广义逆矩阵与正则化方法

[德] FriederKuhnert

陈 杰 译

高等教育出版社

广义逆矩阵与正则化方法

〔德〕Frieder Kuhnert 著
陈　杰　译

高等教育出版社

这本小册子是从［德］Frieder Kuhnert 著《Pseudoinverse Matrizen und die Methode der regularisierung》(1976)一书翻译过来的. 它在几何直观的基础上引入广义逆矩阵，并联系到超定的，不定的，奇异的或病态的方程组的正规解和拟正规解，介绍了 Тихонов 的正则化方法.

本书可以作为数学、计算专业师生的参考书.

广义逆矩阵与正则化方法

［德］Frieder Kuhnert 著

陈　杰　译

*

高等教育出版社出版

新华书店北京发行所发行

通县觅子店印刷厂印装

*

开本 850×1168 1/32　印张 2.75 字数 63,000

1985 年 7 月第 1 版　1985年 7 月第 1 次印刷

印数 00,001—9,500

书号 13010·0977　定价0.73元

译　序

本书是关于广义逆矩阵的入门书．全书篇幅还不到百页，但却写得具有一定的特色．书中把广义逆矩阵的引入建立在几何直观的基础上，给人以颇深的印象．广义逆矩阵的数值计算有所论列．联系到(超定的、不定的、奇异的或病态的)方程组的正规解和拟正规解，本书还介绍了 Тихонов 的正则化方法．大家知道，这个方法远不限于矩阵逆的问题．对于这个方法有兴趣的读者，读过本书也可得到一个初步的概念．

关于广义逆矩阵的专著，七十年代中已有多种出版(参看本书文献目录中的[3]，[6]，[7]和[21])．但对于不想立即深入的读者，本书作为入门的导引也许更能适应多数读者的需要．前几年，我们曾将本书译稿油印，并分赠若干兄弟院校，据我们所知，不少同志对这本小书都给予了一定的评价．这大概因为我们国内到那时为止(迄今似乎还是这样)还没有一本关于广义逆矩阵的中文专书出版，同时也由于这本小书的特色还能引人入胜的缘故吧！事实上，本书不仅可以作为一般读者的一本有益的参考读物，略加补充，也可作为本科高年级学生一个学期的选修课教材．

现在本书有了出版的机会，这是值得感谢的．利用这个机会，在编辑同志的帮助下，我们把油印稿中有些误译做了一些订正．对原书中个别明显错误(原书是打字排版的，看来出书比较匆忙)也做了少量修正．但由于译者德文水平有限，对计算方法也缺少修养，错误之处，是难免的．希望读者发现后，不吝赐予指正．

译者

1984 年于内蒙古大学

广义逆矩阵与正则化方法

本书研究实际求解超定的和不定的方程组以及奇异的和病态的方程组的问题。 计算机只记录有限位数以及舍入误差的出现，使得有必要创建稳定的数值算法。首先，本书介绍了广义逆矩阵的概念，用它作为给出超定的、不定的以及奇异的方程组的正规解和拟正规解的一种统一方法. 叙述了广义逆矩阵的准确算法和迭代法. 最后，为得到一种稳定的计算方法，介绍了Тихонов的正则化方法，并给出了有关的误差估计.

本书是为那些对实际求解线性方程组有兴趣的读者编写的。不管他们是从事教学工作、研究工作还是实际工作，只要具有工程学院开设的线代数课的知识水平就可以阅读本书。

序　言

人们越来越多地需要对超定的或不定的以及奇异的线性方程组求数值解.

可以区别两类不同的问题：第一类，方程组是可解的. 这时，通常有无穷多个解. 第二类，方程组不可解，就是说，我们有一个矛盾方程组.

对第一类问题，人们往往并不需要知道整个的解集，需要的只是其中符合某种附加条件的一个子集，通常，我们从这个解集中挑选出所谓正规解，即具有最小范数的解.

对第二类问题，我们只能确定所谓拟解，即在最小二乘误差的意义下最好地满足方程组的解. 由于这种解一般也有无穷多个，而我们同样并不需要全知道，因而也有一个如何挑选的问题. 在这里，通常最令人感兴趣的仍然是那个具有最小范数的拟解，即所谓拟正规解.

正规解和拟正规解可以运用所谓广义逆矩阵的概念来加以刻划. 本书介绍了这个线代数的概念. 这个概念在有关代数教程中通常都不作介绍或者只作为次要内容略提一下它的定义. 我们在第二章中定义了广义逆矩阵，论证了它和线性方程组的正规解或拟正规解的关系，还论述了它与近年来在计算数学的理论与实践中都常用的奇异值概念，以及与奇异值概念相联系的任意长方阵的"对角形表示"问题之间的关系.

在第三章中，我们介绍了求广义逆矩阵的数值方法. 这些方法的意义是有限的，因为在大多数实际场合，需要的并不是广义逆

矩阵本身，而是它和一个向量的乘积，即线性方程组的正规解或拟正规解。但是，这些数值方法提供了广义逆矩阵的某些其它知识，而这些知识在我们叙述其它更进一步的数值方法时是有用的。

第四章中，我们讨论了求线性方程组的正规解和拟正规解的问题。首先我们判定了确定正规解或拟正规解是一个不稳定的问题，因此为了避免在实际计算中舍入误差会灾难性地积累起来，要实现一个有效的数值方法需要进行多方面的考虑。这一章中，我们介绍了用于有限维空间中线代数问题的 Тихонов 正则化方法。这个方法，除去可用于超定的、不定的以及奇异的方程组外，对所谓病态方程组也适用。本章的目的在于推导正则化问题的解的误差公式，公式中有对于初始误差和正则化参数的明确的依赖关系。

第二到第四章之前的第一章中，我们介绍了几章中需要的若干属于线代数范围的常识。

迄今已有一些内容丰富的关于广义逆矩阵的书，这些书中还研究了理论的应用，首先是在数理统计以及其它领域的应用(参看[3],[6],[7],[21])。在这些书中，除去广义逆矩阵外，还介绍了矩阵逆的若干更进一步的推广，但这些推广与线性方程组的正规解或拟正规解，并没有多少直接联系。

本书的手稿是作者1975年春在华沙 Stefan Banach 国际数学中心关于数学模型和计算方法作一个学期逗留时写成的。Banach 中心那种适宜于科学工作的环境以及有机会和社会主义国家科学界的同事们就数学的发展和应用的重要问题进行讨论，对完成这份手稿很有帮助。又作者从能够利用波兰科学院数学研究所组织得很好的专门图书馆这一点也得到了很大的助益。

1975年5月于华沙

目　录

1. 线代数的某些概念

1.1. 矩阵和线性算子

我们用$\mathscr{K}^n$ 记列矩阵 $x=(x_1, \cdots, x_n)^T$ 作成的线性空间，这里 $x_k\in K$ 通常是复数，上标T永远代表矩阵的转置，就是说，对于任意矩阵A, A^T 是对换A的行和列得到的矩阵.

在空间$\mathscr{K}^n$ 中，我们考虑普通的内积．对于 $x=(x_1, \cdots x_n)^T\in\mathscr{K}^n$, $y=(y_1, \cdots, y_n)^T\in\mathscr{K}^n$, 内积

$$(x, y)=\sum_{k=1}^n x_k\bar{y}_k,$$

其中数上的一横表示复数的共轭．我们用

$$\|x\|=\sqrt{(x, x)}=\left(\sum_{k=1}^n |x_k|^2\right)^{\frac{1}{2}}$$

记向量 $x\in\mathscr{K}^n$ 的范数．向量 $x, y\in\mathscr{K}^n$ 互相正交，如果有$(x, y)=0$ 成立.

众所周知，每一组 n 个线性无关的向量 $b_1, b_2, \cdots, b_n\in\mathscr{K}^n$都作成线性空间 $\mathscr{K}^n$ 的一个基，因此，每一个向量 $x\in\mathscr{K}^n$ 都可以唯一地表成一个线性组合

$$x=u_1b_1+\cdots+u_nb_n=\sum_{k=1}^n u_kb_k, \qquad u_k\in K.$$

向量

$e_1=(1, 0, \cdots, 0)^T$, $e_2=(0, 1, 0, \cdots, 0)^T$, $\cdots$, $e_n=(0, \cdots, 0, 1)^T$

是一个基，对于这个基，$x=(x_1, \cdots, x_n)^T$ 有表达式

$$x=\sum_{k=1}^{n} x_k e_k.$$

设 $\mathscr{K}^n$ 和 $\mathscr{K}^m$ 是两个线性空间，A 是一个映射关系，对于每一个向量 $x\in\mathscr{K}^n$，给它对应一个向量 $z=Ax\in\mathscr{K}^m$. 如果对于任意向量 $x, y\in\mathscr{K}^n$ 和任意数 $\alpha, \beta\in K$ 都有关系式

$$A(\alpha x+\beta y)=\alpha Ax+\beta Ay$$

成立，我们说映射关系 A 是一个线性算子.

设 $b_1, \cdots, b_n\in\mathscr{K}^n$, $d_1, \cdots, d_m\in\mathscr{K}^m$ 是这些空间的任意基. 我们有

$$x=u_1b_1+\cdots+u_nb_n,\quad z=v_1d_1+\cdots+v_md_m.$$

设令

$$Ab_k=c_{1k}d_1+\cdots+c_{mk}d_m=\sum_{j=1}^{m} c_{jk}d_j,\quad k=1, \cdots, n,$$

我们有

$$\sum_{j=1}^{m} v_jd_j=z=Ax=\sum_{k=1}^{n} u_kAb_k$$
$$=\sum_{k=1}^{n} u_k\sum_{j=1}^{m} c_{jk}d_j=\sum_{j=1}^{m}\left(\sum_{k=1}^{n} c_{jk}u_k\right)d_j.$$

由于向量 z 对基 $d_1, \cdots, d_m$ 的线性表达式是唯一的，由上面的等式得到 v_j 和 u_k 之间的下列关系式:

$$\sum_{k=1}^{n} c_{jk}u_k=v_j,\quad j=1, \cdots, m.$$

引进记号

$$u=(u_1, \cdots, u_n)^T,\quad v=(v_1, \cdots, v_m)^T,$$

$$C=\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{pmatrix},$$

上述关系可写成矩阵形式

$$Cu=v.$$

我们把这个(m,n)矩阵C称作线性算子A在基$b_1,\cdots,b_n\in\mathscr{K}^n$和$d_1,\cdots,d_m\in\mathscr{K}^m$中的表现。

对于特殊的基

$$e_1=(1,0,\cdots,0)^T,\quad e_2=(0,1,0,\cdots,0)^T,\cdots,$$
$$e_n=(0,\cdots,0,1)^T\in\mathscr{K}^n$$
$$f_1=(1,0,\cdots,0)^T,\quad f_2=(0,1,0,\cdots,0)^T,\cdots,$$
$$f_m=(0,\cdots,0,1)^T\in\mathscr{K}^m$$

我们有

$$x=\sum_{k=1}^{n}x_ke_k,\qquad z=\sum_{j=1}^{m}z_jf_j,$$

从而有(m,n)矩阵$A=(a_{jk})$使关系式

$$Ax=z\tag{1.1}$$

成立。

反过来，如果给定了一个(m,n)矩阵$A=(a_{jk})$，则公式(1.1)确定一个线性算子A，它在基$e_1,\cdots,e_n\in\mathscr{K}^n$和$f_1,\cdots,f_m\in\mathscr{K}^m$中的表现正好是矩阵$A$. 在这种相互关系的意义下，今后我们就用矩阵$A$来记线性算子$A$.

设$A=(a_{jk})$和$B=(b_{jk})$是两个(m,n)矩阵. 如果对线性空间$\mathscr{K}^n$的任意基$g_1,\cdots,g_n\in\mathscr{K}^n$中的任意向量都有等式

$$Ag_k=Bg_k,\quad k=1,\cdots,n\tag{1.2}$$

成立，则矩阵A和B相等：$A=B$.

事实上，对于$e_k=(0,\cdots,0,1,0,\cdots,0)^T$我们有

$$e_k=\alpha_1g_1+\cdots+\alpha_ng_n.$$

因此，根据(1.2)，有$Ae_k=Be_k$，也就是

$$(a_{1k},a_{2k},\cdots,a_{mk})^T=(b_{1k},b_{2k},\cdots,b_{mk})^T.$$

由此知 $a_{jk}=b_{jk}$, $j=1, \cdots, m$. 又由于 $k=1, \cdots, n$ 是任意的, 等式 $A=B$ 成立.

如果矩阵 A 是两个矩阵 G 和 H 的乘积, $A=G\cdot H$, 则矩阵 A 的列向量是矩阵 G 的列向量的线性组合, 而 A 的行向量则是矩阵 H 的行向量的线性组合.

设 $A=(a_{jk})$ 为一 (m, n) 矩阵, 则由 A 行列互换后再取共轭得到的 (n, m) 矩阵 $A^*=(\bar{a}_{kj})$ 称作 A 的共轭转置矩阵. 它是一个线性算子, 使得每一个向量 $z\in\mathscr{K}^m$ 对应于一个向量 $u=A^*z\in\mathscr{K}^n$. 对于任意向量 $x\in\mathscr{K}^n$ 和 $z\in\mathscr{K}^m$, 等式

$$(Ax, z)=(x, A^*z)$$

永远成立, 这里等式左端是线性空间 $\mathscr{K}^m$ 中的内积, 而右端则是线性空间 $\mathscr{K}^n$ 中的内积.

矩阵 A 的范数 $\|A\|$ 由下式定义:

$$\|A\|=\sup_{x\neq 0}\frac{\|Ax\|}{\|x\|}=\sup_{\|x\|=1}\|Ax\|=\max_{\|x\|=1}\|Ax\|,$$

这里, 向量 $z=Ax$ 的范数是在线性空间 $\mathscr{K}^m$ 中取的, 而向量 x 的范数则是在空间 $\mathscr{K}^n$ 中取的. 范数 $\|A\|$ 等于矩阵 A^*A 的最大特征值的平方根.

线性空间 $\mathscr{K}^n$ 的任意子集 $\mathscr{K}'$ 如果本身是一个线性空间, 则我们说它是 $\mathscr{K}^n$ 的一个线性子空间. 设 $a_1, \cdots, a_s\in\mathscr{K}^n$, 这些向量的一切线性组合作成一个集

$$\mathscr{D}=\left\{x=\sum_{k=1}^{s}\alpha_k a_k,\ \alpha_k\in K\right\}.$$

$\mathscr{D}$ 显然是 $\mathscr{K}^n$ 的一个线性子空间, 称为向量 $a_1, \cdots, a_s$ 的生成空间, 记作

$$\mathscr{D}=\operatorname{span}[a_1, \cdots, a_s].$$

如果我们把 $a_1, \cdots, a_s$ 作为列向量写在一起构成一个 (n, s) 矩

阵 A_s, 则子空间$\mathscr{D}$的维数等于矩阵 A_s 的秩:

$$\dim\mathscr{D} = \operatorname{rang} A_s.$$

设 A 为一矩阵, 所有被矩阵 A 映射成零向量的向量 $x\in\mathscr{K}^n$ 作成的集是一个线性子空间, 记作

$$N(A)=\{x\in\mathscr{K}^n: Ax=0\},$$

称为矩阵 A 的零空间或核. $N(A)$的正交补空间记作$S(A)$:

$$S(A)=\{u\in\mathscr{K}^n: (u, x)=0,\ x\in N(A)\}.$$

最后, 我们用记号 $R(A)$ 记矩阵 A 的像空间:

$$R(A)=\{z\in\mathscr{K}^m: z=Ax,\ x\in\mathscr{K}^n\},$$

这是空间 $\mathscr{K}^m$ 的一个线性子空间.

设 $A=(a_{jk})$是一个 n 阶方阵. 我们可以把它看作是从 $\mathscr{K}^n$ 仍到 $\mathscr{K}^n$ 本身的一个线性算子. 方阵

$$I=\begin{pmatrix} 1 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

称为单位阵. 对于有 $\det A\neq 0$ 的方阵 A, 存在逆阵 A^{-1}, 它满足条件

$$AA^{-1}=A^{-1}A=I.$$

当 $A=A^*$ 时, 方阵A称为一个 Hermite 阵, 方阵U满足 $U^*=U^{-1}$ 时则称为酉阵.

一类重要的 Hermite 阵是所谓正交射影阵 P, 它除去 $P^*=P$ 外还满足条件 $P^2=P$.

设 $N(P)$ 是矩阵P的零空间, 则对任意向量 $x\in\mathscr{K}^n$ 和 $u\in N(P)$, 都有

$$(Px, u)=(x, P^*u)=(x, Pu)=(x, 0)=0,$$

因此$Px\in S(P)$. 另一方面, 我们还知道向量 Px, $x\in\mathscr{K}^n$, 构成整个子空间 $S(P)$, 这是因为, 如果对一切 $x\in\mathscr{K}^n$, 向量 $w\in\mathscr{K}^n$ 总满足

条件$(Px, w)=0$, 则我们有

$$0=(Px, w)=(x, Pw),$$

由此得出 $Pw=0$, 也就是说 $w\in N(P)$.

从关系式

$$Px=P^2x=P(Px)$$

知道，对于线性子空间 $S(P)$的每一个向量 $v=Px$, $x\in\mathscr{K}^n$, 都有等式

$$Pv=v$$

成立.

如果我们把向量 $x\in\mathscr{K}^n$ 表成一个和

$$x=v+u,\quad v\in S(P),\quad u\in N(P),\quad (v, u)=0,$$

则我们有

$$Px=Pv+Pu=v+0=v,$$

因此, 矩阵P把每一个向量 $x\in\mathscr{K}^n$ 映射成它在子空间$S(P)$上的射影 v.

设$\mathscr{P}$是空间$\mathscr{K}^n$的一个线性子空间，又$\mathscr{Q}$是它的正交补空间. 我们用下式定义一个线性算子 $\boldsymbol{P}$:

$$\boldsymbol{P}v=v,\quad v\in\mathscr{P},$$
$$\boldsymbol{P}u=0,\quad u\in\mathscr{Q}.$$

这个算子在基 $e_1, \cdots, e_n\in\mathscr{K}^n$ 上的表现P是一个正交射影阵, 因为对于任意$x, y\in\mathscr{K}^n$, 从表达式

$$x=v+u, y=v_1+u_1, v, v_1\in\mathscr{P}, u, u_1\in\mathscr{Q}$$

知道

$$\begin{aligned}
&Px=Pv+Pu=v,\\
&P^2x=Pv=Pv+Pu=Px;\\
&(Px, y)=(v, v_1+u_1)=(v+u, v_1)\\
&\qquad=(x, Pv_1)=(x, Pv_1+Pu_1)=(x, Py),
\end{aligned}$$

由此立刻得出 $P^2=P$ 和 $P^*=P$.

1.2. 线性方程组

考虑线性方程组

$$
\begin{aligned}
&a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1,\\
&a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2,\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_n,
\end{aligned}
\tag{1.3}
$$

引入(m, n)矩阵

$$
A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}
$$

和向量 $x=(x_1, \cdots, x_n)^T\in\mathscr{K}^n$, $b=(b_1, \cdots, b_m)^T\in\mathscr{K}^m$, (1.3)可以写成下列等价形式

$$Ax=b \tag{1.4}$$

矩阵的各列可以看成线性空间 $\mathscr{K}^m$ 中的 n 个向量 $a_1, \cdots, a_n$:

$$a_j=(a_{1j}, a_{2j}, \cdots, a_{mj})^T\in\mathscr{K}^m,$$

由此, 方程组(1.3)还可以写成另一等价形式

$$a_1x_1+\cdots+a_nx_n=\sum_{j=1}^{n}a_jx_j=b \tag{1.5}$$

我们考虑相应于(1.3)但右端 $b_1=\cdots=b_m=0$ 的齐次方程组, 这时(1.4)变成

$$Ax=0 \tag{1.6}$$

因此, 齐次方程组的解正好就是矩阵A的零空间 $N(A)$ 中的向量. 如果我们把方程组(1.3) 的左端表成线性空间 $\mathscr{K}^n$ 中的内积的形

式，我们得到齐次方程组 $Ax=0$ 的下列等价形式:

$$
\begin{aligned}
&x_1\bar{a}_{11}+x_2\bar{a}_{12}+\cdots+x_n\bar{a}_{1n}=(x,(a^*)_1)=0,\\
&x_1\bar{a}_{21}+x_2\bar{a}_{22}+\cdots+x_n\bar{a}_{2n}=(x,(a^*)_2)=0,\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&x_1\bar{a}_{m1}+x_2\bar{a}_{m2}+\cdots+x_n\bar{a}_{mn}=(x,(a^*)_m)=0,
\end{aligned}
\tag{1.7}
$$

其中$(a^*)_k=(\bar{a}_{k1},\bar{a}_{k2},\cdots,\bar{a}_{kn})^T\in\mathscr{K}^n$，$k=1,\cdots,m$，正是矩阵 A^* 的列向量.

方程(1.7)说明，齐次方程组(1.6)的解 x 和所有向量$(a^*)_k$都正交. 向量 $(a^*)_1,\cdots,(a^*)_m$ 在空间 $\mathscr{K}^n$ 中生成一个线性子空间，这个子空间正是零空间 $N(A)$的正交补空间，这就说明

$$S(A)=\operatorname{span}[(a^*)_1,\cdots,(a^*)_m],$$

完全类似地可以得到

$$S(A^*)=\operatorname{span}[a_1,\cdots,a_n]\subset\mathscr{K}^m,$$

这个$S(A^*)$与共轭转置矩阵 A^* 的零空间 $N(A^*)$正交，也就是说，它与齐次方程组 $A^*z=0$ 的一切解 z 正交.

现在，表达式(1.5)清楚地说明，非齐次方程组(1.3)在而且只在右端向量 b 可以表成 $a_1,\cdots,a_n$的一个线性组合时才有解，换句话说，当而且仅当 $b\in S(A^*)$，也就是 b和$N(A^*)$正交时，方程(1.3)才有解.

把矩阵A看作把向量 $x\in\mathscr{K}^n$ 映射成向量 $y=Ax\in\mathscr{K}^m$ 的线性算子，类似(1.5)，我们可以得到表达式

$$y=x_1a_1+\cdots+x_na_n=Ax.$$

这个关系式说明，算子A的像 y 由向量 $a_1,\cdots,a_n$ 的一切线性组合作成，换句话说，我们有

$$R(A)=\{y:y=Ax,x\in\mathscr{K}^n\}=\operatorname{span}[a_1,\cdots,a_n]=s(A^*). \tag{1.8}$$

对于矩阵 A^*，有类似的关系式

$$R(A^*) = \{v : v = A^* z,\ z \in \mathcal{K}^m\}$$
$$= \mathrm{span}[(\boldsymbol{a}^*)_1, \cdots, (\boldsymbol{a}^*)_m] = s(A). \tag{1.9}$$