现代统计学丛书

高等计量经济学基础

缪柏其 叶五一 编著

高等教育出版社
HIGHER EDUCATION PRESS

现代统计学丛书

高等
计量经济学
基础

缪柏其 叶五一 编著

GAODENG JILIANGJINGJIXUE JICHU

高等教育出版社·北京
HIGHER EDUCATION PRESS BEIJING

北航

图书在版编目（CIP）数据

高等计量经济学基础 / 缪柏其，叶五一编著 . -- 北京：高等教育出版社，2013. 6

ISBN 978-7-04-037240-3

Ⅰ. ①高… Ⅱ. ①缪… ②叶… Ⅲ. ①计量经济学－研究生－教材 Ⅳ. ① F224.0

中国版本图书馆 CIP 数据核字（2013）第 073018 号

策划编辑 王丽萍　责任编辑 李华英　封面设计 王凌波　版式设计 马敬茹
责任校对 胡晓琪　责任印制 毛斯璐

出版发行	高等教育出版社
社　　址	北京市西城区德外大街4号
邮政编码	100120
印　　刷	北京中科印刷有限公司
开　　本	787 mm× 1092 mm　1/16
印　　张	20.25
字　　数	380 千字
购书热线	010-58581118
咨询电话	400-810-0598
网　　址	http://www.hep.edu.cn http://www.hep.com.cn
网上订购	http://www.landraco.com http://www.landraco.com.cn
版　　次	2013 年 6 月第 1 版
印　　次	2013 年 6 月第 1 次印刷
定　　价	69.00 元

本书如有缺页、倒页、脱页等质量问题，请到所购图书销售部门联系调换
版权所有　侵权必究

物 料 号　37240-00

前言

自 1926 年挪威经济学家 R. Frisch 提出计量经济学 (econometric) 以来, 至今已有 80 多年的历史了。它起源于对经济问题的定量研究。自改革开放以来我国学者开始向西方国家的学者学习对经济学的定量研究, 也促进了我国对计量经济学的教学和研究。在这个大背景下, 中国科学技术大学在 1995 年开办了应用经济学专业下的金融学, 开始了计量经济学的教学与研究。

计量经济学是以数理统计和数学为工具对经济问题进行研究的学科。因此, 这是一门交叉学科, 但它不只等于经济学和数理统计的简单叠加, 而是一种交叉和相互融合。计量经济学对定量研究经济和金融的重要性是不言而喻的, 国内外的专家学者在文献中早已进行了大量的阐述。现在的问题是如何进行这方面的教学。一般在本科生和研究生阶段都会开设这门课程, 关键是如何界定这两个阶段的教学内容和深浅, 特别是研究生这个阶段。由于所研究的数据的规模越来越大, 计算机硬件容量和速度的更新周期越来越短, 数理统计学中有关统计方法和统计工具越来越多, 可用来处理数据的软件也越来越多, 因此计量经济学的研究范围也越来越大, 所用到的统计方法、统计工具和计算机软件也越来越多。那么什么是学生在研究生阶段必须要掌握的? 什么是应该了解的? 国内的教师和学者在不断地探索着这个问题。本教材是我们对这门课程在研究生阶段教学的一种尝试。

计量经济学中研究的变量一般是随机的, 有的变量还是定性变量, 这是比较符合实际的。我们的任务就是用统计方法根据数据来建模, 然后从所建的模型来解释经济变量之间的关系。由于经济是一个动态的复杂系统, 很难用一种模

型来拟合, 所以我们就要尝试用不同的思路、不同的模型来尽可能好地拟合实际数据, 这就需要计量经济学家掌握较宽和较深的经济学和统计学的知识。这是我们对研究生阶段学生要求的基本认识, 也是本书的一种定位。

本书分为五个部分, 第一部分是有关的数学和统计学基础。为了能够理解后序章节和文献中的定量统计方法, 一定的数学和统计学知识, 特别是线性代数和多元统计的知识是必不可少的。而对大部分研究生而言, 在本科没有系统学过, 甚至没有学到这部分内容。所以我们用两章的篇幅来综述在后续章节中要用到的知识点, 特别介绍了随机变量或随机向量之间的相关关系和 Copula。第二部分是有关回归分析的, 包括建模以及参数和非参数估计, 这是计量经济学研究的基本方法之一。在这部分我们用投影这种比较直观的叙述来介绍线性和非线性回归的内容。第三部分是有关时间序列分析的。经济是一个动态系统, 必须用动态的模型来拟合, 时间序列分析是一种有力的工具, 我们从时间域和频率域两个方面简述了有关的内容, 介绍了有用的差分方程。第四部分是面板数据分析。回归是在横截面上分析经济变量之间的因果关系, 时间序列分析主要是研究变量随时间的变化规律, 没有做横截面分析, 面板数据分析则综合了这两种分析, 由于同时考虑经济变量和动态变化, 所以能考虑更广的经济问题, 当然也加大了分析的难度。第五部分是关于因变量为属性数据时的建模问题。

在具体的教学过程中, 可以根据课时和学生是否学过有关内容来灵活安排, 有些内容可以跳过, 如第一章第一节的线性代数基本知识, 第二章第六节的 Copula 函数等。

本教材虽然在研究生教学中使用多年, 但是一定还有不少遗憾之处, 请读者尽可能指出, 以便能不断完善我们的教材。

缪柏其、叶五一

2013 年 2 月 18 日

于中国科学技术大学

试读结束: 需要全本请在线购买: www.ertongbook.com

目录

第一章 线性代数和矩阵基本知识

§1.1 线性代数基本知识

§1.1.1 向量空间

我们知道，线性空间是线性代数研究的最基本的几何对象，我们最关注的是有关线性空间的投影理论. 矩阵是研究线性空间中各种几何问题的最有效的工具. 我们假定读者已有线性代数的基本知识，所以本节我们仅对有关概念作一个简要介绍.

定义 1.1 设 S 是一个空间，有零元素 $\mathbf{0}$，对空间的任意元素定义了加法运算和与实数的数乘运算. 如果运算满足如下条件，则称 S 为线性空间:

设 $\boldsymbol{x} \in S, \boldsymbol{y} \in S, \boldsymbol{z} \in S$，有

$$\boldsymbol{x}+\boldsymbol{y}=\boldsymbol{y}+\boldsymbol{x}; \quad \boldsymbol{x}+(\boldsymbol{y}+\boldsymbol{z})=(\boldsymbol{x}+\boldsymbol{y})+\boldsymbol{z}; \quad \boldsymbol{x}+\mathbf{0}=\boldsymbol{x};$$

存在 $\boldsymbol{\xi} \in S$，使得 $\boldsymbol{\xi}+\boldsymbol{x}=\mathbf{0}$; 对任意实数 c,

$$c(\boldsymbol{x}+\boldsymbol{y})=c\boldsymbol{x}+c\boldsymbol{y}; \quad (c_1+c_2)\boldsymbol{x}=c_1\boldsymbol{x}+c_2\boldsymbol{x}; \quad c_1(c_2\boldsymbol{x})=(c_1c_2)\boldsymbol{x}; \quad 1\cdot\boldsymbol{x}=\boldsymbol{x}.$$

我们称 S 中的元素为向量.

例 1.1 我们通常遇到的 n 维欧氏空间就是线性空间，元素是起点为原点的行向量 $\boldsymbol{a}=(a_1,a_2,\cdots,a_n)$ 或列向量 $\boldsymbol{a}=(a_1,a_2,\cdots,a_n)'$，其中“$'$”表示转置.

例 1.2 设 $L_2 = \{$随机变量 $X : EX^2 < \infty\}$, 则容易验证 L_2 为线性空间.

定义 1.2 (线性相关和线性无关) 设 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m$ 是 S 中的一组向量, 如果存在 m 个不全为零的实数 $\lambda_1, \lambda_2, \cdots, \lambda_m$, 使得 $\lambda_1\boldsymbol{a}_1 + \lambda_2\boldsymbol{a}_2 + \cdots + \lambda_m\boldsymbol{a}_m = \boldsymbol{0}$, 则称这一组向量是线性相关的, 否则称为线性无关的.

在向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m$ 中, 一定存在若干个向量, 不妨记为 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_d$, 它们之间是线性无关的, 而 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m$ 中的其余向量都可以由 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_d$ 线性表出, 这组向量称为原向量组的一个极大线性无关向量组, 其中的向量个数 d 称为 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m$ 的秩. 当然, 极大线性无关向量组不是唯一的.

定义 1.3 (线性空间的基和维数) 设 S 是线性空间, 若存在一组向量 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$, 它们之间是线性无关的, 对 S 中的任一向量 $\boldsymbol{b}$, 存在 n 个实数 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 使得 $\boldsymbol{b} = \lambda_1\boldsymbol{a}_1 + \lambda_2\boldsymbol{a}_2 + \cdots + \lambda_n\boldsymbol{a}_n$, 即 $\boldsymbol{b}$ 可以表为 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$ 的线性组合, 则称向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$ 为线性空间 S 的一组基, n 称为空间的维数, 通常用 $d(S)$ 来表示空间 S 的维数.

由定义, 对任一 n 维线性空间, 一定存在一组基 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$, 基中向量的个数就是空间的维数 n. 如果把每个基看成一个坐标, 由于空间任一向量 $\boldsymbol{x}$ 可以表为一组基的线性组合 $\boldsymbol{x} = a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 + \cdots + a_n\boldsymbol{e}_n$, 我们可以把这组系数 $(a_1, a_2, \cdots, a_n)$ 看成向量 $\boldsymbol{x}$ 的一个表达. 一组基定下来以后, 这种表达是唯一的. 当然, 同一个向量在不同的基下有不同的表达.

我们常见的 n 维欧氏空间的一组基向量为 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$, 其中向量 $\boldsymbol{e}_i$ 的第 i 个分量为 1, 其余分量为零, 一般向量可以用 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$ 来表达.

注 1 基不唯一, 例如把上面 n 维欧氏空间的一组基改为 $\boldsymbol{f}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$, 其中向量 $\boldsymbol{f}_1$ 的第 1 和第 2 个分量为 1, 其余分量为零, 则 $\boldsymbol{f}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 仍是一组基, 上面的 $\boldsymbol{x}$ 在这组基下表为 $(a_1 - a_2, a_2, a_3, \cdots, a_n)$.

注 2 不是所有的线性空间都是有限维的, 如例 1.2 的线性空间 L_2 就不是有限维的.

定义 1.4 (子空间) 设 S 为线性空间, $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_m$ 为 S 中的 m 个向量, 记集合

$$\begin{aligned}\mu &= \mu(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_m)\\ &= \{\alpha_1\boldsymbol{x}_1 + \alpha_2\boldsymbol{x}_2 + \cdots + \alpha_m\boldsymbol{x}_m : \forall(\alpha_1, \alpha_2, \cdots, \alpha_m) \in \mathbb{R}^m\},\end{aligned}$$

容易验证集合 μ 是线性空间, 我们称它是由向量 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_m$ 生成的 S 的一个线性子空间, 简称子空间. 子空间的维数记为 $d(\mu)$, 显然, $d \leqslant n$.

试读结束：需要全本请在线购买： www.ertongbook.com

由线性代数知识可知, 子空间的维数 d 就是向量组 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_m$ 的秩.

定义 1.5 (内积空间) 设 $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in S$, t 为实数, 定义映射 $g : (\boldsymbol{x}, \boldsymbol{y}) \longrightarrow \mathbb{R}$, 满足

$$\begin{aligned}
(\boldsymbol{x}, \boldsymbol{y}) &= (\boldsymbol{y}, \boldsymbol{x}); \\
(\boldsymbol{x}, \boldsymbol{x}) &> 0, \quad \boldsymbol{x} \neq \mathbf{0}; \\
(\boldsymbol{x}, \boldsymbol{x}) &= 0 \Leftrightarrow \boldsymbol{x} = \mathbf{0}; \\
(t\boldsymbol{x}, \boldsymbol{y}) &= t(\boldsymbol{x}, \boldsymbol{y}); \\
(\boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z}) &= (\boldsymbol{x}, \boldsymbol{z}) + (\boldsymbol{y}, \boldsymbol{z}),
\end{aligned}$$

则二元实函数 (x, y) 称为线性空间 S 上的一个内积, 带内积的线性空间称为内积空间.

由内积我们可以定义向量的模 $\|\boldsymbol{x}\| = \sqrt{(\boldsymbol{x}, \boldsymbol{x})}$, 若 $\|\boldsymbol{x}\| = 1$, 我们称 $\boldsymbol{x}$ 为单位向量.

我们通常见到的 $\mathbb{R}^n$ 就是内积空间, 记 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)', \boldsymbol{y} = (y_1, y_2, \cdots, y_n)'$, 内积的定义为 $(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^n x_i y_i$.

例 1.3 在 L_2 中, 我们定义 $(X, Y) = EXY$, 可以验证这是一个内积, 因此 L_2 是一个内积空间.$\|X\| \equiv \sqrt{(X, X)}$ 称为随机变量 X 的模. 也可以把内积定义为 $(X, Y) = \mathrm{Cov}(X, Y) = E(X - EX)(Y - EY)$, 即随机变量 X 和 Y 的协方差, 这相当于把向量的起点都平移到原点.

在内积定义为协方差时, 若 $(\boldsymbol{X}, \boldsymbol{Y}) = 0$, 我们称 $\boldsymbol{X} \perp \boldsymbol{Y}$. 由内积定义, 我们知向量 $\boldsymbol{X}, \boldsymbol{Y}$ 的夹角余弦为

$$\cos\theta = \frac{(\boldsymbol{X}, \boldsymbol{Y})}{\|\boldsymbol{X}\|\|\boldsymbol{Y}\|}.$$

在 L_2 中, 两个随机向量之间的夹角余弦就是相关系数

$$\rho_{XY} = \frac{\mathrm{Cov}(\boldsymbol{X}, \boldsymbol{Y})}{\sqrt{\mathrm{Var}(\boldsymbol{X})\mathrm{Var}(\boldsymbol{Y})}},$$

上面也就是相关系数的几何意义.

§1.1.2 Gram-Schmidt 正交化程序

设 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 是 n 维内积空间 S 的一组基, 如果 $\boldsymbol{e}_i \perp \boldsymbol{e}_j, \forall i \neq j$, 且对一切 $i = 1, 2, \cdots, n$, $\|\boldsymbol{e}_i\| = 1$, 我们称 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 是 S 的一组标准正交基.

对 n 维内积空间 S 的任一组基 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$, 我们可以构造出一组标准正交基 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$. 这就是如下的 Gram-Schmidt 正交化程序:

$$
\begin{aligned}
&\widetilde{\boldsymbol{e}}_1=\boldsymbol{x}_1,\ \boldsymbol{e}_1=\frac{\boldsymbol{x}_1}{\|\boldsymbol{x}_1\|};\\
&\widetilde{\boldsymbol{e}}_2=\boldsymbol{x}_2-(\boldsymbol{x}_2,\boldsymbol{e}_1)\boldsymbol{e}_1,\quad \boldsymbol{e}_2=\frac{\widetilde{\boldsymbol{e}}_2}{\|\widetilde{\boldsymbol{e}}_2\|};\\
&\widetilde{\boldsymbol{e}}_3=\boldsymbol{x}_3-(\boldsymbol{x}_3,\boldsymbol{e}_1)\boldsymbol{e}_1-(\boldsymbol{x}_3,\boldsymbol{e}_2)\boldsymbol{e}_2,\ \boldsymbol{e}_3=\frac{\widetilde{\boldsymbol{e}}_3}{\|\widetilde{\boldsymbol{e}}_3\|};\\
&\cdots\cdots\cdots\cdots\\
&\widetilde{\boldsymbol{e}}_n=x_n-(\boldsymbol{x}_n,\boldsymbol{e}_1)\boldsymbol{e}_1-(\boldsymbol{x}_n,\boldsymbol{e}_2)\boldsymbol{e}_2-\cdots-(\boldsymbol{x}_n,\boldsymbol{e}_{n-1})\boldsymbol{e}_{n-1},\ \boldsymbol{e}_n=\frac{\widetilde{\boldsymbol{e}}_n}{\|\widetilde{\boldsymbol{e}}_n\|}.
\end{aligned}
$$

由此可知, 每个 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 都可以用基向量 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 的线性组合表达出来.

§1.1.3 向量的正交投影和 Bessel 不等式

一、向量的正交投影

我们需要下面的正交分解定理.

定理 1.1 (正交分解定理) *设 $W\subset S$ 是 S 的一个 r 维子空间, $\boldsymbol{x}\notin W$, 则存在向量 $\boldsymbol{y},\boldsymbol{z}$, 满足 $\boldsymbol{z}\in W$, $\boldsymbol{y}\in S$, $\boldsymbol{y}\neq\mathbf{0}$, 且 $\boldsymbol{y}\perp W$, 使得 $\boldsymbol{x}=\boldsymbol{y}+\boldsymbol{z}$, 且该分解是唯一的.*

证 设 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_r$ 为 W 的一组基, 由于 $\boldsymbol{x}$ 不在空间 W 中, 故 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_r,\boldsymbol{x}$ 是线性无关的. 由 Gram-Schmidt 正交化程序, 由 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_r,\boldsymbol{x}$ 可以构造一组正交基 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_r,\boldsymbol{e}_{r+1}$, 使得

$$\boldsymbol{x}=\alpha_{r+1}\boldsymbol{e}_{r+1}+(\alpha_1\boldsymbol{e}_1+\cdots+\alpha_r\boldsymbol{e}_r),$$

而

$$\boldsymbol{z}=\alpha_1\boldsymbol{e}_1+\cdots+\alpha_r\boldsymbol{e}_r\in W,\quad \boldsymbol{y}=\alpha_{r+1}\boldsymbol{e}_{r+1}\perp W,$$

$\boldsymbol{z}=\alpha_1\boldsymbol{e}_1+\cdots+\alpha_r\boldsymbol{e}_r$ 称为向量 $\boldsymbol{x}$ 在子空间 W 上的投影. 投影在统计中常常是一种估计量, 而向量 $\boldsymbol{y}$ 常常称为残差. 它有如下重要性质:

$$\|\boldsymbol{y}\|=\inf_{\boldsymbol{u}\in W}\|\boldsymbol{x}-\boldsymbol{u}\|.$$

注意到, 若 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_r$ 是 W 的一组标准正交基, 内积 $(\boldsymbol{x},\boldsymbol{e}_i)=\alpha_i$ 可视为向量 $\boldsymbol{x}$ 在 $\boldsymbol{e}_i$ 上的投影. 设 $\boldsymbol{x}$ 有正交投影分解 $\boldsymbol{x}=\boldsymbol{y}+\boldsymbol{z}$, $\boldsymbol{y}\perp W$, $\boldsymbol{z}\in W$, 则

$\forall \boldsymbol{u} \in W, (\boldsymbol{u}-\boldsymbol{z}, \boldsymbol{y})=0$, 故

$$
\begin{aligned}
\|\boldsymbol{x}-\boldsymbol{u}\|^2 &= (\boldsymbol{x}-\boldsymbol{u}, \boldsymbol{x}-\boldsymbol{u}) = (\boldsymbol{y}+\boldsymbol{z}-\boldsymbol{u}, \boldsymbol{y}+\boldsymbol{z}-\boldsymbol{u}) \\
&= (\boldsymbol{y}, \boldsymbol{y}) + 2(\boldsymbol{y}, \boldsymbol{z}-\boldsymbol{u}) + (\boldsymbol{z}-\boldsymbol{u}, \boldsymbol{z}-\boldsymbol{u}) \\
&= (\boldsymbol{y}, \boldsymbol{y}) + \|\boldsymbol{z}-\boldsymbol{u}\|^2 \geqslant \|\boldsymbol{y}\|^2,
\end{aligned}
$$

等号成立当且仅当 $\boldsymbol{u}=\boldsymbol{z}$, 这说明向量 $\boldsymbol{x}$ 在子空间 W 上的投影是使残差最小的 W 中的向量. 这个几何直观对理解统计中常用统计量是非常有帮助的.

二、Bessel 不等式

设 $\boldsymbol{f}_1, \boldsymbol{f}_2, \cdots, \boldsymbol{f}_n$ 为内积空间 S (不必是有限维的) 中的一组标准正交向量, $\boldsymbol{\beta} \in S$, 则

$$
|(\boldsymbol{f}_1, \boldsymbol{\beta})|^2 + \cdots + |(\boldsymbol{f}_n, \boldsymbol{\beta})|^2 \leqslant \|\boldsymbol{\beta}\|^2,
$$

等号成立 $\Leftrightarrow \boldsymbol{\beta} \in \mu(\boldsymbol{f}_1, \cdots, \boldsymbol{f}_n)$.

上面不等式的几何意义是向量长度的平方 (相当于三维空间立方体中对角线的平方) 不小于任意直角边的平方和.

证 设 $\boldsymbol{\delta} = \sum_{i=1}^n (\boldsymbol{f}_i, \boldsymbol{\beta})\boldsymbol{f}_i$, 则 $\boldsymbol{\delta}$ 是 $\boldsymbol{\beta}$ 在 $\mu(\boldsymbol{f}_1, \cdots, \boldsymbol{f}_n)$ 上的投影. 由向量的正交分解, $\boldsymbol{\beta} = \boldsymbol{\delta} + \boldsymbol{r}$, $(\boldsymbol{\delta}, \boldsymbol{r}) = 0$, 因此

$$
\|\boldsymbol{\beta}\|^2 = \|\boldsymbol{\delta}\|^2 + \|\boldsymbol{r}\|^2 \geqslant \|\boldsymbol{\delta}\|^2 = \sum_{i=1}^n |(\boldsymbol{f}_i, \boldsymbol{\beta})|^2.
$$

注 这里空间 S 不必是有限维的, 例如由 $(-\pi, \pi)$ 上的 Fourier 正交基 $(1, \sin nx, \cos nx)$ $n=1,2,\cdots$ 表达的 Hilbert 空间是无穷维的.

由上面讨论知, 若 S_1 为 S 的一个子空间, 设 $S_2 = \{\boldsymbol{x}: \boldsymbol{x} \perp S_1, \boldsymbol{x} \in S\}$, 则我们称 $S_2 \perp S_1$, 或称为 S_1 的正交补, 记为 $S_1^{\perp}$, 此时, 空间可以记为 $S = S_1 \oplus S_1^{\perp}$, 称为子空间 S_1 与 $S_1^{\perp}$ 的直和. 子空间维数与空间 S 维数的关系是 $d(S) = d(S_1) + d(S_1^{\perp})$.

我们也容易把直和的概念推广到 k 个子空间的直和, 这里不再细述了.

§1.2 矩阵的一般理论和性质

一、矩阵的运算

把 m 个 n 维的行向量 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m$ 叠起来, 就构成一个 $m \times n$ 的矩阵. 一般记为 $\boldsymbol{A} = (a_{ij})_{m \times n}$. 我们也可以把 $m \times n$ 的矩阵看成 n 个 m 维的列向量

$\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_n$ 并列在一起. 如果 $m = n$, 称其为方阵, 方阵中的每个 a_{ii} 称为对角元. 如果方阵中非对角元都是零, 称其为对角阵, 常记为 $\boldsymbol{A} = \mathrm{diag}(a_{11}, \cdots, a_{nn})$. 对矩阵, 我们可以定义加法、数乘和矩阵乘法:

加法 (都是 $m \times n$ 的矩阵):

$$\boldsymbol{A} + \boldsymbol{B} = (a_{ij} + b_{ij})_{m\times n},$$

它有如下性质:

$$\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{B} + \boldsymbol{A}, \quad \boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C}) = (\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C}.$$

数乘: 设 c 为任一常数, 定义 $c\boldsymbol{A} = (ca_{ij})_{m\times n}$, 数乘满足分配律和结合律:

$$(c + d)\boldsymbol{A} = c\boldsymbol{A} + d\boldsymbol{A}, \quad c(\boldsymbol{A} + \boldsymbol{B}) = c\boldsymbol{A} + c\boldsymbol{B}.$$

矩阵乘法: 设 $\boldsymbol{A} = (a_{ij})_{p\times q}, \boldsymbol{B} = (b_{ij})_{q\times r}$, 定义 $\boldsymbol{C} = \boldsymbol{AB} = (c_{ij})_{p\times r}$, 其中

$$c_{ij} = \sum_{k=1}^{q} a_{ik}b_{kj}.$$

矩阵乘法满足结合律和分配律:

$$\begin{aligned}
&\boldsymbol{A}(\boldsymbol{BC}) = (\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{ABC}, \\
&\boldsymbol{A}(\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{AB} + \boldsymbol{AC}, \quad (\boldsymbol{A} + \boldsymbol{B})(\boldsymbol{C} + \boldsymbol{D}) = \boldsymbol{A}(\boldsymbol{C} + \boldsymbol{D}) + \boldsymbol{B}(\boldsymbol{C} + \boldsymbol{D}).
\end{aligned}$$

交换律一般不成立. 如果矩阵中的每个元素都是零, 称其为零矩阵, 记为 $\mathbf{0} = (0)_{m\times n}$, 如果对角阵中每个对角元都是 1, 称其为单位阵, 记为 $\boldsymbol{I}$. 零矩阵和单位阵有如下显然的性质 (只要能进行运算):

$$\begin{aligned}
\boldsymbol{A} + \mathbf{0} &= \boldsymbol{A}, \\
\boldsymbol{IA} = \boldsymbol{AI} &= \boldsymbol{A}.
\end{aligned}$$

矩阵的转置: 设 $\boldsymbol{A} = (a_{ij})_{m\times n}$, 定义 $b_{ij} = a_{ji}, \boldsymbol{B} = (b_{ij})_{n\times m}$, 我们称矩阵 $\boldsymbol{B}$ 为矩阵 $\boldsymbol{A}$ 的转置, 记为 $\boldsymbol{A}^T$ 或 $\boldsymbol{A}'$, 矩阵转置满足

$$(\boldsymbol{AB})' = \boldsymbol{B}'\boldsymbol{A}', \quad (\boldsymbol{ABC})' = \boldsymbol{C}'\boldsymbol{B}'\boldsymbol{A}'.$$

若实方阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}' = \boldsymbol{A}$, 我们称方阵 $\boldsymbol{A}$ 是对称阵.

共轭转置: 设矩阵 $\boldsymbol{A} = (a_{ij})_{p\times q}$ 的元素 a_{ij} 是复数, 定义 $\boldsymbol{A}^* = \bar{\boldsymbol{A}}' = (\bar{a}_{ji})_{q\times p}$, 其中 $\bar{a}_{ji}$ 表示元素 a_{ji} 的共轭. 共轭转置这一运算满足 $(\boldsymbol{AB})^* = \boldsymbol{B}^*\boldsymbol{A}^*$, $(\boldsymbol{ABC})^* = \boldsymbol{C}^*\boldsymbol{B}^*\boldsymbol{A}^*$.

若复方阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^*=\boldsymbol{A}$, 我们称方阵 $\boldsymbol{A}$ 是 Hermit 阵.

逆矩阵: 设 $\boldsymbol{A}$ 为方阵, 若存在方阵 $\boldsymbol{B}$, 使得 $\boldsymbol{AB}=\boldsymbol{BA}=\boldsymbol{I}$, 则方阵 $\boldsymbol{B}$ 称为方阵 $\boldsymbol{A}$ 的逆, 记为 $\boldsymbol{A}^{-1}$. 逆运算有如下性质:

$$\begin{aligned}
&(\boldsymbol{AB})^{-1}=\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}, \quad (\boldsymbol{ABC})^{-1}=\boldsymbol{C}^{-1}\boldsymbol{B}^{-1}\boldsymbol{A}^{-1},\\
&(\boldsymbol{A}')^{-1}=(\boldsymbol{A}^{-1})', \quad (\boldsymbol{A}^*)^{-1}=(\boldsymbol{A}^{-1})^*.
\end{aligned}$$

注 $\boldsymbol{AB}=\boldsymbol{0} \nRightarrow \boldsymbol{A}=\boldsymbol{0}$ 或 $\boldsymbol{B}=\boldsymbol{0}$.

U 阵 (unitary matrix) 和正交阵 (orthogonal matrix): 一个复元素的方阵 $\boldsymbol{A}$ 称为 U 阵, 若 $\boldsymbol{A}^*\boldsymbol{A}=\boldsymbol{AA}^*=\boldsymbol{I}$, 此时, $\boldsymbol{A}^*=\boldsymbol{A}^{-1}$; 当 $\boldsymbol{A}$ 的元素为实数时, 称为正交阵, 此时 $\boldsymbol{A}^*=\boldsymbol{A}'$, 故正交阵满足 $\boldsymbol{A}'\boldsymbol{A}=\boldsymbol{AA}'=\boldsymbol{I}$.

分块矩阵: 若矩阵 $\boldsymbol{A}$ 可以分为 $r\times s$ 个矩形小块, 如

$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{P} & \boldsymbol{Q}\\ \boldsymbol{R} & \boldsymbol{S}\end{pmatrix},$$

则 $\boldsymbol{A}$ 称为一个分块矩阵. 设分块矩阵 $\boldsymbol{B}$ 为

$$\boldsymbol{B}=\begin{pmatrix}\boldsymbol{E} & \boldsymbol{F}\\ \boldsymbol{G} & \boldsymbol{H}\end{pmatrix},$$

只要运算能进行, 我们有分块矩阵的乘法

$$\begin{aligned}
\boldsymbol{AB}&=\begin{pmatrix}\boldsymbol{P} & \boldsymbol{Q}\\ \boldsymbol{R} & \boldsymbol{S}\end{pmatrix}\begin{pmatrix}\boldsymbol{E} & \boldsymbol{F}\\ \boldsymbol{G} & \boldsymbol{H}\end{pmatrix}\\
&=\begin{pmatrix}\boldsymbol{PE}+\boldsymbol{QG} & \boldsymbol{PF}+\boldsymbol{QH}\\ \boldsymbol{RE}+\boldsymbol{SG} & \boldsymbol{RF}+\boldsymbol{SH}\end{pmatrix}.
\end{aligned}$$

分块矩阵求逆: 设分块矩阵

$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22}\end{pmatrix},$$

若 $\boldsymbol{A}$ 可逆, 且 $\boldsymbol{A}_{11},\boldsymbol{A}_{22}$ 都可逆, 记

$$\begin{aligned}
\boldsymbol{A}_{11\cdot 2}&=\boldsymbol{A}_{11}-\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21},\\
\boldsymbol{A}_{22\cdot 1}&=\boldsymbol{A}_{22}-\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12},
\end{aligned}$$

则

$$\boldsymbol{A}^{-1}=\begin{pmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22}\end{pmatrix}^{-1}=\begin{pmatrix}\boldsymbol{A}_{11\cdot 2}^{-1} & -\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22\cdot 1}^{-1}\\ -\boldsymbol{A}_{22\cdot 1}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1} & \boldsymbol{A}_{22\cdot 1}^{-1}\end{pmatrix},$$

其中

$$A_{11\cdot 2}^{-1} = (A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1} = A_{11}^{-1} + A_{11}^{-1}A_{12}A_{22\cdot 1}^{-1}A_{21}A_{11}^{-1},$$
$$A_{11}^{-1}A_{12}A_{22\cdot 1}^{-1} = A_{11\cdot 2}^{-1}A_{12}A_{22}^{-1},$$
$$A_{22\cdot 1}^{-1}A_{21}A_{11}^{-1} = A_{22}^{-1}A_{21}A_{11\cdot 2}^{-1},$$

且

$$A_{22\cdot 1}^{-1} = (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} = A_{22}^{-1} + A_{22}^{-1}A_{21}A_{11\cdot 2}^{-1}A_{12}A_{22}^{-1}.$$

例 1.4　在分块矩阵求逆公式中取 $\boldsymbol{A}_{12} = \boldsymbol{a}, \boldsymbol{A}_{22} = -1, \boldsymbol{A}_{21} = \boldsymbol{b}'$, 其中 $\boldsymbol{a}, \boldsymbol{b}$ 为列向量, 则我们有

$$(\boldsymbol{A}_{11} + \boldsymbol{a}\boldsymbol{b}')^{-1} = \boldsymbol{A}_{11}^{-1} - \boldsymbol{A}_{11}^{-1}\boldsymbol{a}\boldsymbol{b}'\boldsymbol{A}_{11}^{-1}/(1 + \boldsymbol{b}'\boldsymbol{A}_{11}^{-1}\boldsymbol{a}).$$

二、矩阵的相抵

对矩阵 $\boldsymbol{A}_{m\times n}$, 我们可以作如下的初等变换:

某行乘以一个数 λ. 取对角阵 $\boldsymbol{M}_i(\lambda) = \mathrm{diag}(1, \cdots, 1, \lambda, 1, \cdots, 1)$, 即第 i 个对角元为 λ, 其余为 1, 则矩阵 $\boldsymbol{M}_i(\lambda)$ 左乘 A 就完成此功能.

把 s 行的 λ 倍加到第 r 行. 令 $\boldsymbol{P}_{rs}(\lambda) = (p_{ij})_{m\times m}$, 其中 $p_{ii} = 1, i = 1, \cdots, m$, $p_{rs} = \lambda$, 其余元都为零, 则矩阵 $\boldsymbol{P}_{rs}(\lambda)$ 左乘 A 就完成此功能.

交换 r, s 行. 令 $m \times m$ 矩阵 $\boldsymbol{T}_{rs} = (t_{ij})$, 其中 $t_{rs} = t_{sr} = 1, t_{rr} = t_{ss} = 0$, 矩阵的其余元为零, 则矩阵 $\boldsymbol{T}_{rs}$ 左乘 $\boldsymbol{A}$ 就完成此功能.

注 1　我们也可以对矩阵中的列作上述三种变换, 此时只要把上述的三种初等变换矩阵右乘矩阵 $\boldsymbol{A}$ 就完成此功能.

注 2　可以验证 $(\boldsymbol{M}_i(\lambda))^{-1} = \boldsymbol{M}_i(\lambda^{-1}), (\boldsymbol{P}_{rs}(\lambda))^{-1} = \boldsymbol{P}_{rs}(-\lambda), \boldsymbol{T}_{rs}^{-1} = \boldsymbol{T}_{rs}$.

由上面对行和列的三种初等变换, 在线性代数中我们可以得到矩阵的相抵标准型: 对任意的矩阵 $\boldsymbol{A}_{m\times n}$, 存在 $m \times m$ 可逆方阵 $\boldsymbol{P}$ 和 $n \times n$ 可逆方阵 $\boldsymbol{Q}$, 使得

$$\boldsymbol{PAQ} = \begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix},$$

其中 r 称为矩阵的秩 (rank), 记为 $\mathrm{rank}(A)$, 它可以看作矩阵 $\boldsymbol{A}$ 的 n 个列向量中极大线性无关组中向量的个数, 也可以看作矩阵 $\boldsymbol{A}$ 的 m 个行向量中极大线性无关组中向量的个数.

试读结束：需要全本请在线购买： www.ertongbook.com