个点以外的所有点
- $E \subset X$ 有聚点 p ,必存在 $\{p_n\} \subset E$ 以 p 为极限

⁹ 子列的基本性质

- 序列的全部收敛子列的极限构成的集合是闭集

¹⁰ Cauchy 列的基本性质

- 收敛序列必是柯西列

¹¹ 本质是紧度量空间上的柯西列必收敛

¹² 上、下极限的基本性质

- $E := \{\text{全部收敛子列的极限值}\}$
- $s^* \in E, s_* \in E$
 - $\forall x > s^* \Rightarrow (\exists N \in \mathbb{N} \forall n > N (x > s_n))$
 - $\forall x < s_* \Rightarrow (\exists N \in \mathbb{N} \forall n > N (x < s_n))$
 - $(\exists N \in \mathbb{N} \forall n > N (s_n \geq t_n)) \Rightarrow (s^* \geq t^* \wedge s_* \geq t_*)$

¹³ 数项级数的基本性质

- 收敛的数项级数的通项收敛且极限为零
- 通项为非负数的实数项级数的收敛性等价于有界性

¹⁴ 比较审敛法

- 敛:除了有限项,通项的绝对值小于等于某收敛级数的通项
- 散:除了有限项,通项大于等于某发散级数的通项的绝对值

¹⁵ 柯西稠密审敛法:若通项 a_n 是单调递减的非负数列,则 $\sum a_n$ 与 $\sum 2^k a_{2^k}$ 敛散性相同。

¹⁶ 根审敛法: $\alpha := \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|}$

- $\alpha < 1, \sum a_n$ 收敛
- $\alpha > 1, \sum a_n$ 发散
- $\alpha = 1, \sum a_n$ 敛散性未知

¹⁷ 比值审敛法:

- 若 $\limsup_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| < 1$, 则 $\sum a_n$ 收敛
- 若 $\exists N \in \mathbb{N} \forall n > N \left(\left| \frac{a_{n+1}}{a_n} \right| \geq 1 \right)$, 则 $\sum a_n$ 发散

另外,若根审敛法无法判断,则比值审敛法同样无效。

参数	级数	敛散性
$0 \leq x < 1$	$\sum x^n = \frac{1}{1-x}$	收敛
$x \geq 1$	$\sum x^n$	发散
$p \leq 1$	$\sum \frac{1}{n^p}$	发散
$p > 1$	$\sum \frac{1}{n^p}$	收敛
$p \leq 1$	$\sum \frac{1}{n(\log n)^p}$	发散
$p > 1$	$\sum \frac{1}{n(\log n)^p}$	收敛
\searrow	$\sum \frac{1}{n \log n \log \log n}$	发散
\searrow	$\sum \frac{1}{n \log n (\log \log n)^2}$	收敛

定义函数项级数, 定义幂级数(实/复数域)。重要结论是幂级数收敛域是圆形, 及其收敛半径的表达式。¹⁸

定义级数的绝对收敛。重要结论是级数绝对收敛是收敛的充分非必要条件。

$$^{18} R := \frac{1}{\limsup_{n \rightarrow \infty} \sqrt[n]{|c_n|}}$$

连续

定义映射在某点处的极限。重要结论是沟通了序列极限与映射极限的 Heine 定理¹⁹, 以及实数、复数、Euclidean 空间函数极限运算与算术运算的关系。

定义映射在某点处的连续性。重要定理是连续性与极限的关系²⁰; 连续映射的复合仍连续; 实、复数、Euclidean 空间里的各种基本运算(加、减、乘、除、向量加法、数乘、内积、直积、限定)保持映射的连续性。

连续映射具有丰富的拓扑性质, 可以利用开、闭集作连续性的等价刻画。重要定理是连续映射与紧集、连通集的关系: 连续映射将紧集映射为紧集(有界性定理、最值定理), 将连通集映射为连通集(介值定理)。

定义一致连续, 显然一致连续蕴含连续, 重要定理是紧集上的连续与一致连续等价。

对于以实数为自变量的函数, 可以定义左极限与右极限, 从而为不连续点分类。重要定理是关于单调函数的不连续性。

定义实、复数、Euclidean 空间的扩充与广义邻域, 定义广义极限(无穷处的极限)与无穷极限(极限值是无穷)。重新考察了极限(包含广义与无穷)与基本运算的关系, 从而为无穷定义了几种运算, 悬而未决的不定式是重点²¹。

¹⁹ 映射极限的基本性质

- 映射在某点若存在极限则极限唯一

²⁰ 特例: 孤立点

²¹ 常见不定式

- $\frac{\infty}{\infty}$
- $\frac{0}{0}$

导数

定义实函数²²的导数与单侧导数^{23 24}。重要定理是求导运算与算术运算的关系;复合函数的求导链式法则。

定义极小、极大值。重要定理是若极值点可导²⁵,则导数为零。

三大微分中值定理²⁶,掌握其应用。

导数未必连续,但必有介值性。

定义高阶导数。

L'Hospital's 法则提供了某种不定式极限的考察方法,注意其应用条件。²⁷

Taylor's 定理阐明了最佳的局部多项式近似与各阶导数的关系。²⁸

²² 自、因变量皆为实数

²³ 理解导数的两种方式:

- 瞬时变化率 (Newton)
- 最佳的局部线性近似 (Leibniz)

²⁴ 导数的基本性质

- 若可导,导数唯一
- 可导蕴含连续

²⁵ 极值点处可能不可导

²⁶ 微分中值定理的前提条件

- 闭区间连续 $f, g \in C^0[a, b]$
- 开区间可导 $\forall x \in (a, b) \exists f'(x), g'(x)$

Rolle 中值定理

$$f(a) = f(b) \Rightarrow (\exists \xi \in [a, b] f'(\xi) = 0)$$

Lagrange 中值定理

$$\exists \xi \in (a, b) f(\xi) = \frac{f(b) - f(a)}{b - a}$$

Cauchy 中值定理

$$\exists \xi \in (a, b) ((f(a) - f(b))g'(\xi) = (g(a) - g(b))f'(\xi))$$

²⁷ L'Hospital's 法则

- $a \in \bar{\mathbb{R}}, f(x)$ 和 $g(x)$ 都在 a 点的单侧开邻域可导
- $\lim_{x \rightarrow a} \frac{f'(x)}{g'(x)}$ 存在

并且满足以下其一

1. $x \rightarrow a, f(x) \rightarrow \infty \wedge g(x) \rightarrow \infty$
2. $x \rightarrow a, f(x) \rightarrow 0 \wedge g(x) \rightarrow 0$

则

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a} \frac{f'(x)}{g'(x)}$$

²⁸ $f \in C^n$

$$f(x) = f(a) + \sum_{k=1}^n \frac{f^{(k)}(a)}{k!} (x-a)^k + o((x-a)^n)$$

Riemann-Stieltjes 积分

积分两要素:被积对象与积分区域。Riemann-Stieltjes 积分是实函数在闭区间上的积分。

利用分割、Darboux 上/下和、Darboux 上/下积分,定义 R-S 积分²⁹。众多可积性定理给出了 R-S 可积的充分条件³⁰。

易证 R-S 积分的基本性质³¹。

几个经典的 α 构造展示了 R-S 积分相对 Riemann 积分的优越之处——利用单位阶跃函数将积分值转化为级数极限,利用 α 的导数将 R-S 积分转化为 Riemann 积分(重要),利用 α 的复合可完整地说明积分的换元公式(重要)。³²

定义变上限积分³³以后,微积分基本定理呼之欲出。

分部积分公式对应的是乘法的导数运算。

²⁹ Darboux 定义的基本性质

- 若分割更加精细,则达布上和会更小,达布下和会更大(也可以不变)
- 达布下积分小于等于达布上积分
- R-S 可积等价于存在分割使达布上下和任意地接近
- R-S 可积等价于“达布任意和”与积分值可以任意地接近(定积分求的是面积的精确值,只要你承认长方形的面积等于长乘以高)

³⁰ R-S 可积性定理

- 闭区间上的连续函数可积
- 闭区间上的单调函数可积(α 连续)
- 闭区间上的有界且间断点可数的函数可积(α 在被积函数的间断点处连续)
- 可积函数的复合函数仍可积
- 可积函数的加、减、乘、除仍可积

³¹ R-S 积分的基本性质

- 双线性
- 区间可加性
- 抽样性质($\alpha :=$ 单位阶跃函数,等价于与 Dirac 函数卷积)

³² R-S 积分的很多限制条件,工科生可能永远不会遇到。目前理论有限,对于各种条件的说明也不够精确。

³³ 首先,变上限积分是一个函数,是借助 R-S 可积函数自然构造出来的一个函数。它的性质很好:在整个闭区间上连续;若被积函数在某处连续,则它在该处可导。

Part II

线性代数

线性映射

线性映射的概念

线性映射构成的线性空间

有限维线性映射的基本定理

线性映射的矩阵

线性变换与其逆的概念

积空间与商空间

对偶与秩

在线性代数中,我们接触的第一个对偶关系是向量和线性泛函的对偶。³⁴ 顺理成章,有

³⁴ 类似“点动成线,线动成点”。

- 线性空间的对偶(对偶空间)属于线性泛函空间;
- 线性空间的基的对偶(对偶基)属于线性泛函空间的基;
- 线性空间之间的映射的对偶(对偶映射)属于线性泛函空间之间的映射。

严格定义如下:³⁵

³⁵ 需额外证明对偶基是基。重点关注对偶映射的构造方式。

$$V' := \mathcal{L}(V, \mathbb{F})$$

$$\{v_k\} =: \text{A basis of } V, \quad \varphi_k(v_j) := \begin{cases} 1 & j = k \\ 0 & j \neq k \end{cases}$$

$$T: V \mapsto W, \quad T'(\varphi) := \varphi \circ T \in \mathcal{L}(W', V')$$

根据定义,可得一些基本性质,见表格??。

随后利用零化子空间³⁶研究对偶映射的核与值域。定义为:³⁷

$$U \subset V, \quad U_V^0 := \{\varphi \in V' : \forall u \in U(\varphi(u) = 0)\}$$

脚手架:子空间与其零化子空间的维度之和是母空间维度。³⁸

$$\dim V = \dim U + \dim U^0$$

应用秩零定理,可以得到结论,见表格??。³⁹

命名	内容
核与值域的对偶	$\text{null } T' = (\text{range } T)^0$ $\text{range } T' = (\text{null } T)^0$
单与满的对偶	$T' \text{ is surjective} \Leftrightarrow T \text{ is injective}$ $T' \text{ is injective} \Leftrightarrow T \text{ is surjective}$
维度关系	$\dim \text{range } T' = \dim \text{range } T$ $\dim \text{null } T' = \dim \text{null } T + \dim W - \dim V$

最后研究对偶映射的矩阵。结论是矩阵的转置与线性映射的对偶等价。自然,对偶映射运算律对应转置矩阵运算律。

在这里适合给出矩阵的秩的概念,先定义行秩与列秩,⁴⁰

$$\text{column rank } A := \dim \text{span}\{\text{column vectors of } A\}$$

$$\text{rank rank } A := \dim \text{span}\{\text{rank vectors of } A\}$$

显然矩阵的列秩等于对应线性变换的值域的维度。再根据 $\dim \text{range } T = \dim \text{range } T'$,直接推出行秩等于列秩。

命名	内容
对偶同构律	$\dim V' = V$ $ST = T'S'$
对偶映射运算律	$(\lambda T)' = \lambda T'$ $(S + T)' = S' + T'$

表 1: 线性代数的对偶基本性质

³⁶ annihilator

³⁷ 可以为普通的子集 U 定义零化子空间,未必是子空间;定义没有揭示零化子空间是子空间,需额外证明。

³⁸ 证明思路: 关键在于零化子空间可视作对偶映射的零空间,即 $\forall T \in \mathcal{L}(U, V)$ ($\text{null } T' = U^0$), 再取特例 $T(u) = u$, 易得 $\dim \text{range } T' = \dim U$, 再应用秩零定理得证。

³⁹ 对偶映射的核与值域的关系对于有限维线性代数仍旧成立。(无需秩零定理即可证明)

⁴⁰ 行向量属于 $\mathbb{F}^{1,n}$, 列向量属于 $\mathbb{F}^{m,1}$

特征值、特征向量与不变子空间

本章研究线性变换。⁴¹由于线性变换只涉及一个线性空间，很自然的想法就是对这个空间作直和分解，分别考察线性变换的作用效果。⁴²但是，该线性变换对于各个子空间而言，未必是线性变换，可能是线性映射。首先，仅考虑不变子空间的情况——如果某个子空间包含了该线性变换在其上的值域，我们称之为不变子空间。

$$T \in \mathcal{L}(V), \quad \text{a subspace } U \text{ is invariant} := (u \in U \Rightarrow Tu \in U)$$

显然，对于任何线性变换，其零空间和值域都是不变子空间，这是平凡的例子。我们接着考虑最简单也最理想的不变子空间（一维）⁴³。根据线性空间与线性映射的定义，只能有这种情况： $\forall v \in V \exists \lambda \in \mathbb{F} \quad Tv = \lambda v$ 。特征值和特征向量的概念出现了。⁴⁴

$$\begin{aligned} \lambda \in \mathbb{F} \wedge (\exists v \in V (v \neq 0 \wedge Tv = \lambda v)) \\ := \lambda \text{ is an eigen value of } T \end{aligned}$$

$$\begin{aligned} \lambda \text{ is an eigen value of } T \wedge v \in V \wedge v \neq 0 \wedge Tv = \lambda v \\ := v \text{ is an eigen vector of } T \text{ corresponding to } \lambda \end{aligned}$$

根据定义，可以得到关于特征值与特征向量的几个等价条件。⁴⁵可以发现 $T - \lambda I$ ，这个线性变换是非常重要的脚手架，定义特征子空间：

$$E(\lambda, T) := \text{null}(T - \lambda I)$$

也容易证明特征值和特征向量的一些基本性质。⁴⁶

可以利用线性变换的不变子空间诱导出两种变换——限制变换和商变换。⁴⁷暂时没派上用场，先放这儿。

$$T|_U(u) := Tu$$

⁴¹ 线性变换的物理背景较为丰富。线性“变换”的一个特别之处：可以定义其幂运算，构造多项式。值得一提的性质是： $p(T)q(T) = q(T)p(T)$

⁴² 线性代数是关于表示的学问，直和表示空间和基的线性组合表示向量的思想是共通的。

⁴³ 直和分解的理想结果就是一维子空间

⁴⁴ “只能有这种情况”，所以“存在”即可，但是这个存在显然不能是零向量。

⁴⁵ 以下命题等价：

- λ 是 T 的一个特征值；
- $T - \lambda I$ 不是单射；
- $T - \lambda I$ 不是满射；
- $T - \lambda I$ 不可逆；

以下命题等价：

- v 是 T 关于特征值 λ 的特征向量；
- $v \in \text{null}(T - \lambda I) \wedge v \neq 0$

⁴⁶

- 不同特征值对应的特征向量必然线性无关（不同特征值对应的特征子空间之和是直和）；
- 不同特征值的个数至多是维度数目（不同特征值对应的特征子空间之和的维度小于等于母空间维度）。

⁴⁷ 根据定义有 $T|_U \in \mathcal{L}(U)$, $T/U \in \mathcal{L}(V/U)$ 。好好体会这种映射是线性、是变换。

$$(T/U)(v+U) := Tv + U$$

首先应关心存在性问题: 不变子空间/特征值/特征向量广泛存在于各种线性空间和线性变换吗? 一个重要结论是: **任何非零复线性空间上的任何线性变换, 必然存在特征值。**^{48,49}

下面研究线性变换的矩阵。回忆一下, 矩阵表达不仅取决于线性映射本身, 还取决于定义空间与到达空间的基的选取。对于线性变换而言, 定义空间与到达空间相同, 所以选取一组基即可。选基是一门很重要的学问, 我们的目标是为线性变换选取一组基, 使得对应矩阵的形式最简单。问题来了: 什么是最简单? 具体怎么选?

上面的定理已经保证了复线性空间至少有一个特征值, 不如把对应的特征向量作为一个基向量。显然, 它在矩阵中对应的列向量, 除了特征值以外, 其余分量全部是零。零这么多, 当然简单。然后我们进行基的扩张, 如果能够确保每次添加的基向量的变换结果在已有基向量张成的空间内, 我们就可以得到“上三角矩阵”。这可行吗? 答案是可行。另一个重要结论是, **对于任何非零复线性空间的任何线性变换, 都存在一组基使其矩阵为上三角矩阵。**⁵⁰

上三角矩阵这么简单, 有什么特别性质呢? 除了基本的几个等价条件⁵¹, 其可逆性的判定非常简单: **上三角矩阵 (或其表示的线性变换) 可逆, 当且仅当对角线元素皆为非零。**另外, 可以轻易地根据上三角矩阵找到对应线性变换的所有特征值——**线性变换的特征值就是其上三角矩阵的对角线元素。**⁵²

一旦特征值已知, 通过高斯消元法即可得到对应的特征向量。

更进一步, 上三角矩阵能不能更简单? 通过之前的推导, 我们可以很自然地想到对角矩阵, 这是线性变换“完全解耦”的表达形式。但是, **对于复线性空间, 并非所有的线性变换都存在对角矩阵。**^{53,54}

本章到此结束, 学点新东西后再杀回来。

⁴⁸ “非零”指不是仅有零元素的线性空间。

⁴⁹ 这说明我们对于不变子空间、特征值、特征向量的研究是有价值的。这个命题的证明方法很多。尽管此时还未引入行列式与特征多项式的概念, 但仍可以应用代数基本定理证明此结论。

⁵⁰ 证明主要基于数学归纳法, 对 $\text{range}(T - \lambda I)$ 的基进行扩张。

⁵¹

- T 关于基 v_1, \dots, v_n 的矩阵是上三角矩阵;
- $Tv_j \in \text{span}(v_1, \dots, v_j), \quad j = 1, \dots, n;$
- $\text{span}(v_1, \dots, v_n)$ 是 T 的不变子空间;

⁵² 显然, 可逆等价于所有特征值不等于 0。

⁵³ 关于可对角化的等价条件

- T 可对角化;
- V 存在一组基, 全部为 T 的特征向量;
- 存在 n 个一维不变子空间, 并且 $V = U_1 \oplus \dots \oplus U_n$;
- $V = E(\lambda_1, T) \oplus \dots \oplus E(\lambda_m, T)$;
- $\dim V = \dim E(\lambda_1, T) + \dots + \dim E(\lambda_m, T)$

⁵⁴ 关于可对角化的充分条件: 如果 T 含有 $\dim V$ 个不同的特征值, T 可对角化。

内积空间

线性空间是代数意义上对欧氏空间 \mathbb{R}^n 的推广,然而欧氏空间还有着丰富的几何结构,最基本的是长度和角度,这可不能放过。我们将把点积运算推广为内积运算,诱导出范数(长度)和夹角。而范数又可以诱导出距离函数,从而使极限、微积分成为可能。

如何为一般的线性空间⁵⁵定义内积? 类比之前对一般的集合赋予线性结构的方法,我们提取出点积的“本质”,为内积运算提出一些抽象的要求⁵⁶——首先内积是一个二元运算,得到的结果应在底域之内。还需满足:半正定性、可加性、齐次性、共轭对称性。⁵⁷ 随后可推出一些内积的基本性质。⁵⁸

范数可以由内积诱导,可验证内积诱导的范数满足范数的公理化定义。⁵⁹

$$\|v\| := \sqrt{\langle v, v \rangle}$$

角度可以由内积与范数诱导。实际上,我们最关心的是角度为 90° 的情况,也就是内积为零,这也是正交的定义。⁶⁰

$$\theta(u, v) := \arccos \frac{\langle u, v \rangle}{\|u\| \|v\|}$$

$$u, v \text{ are orthogonal} := \langle u, v \rangle = 0$$

万事俱备,可验证欧式几何最经典的勾股定理⁶¹;三角不等式⁶²;平行四边形等式成立;⁶³ 以及 Cauchy-Schwarz 不等式成立。⁶⁴

上一章研究不变子空间、特征子空间,用更优秀(精简)的矩阵揭示了线性变换的“本征”,这可以理解为选择耦合度低的方式去表达一个变换。本章研究内积与正交,是想用更优秀(易得)的坐标去定位一个向量,这可以理解为用耦合度低的方式去表达一个向量。基础确保起码能表达,正交基带来更好的表达。标准正交基广泛存在吗? 标准正交基怎么找? 标准正交基除了可以通过 n 个内积运算找到坐标,还有什么性质?

⁵⁵ 暂时只考虑底域为复数域或实数域上的线性空间。

⁵⁶ 公理化定义方法

⁵⁷ 内积的公理化定义

1. $\forall v \in V \langle v, v \rangle \geq 0$
2. $\langle v, v \rangle = 0 \Leftrightarrow v = 0$
3. $\forall u, v, w \in V \langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$
4. $\forall u, v \in V \forall \lambda \in \mathbb{F} \langle \lambda u, v \rangle = \lambda \langle u, v \rangle$
5. $\forall u, v \in V \langle u, v \rangle = \overline{\langle v, u \rangle}$

⁵⁸ 内积的基本性质

1. 固定一个元素后,内积是一个线性映射(线性泛函)
2. $\forall u \in V \langle 0, u \rangle = 0$
3. $\forall u \in V \langle u, 0 \rangle = 0$
4. $\forall u, v, w \in V \langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle$
5. $\forall u, v \in V \forall \lambda \in \mathbb{F} \langle u, \lambda v \rangle = \bar{\lambda} \langle u, v \rangle$

⁵⁹ 范数的公理化定义

1. $p(v) \geq 0$ (半正定性)
2. $p(av) = |a|p(v)$ (绝对一次齐次性)
3. $p(u + v) \leq p(u) + p(v)$ (三角不等式/次可加性)

⁶⁰ 内积为零,意味着某种不相关,无耦合关系。

⁶¹ $\langle u, v \rangle = 0 \Rightarrow \|u + v\|^2 = \|u\|^2 + \|v\|^2$

⁶² 范数的公理化定义包含次可加性

⁶³ $\|u + v\| + \|u - v\| = 2(\|u\|^2 + \|v\|^2)$ (四边边长的平方和等于两条对角线的平方和)

⁶⁴ $|\langle u, v \rangle| \leq \|u\| \|v\|$ (等号成立当且仅当 u 是 v 的数乘)

正交基

Gram-Schmidt 正交化程序直接给出了寻找标准正交基的方法(当然,这意味着**任何有限维内积空间都存在标准正交基**)。

我们上一章有结论——任意复指数空间的任意线性变换,都可以选取一组基使其矩阵为上三角形矩阵。我还想让这组基是标准正交的,我全都要,可以吗? ㄟ ㄏ。

Schur's 定理:任意复指数空间的任意线性变换,都可以选取一组标准正交基使其矩阵为上三角形矩阵;若线性空间的底域为 \mathbb{R} , 该定理能否成立取决于上三角形矩阵的存在性。⁶⁵

内积空间的线性泛函

显然,固定一个向量,内积运算就是一个线性泛函。逆天,其逆命题居然是成立的。

Riesz 表示定理:任何有限维线性空间的任何线性泛函都可以表示成内积的形式。

$$\forall \varphi \in \mathcal{L}(V, \mathbb{F}) \exists u \in V \forall v \in V (\varphi(v) = \langle v, u \rangle)$$

$$u = \overline{\varphi(e_1)}e_1 + \cdots + \overline{\varphi(e_n)}e_n$$

正交补与最优化问题

给定一个子集 U , 与全部元素正交的集合是其正交补 U^\perp 。⁶⁶

$$U^\perp := \{v \in V : \forall u \in U \langle v, u \rangle = 0\}$$

正交补的基本性质

- 正交补是子空间;
- $\{0\}^\perp = V \wedge V^\perp = \{0\}$
- $U \cap U^\perp \subset \{0\}$
- $U \cap W \Rightarrow W^\perp \cap U^\perp$

正交补提供了一种直和分解的自然方法。

$$U \text{ is a subspace of } V \Rightarrow V = U \oplus U^\perp$$

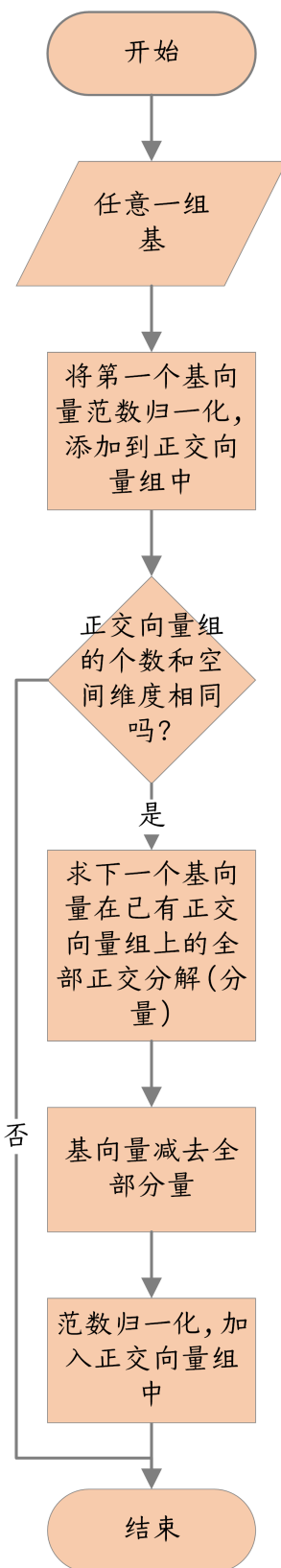


图 2: Gram-Schmidt 正交化程序

⁶⁵ 直接对上三角矩阵这组基进行 Gram-Schmidt 正交化即可, 上三角的等价条件并不会被破坏。

⁶⁶ U 未必是子空间

可以推出

$$\dim U^\perp = \dim V - \dim U$$

$$U = (U^\perp)^\perp$$

内积空间上的线性变换

自伴和正规变换

伴随映射⁶⁷

$$T : V \mapsto W, \quad T^* := \langle Tv, w \rangle = \langle v, T^*w \rangle$$

⁶⁷ 这是一种隐式定义, 用 Riesz 表示定理理解——固定 w , 泛函为 $\varphi(v) = \langle Tv, w \rangle$, 然后 $u := T^*w$

基本性质

- 伴随映射是线性映射 $T^* \in \mathcal{L}(W, V)$

