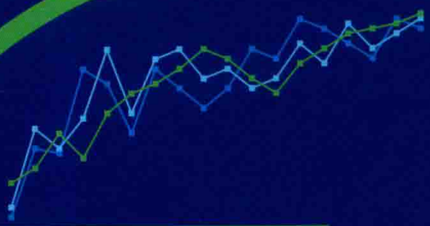

回归分析与线性统计模型

林建忠◎编

上海交通大学出版社
SHANGHAI JIAO TONG UNIVERSITY PRESS

回归分析与线性统计模型

林建忠◎编

上海交通大学出版社
SHANGHAI JIAO TONG UNIVERSITY PRESS

内容提要

本书介绍了几种典型的线性统计模型及其建模分析方法,不仅详细讲解了各种理论公式的推导过程,还就具体的案例数据结合统计软件展示数据分析的各个步骤.此外,每章还配备一定数量的理论习题与上机实验题.

本书可作为普通高等院校应用统计硕士专业学位研究生基础课程教材,也可作为数学专业大四学生和其他学科研究生统计课程的教学参考书,以及业界数据分析师的参考用书.

图书在版编目(CIP)数据

回归分析与线性统计模型/ 林建忠编.—上海:上海交通大学出版社,2018
ISBN 978-7-313-20166-9

Ⅰ.①回… Ⅱ.①林… Ⅲ.①回归分析—高等学校—教材②线性回归—统计模型—高等学校—教材 Ⅳ.①O212.1

中国版本图书馆 CIP 数据核字(2018)第 209552 号

回归分析与线性统计模型

主　　编:林建忠
出版发行:上海交通大学出版社
地　　址:上海市番禺路 951 号
邮政编码:200030
电　　话:021-64071208
出 版 人:谈　毅
印　　制:上海景条印刷有限公司
经　　销:全国新华书店
开　　本:710 mm×1000 mm　1/16
印　　张:21.5
字　　数:392 千字
版　　次:2018 年 9 月第 1 版
印　　次:2018 年 9 月第 1 次印刷
书　　号:ISBN 978-7-313-20166-9/O
定　　价:58.00 元

前言

在大数据时代，探索由工作进程所产生的数据集中各变量之间的关系已成为十分重要的问题.大数据时代的数据集不仅变量繁多，而且变量性状多样化及变量间层次结构复杂化.数据分析的实践表明，拟合这类数据集中多变量复杂关系的模型往往首先考虑的是线性统计模型.广义上说，线性统计模型包括线性回归模型、可以线性化的非线性模型、方差分析模型、广义线性模型和混合效应类线性模型等一大类应用广泛的统计模型.对实际数据建立统计模型过程中所使用的两个基本建模分析方法包括回归分析方法和方差分析方法.

回归分析法是确定两种或两种以上变量间相互依赖的定量关系的一种统计分析方法.由于实际数据集中的变量与变量之间的定量关系事先是未知的，回归分析方法强调从多个视角评判数据所拟合的统计模型的逻辑合理性，运用十分广泛.线性统计模型的理论及其相应的分析方法也是进一步学习和研究其他统计建模方法的基础.正是由于这些原因，线性统计模型及其建模分析方法不仅成为统计专业本科生和研究生的基础课，而且也是生物、医学、经济、管理、商业、金融、工程技术及社会科学等学科的本科生和研究生统计课程的重要内容.

本书以回归分析方法作为线性回归模型建模分析的主线，然后将这一方法逐步引入到方差分析模型、协方差分析

模型和混合效应类线性模型的建模分析过程中.在这个过程中,方差分析方法是自然代入的.全书共12章.第1章预备知识汇总了本书所需的矩阵和随机向量知识;第2、3章系统讨论了多元线性回归模型的参数估计、假设检验和模型设定的恰当性检验;第4、5、6章介绍了模型设定不恰当及病态数据情形下的几种模型校正方法;第7章就实际数据展示了如何搭建和验证各种可能扩展的线性模型;第8章讲解了能够处理具有一般随机特征响应变量的广义线性模型;第9、10章系统探讨了能够分析因变量与定性型自变量线性关系的方差分析模型和协方差分析模型,第11、12章介绍了能够分析具有随机效应和分层结构数据特征的几种混合效应类线性模型及面板数据模型.

线性统计模型的类型十分丰富,本书旨在为应用统计硕士专业学位的研究生介绍这一领域最常用的几类模型,使读者对线性统计模型的建模分析方法有比较系统的了解,以便为今后更深入的学习打下良好的基础.正因为此,读者只要掌握了大学阶段的多元微积分、线性代数和初等概率统计知识,就可以阅读本书的内容.考虑到目前统计软件及介绍统计软件的书籍和网上资料繁多,因此本书对统计软件的具体操作不再详细介绍.本书的部分内容也可作为数学专业大四学生和其他学科研究生统计课程的教学参考书,以及业界数据分析师的参考用书.

在本书完稿之际,我要感谢所有关心和支持我写作和出版此书的人们.香港中文大学统计系张绍洪教授带我进入线性模型统计建模领域.自2005年以来,本教材的大部分内容在上海交通大学数学科学学院的《数据分析》和《回归分析与线性统计模型》课程中进行讲授,部分参加课程学习的同学也对课程的成功和本书的成书做出了贡献,其中张铁和杨宇弘两同学为本书各提供了一个案例.本书的写作和

出版得到了教育部上海交通大学2016年深化专业学位研究生教育综合改革项目(项目编号:ZXZY071007)的重点资助,以及国家自然科学基金国际(地区)合作与交流项目“开放网络下医疗资源配置和优化的模型、算法及应用研究”(项目编号:71520107003)和上海交通大学2017年研究生“学科人才培养建设与创新-专业课建设”经费(经费号:WF610107101)的资助.

本书相关数据存放在本书作者在上海交通大学数学科学学院的个人网页内.

由于作者水平所限,书中存在的缺点和错误,恳请同行和广大读者批评指正.

林建忠

上海交通大学数学科学学院

2018年2月

目 录

1 预备知识

线性统计模型中的回归分析和方差分析等方法广泛使用了矩阵和随机向量的数学理论，国内外各类线性统计模型的著作都对这些结果进行了不同程度地介绍，如 *Matrix Analysis for Statistics*[1]《近代回归分析》[2]《线性模型引论》[3]《统计学中的矩阵代数》[4]等.本章仅就与本书有关的相关结果加以介绍.

1.1 实向量线性空间

一般的线性空间 S 是由向量组成的一个集合，集合中的元素是一般意义的向量，它可以是 n 维实数向量，也可以是满足一定条件的函数等.在线性空间 S 这个集合中定义了两种运算，即向量与向量之间的加法运算和向量与实数之间的乘法运算.此外，要求 S 中任意两个向量求和之后所得的向量仍然是 S 中的向量，即线性空间对向量加法运算封闭；S 中任意向量与任意实数相乘后仍然是 S 中的向量，即线性空间对向量与实数之间的数乘运算封闭.另外，向量之间的加法运算满足结合律和交换律，数乘运算满足结合律和分配律.本书仅简单介绍由 n 维实数向量按通常意义下的向量加法和数乘法所形成的线性空间 $\mathbf{R}^n$，简称 n 维实向量空间，直观上可以将它看成是二、三维实向量空间的自然推广.

考虑 n 维实数空间 $\mathbf{R}^n$ 中向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_k$ 的一切可能的线性组合构成的集合

$$S_0 = \left\{\boldsymbol{x} = \sum_{i=1}^{k} \alpha_i \boldsymbol{a}_i, \alpha_1, \alpha_2, \cdots, \alpha_k \in \mathbf{R}\right\},$$

S_0 也是线性空间，称为 $\mathbf{R}^n$ 的子空间.若记 $\mathbf{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_k)$，则 $S_0 = \{\boldsymbol{x} = \mathbf{A}t, t \in \mathbf{R}_k\}$，它是矩阵 $\mathbf{A}$ 的列向量张成的子空间，记为 $S_0 = \mathcal{M}(A)$. 如果子空间 $\boldsymbol{S}_0$ 由一组线性无关的向量 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_k$ 张成，则称 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_k$ 为 S_0 的一组基，k 称为 S_0 的维数，记作 $k = \dim(S_0)$. 记矩阵 $\mathbf{A}$ 的秩为 $\mathrm{rk}(\mathbf{A})$，容易证明 $\dim \mathcal{M}(\mathbf{A}) = \mathrm{rk}(\mathbf{A})$.

对 $\mathbf{R}^n$ 中的任意两个向量 $\boldsymbol{a}=(a_1, a_2, \cdots, a_n)^{\mathrm{T}}$，$\boldsymbol{b}=(b_1, b_2, \cdots, b_n)^{\mathrm{T}}$，定义它们的内积为 $(\boldsymbol{a}, \boldsymbol{b})=\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}=\sum_{i=1}^{n}a_i b_i$.若 $(\boldsymbol{a}, \boldsymbol{b})=0$，则称 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 正交，记为 $\boldsymbol{a}\perp\boldsymbol{b}$.若 $\boldsymbol{a}$ 与子空间 S 中的每一个向量正交，则称 $\boldsymbol{a}$ 正交于 S，记为 $\boldsymbol{a}\perp S$.称 $(\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a})^{1/2}=\sqrt{\sum_{i=1}^{n}a_i^2}$ 为向量 $\boldsymbol{a}$ 的长度，记为 $\|\boldsymbol{a}\|$.设 S 为一子空间，容易证明

$$S^{\perp}=\{\boldsymbol{x}:\boldsymbol{x}\perp S\}$$

也是线性空间，称为 S 的正交补空间.设 $\mathbf{A}$ 为 $n\times k$ 矩阵，记 $\mathbf{A}^{\perp}$ 为满足条件 $\mathbf{A}^{\mathrm{T}}\mathbf{A}^{\perp}$ 且具有最大秩的矩阵，则

$$\mathcal{M}(\mathbf{A}^{\perp})=\mathbf{A}^{\perp}. \tag{1-1}$$

对于一个线性空间 S，如果存在 k 个子空间 $S_1, S_2, \cdots, S_k$，使得对任意 $\boldsymbol{a}\in S$，可唯一分解为

$$\boldsymbol{a}=\boldsymbol{a}_1+\boldsymbol{a}_2+\cdots+\boldsymbol{a}_k,\ \boldsymbol{a}_i\in S_i(i=1, 2, \cdots, k),$$

则称 S 为 $S_1, S_2, \cdots, S_k$ 的直和，记为 $S=S_1\oplus S_2\oplus\cdots\oplus S_k$.若进一步假设，对任意的 $\boldsymbol{a}_i\in S_i$，$\boldsymbol{a}_j\in S_j$，$i\neq j$ 有 $\boldsymbol{a}_i\perp\boldsymbol{a}_j$，则称 S 为 $S_1, S_2, \cdots, S_k$ 的正交直和，记为 $S=S_1\dot{+}S_2\dot{+}\cdots\dot{+}S_k$，特别 $\mathbf{R}^n=S\dot{+}S^{\perp}$，对 $\mathbf{R}^n$ 的任意子空间 S 成立.设 $\mathbf{A}=(\mathbf{A}_1\vdots\cdots\vdots\mathbf{A}_k)$，$\mathcal{M}(\mathbf{A}_i)\mathcal{M}(\mathbf{A}_j)=\{\mathbf{0}\}$，$i\neq j$，则

$$\mathcal{M}(\mathbf{A})=\mathcal{M}(\mathbf{A}_1)\oplus\mathcal{M}(\mathbf{A}_2)\oplus\cdots\oplus\mathcal{M}(\mathbf{A}_k).$$

若进一步假设 $\mathbf{A}_i^{\mathrm{T}}\mathbf{A}_j=\mathbf{0}$，$i\neq j$，则

$$\mathcal{M}(\mathbf{A})=\mathcal{M}(\mathbf{A}_1)\dot{+}\mathcal{M}(\mathbf{A}_2)\dot{+}\cdots\dot{+}\mathcal{M}(\mathbf{A}_k),$$

对实向量线性空间的完整介绍可以参考其他相关书籍[1,5].

以下介绍几个以后讨论中经常用到的事实.

定理 1.1 对任意矩阵 $\mathbf{A}$，恒有 $\mathcal{M}(\mathbf{A})=\mathcal{M}(\mathbf{A}\mathbf{A}^{\mathrm{T}})$.

证明 显然 $\mathcal{M}(\mathbf{A}\mathbf{A}^{\mathrm{T}})\subset\mathcal{M}(\mathbf{A})$，故只需证 $\mathcal{M}(\mathbf{A})\subset\mathcal{M}(\mathbf{A}\mathbf{A}^{\mathrm{T}})$.事实上，对任给 $\boldsymbol{x}\perp\mathcal{M}(\mathbf{A}\mathbf{A}^{\mathrm{T}})$，有 $\boldsymbol{x}^{\mathrm{T}}\mathbf{A}\mathbf{A}^{\mathrm{T}}=\mathbf{0}$.右乘 $\boldsymbol{x}$，得 $\boldsymbol{x}^{\mathrm{T}}\mathbf{A}\mathbf{A}^{\mathrm{T}}\boldsymbol{x}=\|\mathbf{A}^{\mathrm{T}}\boldsymbol{x}\|^2=0$，故 $\mathbf{A}^{\mathrm{T}}\boldsymbol{x}=\mathbf{0}$.于是 $\boldsymbol{x}\perp\mathcal{M}(\mathbf{A})$.定理证毕.

定理 1.2 设 $\mathbf{A}_{n\times m}$，$\boldsymbol{H}_{k\times m}$，则

(1) $S=\{\mathbf{A}\boldsymbol{x}:\boldsymbol{H}\boldsymbol{x}=\mathbf{0}\}$ 是 $\mathcal{M}(\mathbf{A})$ 的子空间.

(2) $\dim(S)=\mathrm{rk}\begin{bmatrix}\mathbf{A}\\ \boldsymbol{H}\end{bmatrix}-\mathrm{rk}(\boldsymbol{H})$.

证明 (1) 显然成立,证明略;(2) 不妨设 $\mathrm{rk}(\boldsymbol{H})=k$(自然数 $k \leqslant m$),则存在 $m \times m$ 可逆阵 $\boldsymbol{Q}$,使得 $\boldsymbol{HQ}=(\boldsymbol{I}_k \vdots \boldsymbol{O})$,其中 $\boldsymbol{I}_k$ 为 k 阶单位矩阵.于是

$$
\begin{aligned}
\dim(S) &= \dim\left\{\begin{bmatrix}\boldsymbol{A}\\ \boldsymbol{H}\end{bmatrix}\boldsymbol{x} : \boldsymbol{Hx}=\boldsymbol{0}\right\} = \dim\left\{\begin{bmatrix}\boldsymbol{A}\\ \boldsymbol{H}\end{bmatrix}\boldsymbol{Qx} : \boldsymbol{HQx}=\boldsymbol{0}\right\} \\
&= \dim\left\{\begin{bmatrix}\boldsymbol{U}_1 & \boldsymbol{U}_2\\ \boldsymbol{I}_k & \boldsymbol{O}\end{bmatrix}\boldsymbol{x} : (\boldsymbol{I}_k \vdots \boldsymbol{O})\boldsymbol{x}=\boldsymbol{0}\right\} = \dim\{\boldsymbol{U}_2\boldsymbol{x}_2 : \boldsymbol{x}_2 \text{ 任意}\} \\
&= \mathrm{rk}(\boldsymbol{U}_2) = \mathrm{rk}\begin{bmatrix}\boldsymbol{U}_1 & \boldsymbol{U}_2\\ \boldsymbol{I}_k & \boldsymbol{O}\end{bmatrix} - \mathrm{rk}(\boldsymbol{I}_k) = \mathrm{rk}\begin{bmatrix}\boldsymbol{A}\\ \boldsymbol{H}\end{bmatrix} - \mathrm{rk}(\boldsymbol{H}),
\end{aligned}
$$

其中 $(\boldsymbol{U}_1 \vdots \boldsymbol{U}_2)=\boldsymbol{AQ}$, $\boldsymbol{x}=\begin{bmatrix}\boldsymbol{x}_1\\ \boldsymbol{x}_2\end{bmatrix}$, $\boldsymbol{x}_1$ 为 $k \times 1$ 向量,$\boldsymbol{x}_2$ 为 $(m-k)\times 1$ 向量,定理证毕.

推论 1.1 设 $\mathcal{M}(\boldsymbol{A}) \cap \mathcal{M}(\boldsymbol{B})=\{\boldsymbol{0}\}$,则 $\mathcal{M}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{B}^{\perp})=\mathcal{M}(\boldsymbol{A}^{\mathrm{T}})$.

证明 因为

$$
\mathcal{M}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{B}^{\perp}) = \{\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x},\ \boldsymbol{x}=\boldsymbol{B}^{\perp}t,\ t \text{ 为任意数}\} = \{\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x},\ \boldsymbol{B}^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{0}\},
$$

依定理 1.2 及假设条件,有

$$
\begin{aligned}
\dim \mathcal{M}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{B}^{\perp}) &= \mathrm{rk}\begin{bmatrix}\boldsymbol{A}^{\mathrm{T}}\\ \boldsymbol{B}^{\mathrm{T}}\end{bmatrix} - \mathrm{rk}(\boldsymbol{B}^{\mathrm{T}}) = \mathrm{rk}(\boldsymbol{A} \vdots \boldsymbol{B}) - \mathrm{rk}(\boldsymbol{B}) \\
&= \mathrm{rk}(\boldsymbol{A}) = \dim(\mathcal{M}(\boldsymbol{A}^{\mathrm{T}})),
\end{aligned}
$$

但

$$
\mathcal{M}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{B}^{\perp}) \subset \mathcal{M}(\boldsymbol{A}^{\mathrm{T}}),
$$

于是

$$
\mathcal{M}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{B}^{\perp}) = \mathcal{M}(\boldsymbol{A}^{\mathrm{T}}),
$$

定理证毕.

1.2 矩阵求逆引理

分块矩阵的求逆公式和矩阵之和的求逆公式不仅在统计学的理论推导中经

常使用，而且在信号处理、神经网络、自动控制和系统理论中都具有广泛的应用，现分别介绍如下.

1.2.1 分块矩阵的求逆公式

引理 1.1 设矩阵

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

可逆.若 $|\boldsymbol{A}_{11}|\neq 0$，则

$$\boldsymbol{A}^{-1}=\begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22}\end{bmatrix}^{-1}=\begin{bmatrix}\boldsymbol{A}_{11}^{-1}+\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22.1}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1} & -\boldsymbol{A}_{11}^{-1}A_{12}\boldsymbol{A}_{22.1}^{-1}\\ -\boldsymbol{A}_{22.1}^{-1}\boldsymbol{A}_{21}A_{11}^{-1} & A_{22.1}^{-1}\end{bmatrix},\tag{1-2}$$

若 $|\boldsymbol{A}_{22}|\neq 0$，则

$$\boldsymbol{A}^{-1}=\begin{bmatrix}\boldsymbol{A}_{11.2}^{-1} & -\boldsymbol{A}_{11.2}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\\ -\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11.2}^{-1} & \boldsymbol{A}_{22}^{-1}+\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11.2}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\end{bmatrix},\tag{1-3}$$

式(1-2)中 $\boldsymbol{A}_{22.1}=\boldsymbol{A}_{22}-\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}$，$\boldsymbol{A}_{11.2}=\boldsymbol{A}_{11}-\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}$.

证明 注意到在矩阵 $\boldsymbol{A}$ 左乘一个单位下三角阵(亦称 Frobenius 矩阵)可以将 $\boldsymbol{A}$ 中的 $\boldsymbol{A}_{21}$ 打成零矩阵，即

$$\begin{bmatrix}\boldsymbol{I} & \boldsymbol{O}\\ \boldsymbol{X} & \boldsymbol{I}\end{bmatrix}\begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22}\end{bmatrix}=\begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{X}\boldsymbol{A}_{11}+\boldsymbol{A}_{21} & \boldsymbol{X}\boldsymbol{A}_{12}+\boldsymbol{A}_{22}\end{bmatrix},\tag{1-4}$$

其中 $\boldsymbol{X}$ 为待定矩阵.为了使右边矩阵右下角子矩阵变为零矩阵，令 $\boldsymbol{X}\boldsymbol{A}_{11}+\boldsymbol{A}_{21}=\boldsymbol{O}$，若 $|\boldsymbol{A}_{11}|\neq 0$，解得 $\boldsymbol{X}=-\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}$. 将此代入式(1-4)得

$$\begin{bmatrix}\boldsymbol{I} & \boldsymbol{O}\\ -\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1} & \boldsymbol{I}\end{bmatrix}\begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22}\end{bmatrix}=\begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{O} & -\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}+\boldsymbol{A}_{22}\end{bmatrix},$$

类似地，在上面矩阵方程中右乘一个单位上三角阵，则可以将 $\boldsymbol{A}$ 中的 $\boldsymbol{A}_{12}$ 打成零矩阵，即

$$\begin{bmatrix} \boldsymbol{I} & \boldsymbol{O} \\ -\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{I} & -\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12} \\ \boldsymbol{O} & \boldsymbol{I} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{O} \\ \boldsymbol{O} & -\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12} + \boldsymbol{A}_{22} \end{bmatrix}, \tag{1-5}$$

此式证明了 $\boldsymbol{A}_{22.1}$ 的可逆性.两边求逆矩阵,容易得到

$$\begin{aligned} \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix}^{-1} &= \begin{bmatrix} \boldsymbol{I} & -\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12} \\ \boldsymbol{O} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{A}_{11}^{-1} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{A}_{22.1}^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{I} & \boldsymbol{O} \\ -\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1} & \boldsymbol{I} \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{A}_{11}^{-1} + \boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22.1}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1} & -\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22.1}^{-1} \\ -\boldsymbol{A}_{22.1}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1} & \boldsymbol{A}_{22.1}^{-1} \end{bmatrix}, \end{aligned}$$

用完全类似的方法可以证明引理的后半部分.

1.2.2 矩阵之和的求逆公式

引理 1.2 (Sherman-Morrison 公式)令 $\boldsymbol{A}$ 是一个 $n \times n$ 的可逆矩阵,并且 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 是两个 $n \times 1$ 向量,使得 $\boldsymbol{A} + \boldsymbol{x}\boldsymbol{y}^{\mathrm{T}}$ 可逆,则

$$(\boldsymbol{A} + \boldsymbol{x}\boldsymbol{y}^{\mathrm{T}})^{-1} = \boldsymbol{A}^{-1} - \frac{\boldsymbol{A}^{-1}\boldsymbol{x}\boldsymbol{y}^{\mathrm{T}}\boldsymbol{A}^{-1}}{1 + \boldsymbol{y}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{x}}, \tag{1-6}$$

该引理称为矩阵求逆引理,是 Sherman 与 Morrison 于 1949 年和 1950 年得到的.

Woodbury 于 1950 年将矩阵求逆引理进一步推广为矩阵之和的求逆公式,即

引理 1.3 (Sherman-Morrison-Woodbury 公式)令 $\boldsymbol{A}$ 是一个 $n \times n$ 的可逆矩阵,并且 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 是两个 $n \times 1$ 向量,使得 $\boldsymbol{A} + \boldsymbol{UBV}$ 可逆,则

$$\begin{aligned} (\boldsymbol{A} + \boldsymbol{UBV})^{-1} &= \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{UB}(\boldsymbol{B} + \boldsymbol{BVA}^{-1}\boldsymbol{UB})^{-1}\boldsymbol{BVA}^{-1} \\ &= \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{U}(\boldsymbol{I} + \boldsymbol{BVA}^{-1}\boldsymbol{U})^{-1}\boldsymbol{BVA}^{-1}, \end{aligned} \tag{1-7}$$

或者

$$(\boldsymbol{A} - \boldsymbol{UV})^{-1} = \boldsymbol{A}^{-1} + \boldsymbol{A}^{-1}\boldsymbol{U}(\boldsymbol{I} - \boldsymbol{VA}^{-1}\boldsymbol{U})^{-1}\boldsymbol{VA}^{-1}, \tag{1-8}$$

构造出这两个公式当然不易,但验证这两个公式并不困难,留给读者作为后面的习题.

1.3 广义逆矩阵

定义 1.1 对于矩阵 $\boldsymbol{A}_{m\times n}$，一切满足方程组

$$\boldsymbol{AXA}=\boldsymbol{A} \tag{1-9}$$

的矩阵 $\boldsymbol{X}$ 称为矩阵 $\boldsymbol{A}$ 的广义逆，记为 $\boldsymbol{A}^{-}$.

定理 1.3 设 $\boldsymbol{A}$ 为 $m\times n$ 矩阵，$\mathrm{rk}(\boldsymbol{A})=r$. 若

$$\boldsymbol{A}=\boldsymbol{P}\begin{bmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{bmatrix}\boldsymbol{Q},$$

这里 $\boldsymbol{P}$ 和 $\boldsymbol{Q}$ 分别为 $m\times m$，$n\times n$ 的可逆阵，则

$$\boldsymbol{A}^{-}=\boldsymbol{Q}^{-1}\begin{bmatrix}\boldsymbol{I}_r & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{bmatrix}\boldsymbol{P}^{-1},$$

这里 $\boldsymbol{B}$，$\boldsymbol{C}$ 和 $\boldsymbol{D}$ 为适当阶数的任意矩阵.

证明 设 $\boldsymbol{X}$ 为 $\boldsymbol{A}$ 的广义逆，则有

$$\begin{aligned}\boldsymbol{AXA}=\boldsymbol{A}&\Leftrightarrow\boldsymbol{P}\begin{bmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{bmatrix}\boldsymbol{QXP}\begin{bmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{bmatrix}\boldsymbol{Q}=\boldsymbol{P}\begin{bmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{bmatrix}\boldsymbol{Q},\\ &\Leftrightarrow\begin{bmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{bmatrix}\boldsymbol{QXP}\begin{bmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{bmatrix}=\begin{bmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{bmatrix}.\end{aligned}$$

若记

$$\boldsymbol{QXP}=\begin{bmatrix}\boldsymbol{B}_{11} & \boldsymbol{B}_{12}\\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22}\end{bmatrix},$$

则上式

$$\Leftrightarrow\begin{bmatrix}\boldsymbol{B}_{11} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{bmatrix}=\begin{bmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{bmatrix}\Leftrightarrow\boldsymbol{B}_{11}=\boldsymbol{I}_r,$$

于是