矩阵最小二乘法理论及其应用

肖明耀

中国计量科学研究院
1978

矩阵最小二乘法理论及其应用

第四机械工业部第十研究院翻印
河北省束鹿县印刷厂印刷
一九七八年五月

前　　言

对一切从事精密科学实验的同志说来，应用最小二乘法来解决实际问题，是目前几乎离不开的手段。计量科学战线上的广大科研与实验人员，除以很大精力投入到他的本专业的工作中去以外，也迫切需要掌握这种方法。这种形势，推动了我们计量院内部的最小二乘法的讲座的工作。这篇资料也就在73年8月边讲边写的情况下初次印出来了。

在这篇资料中，作者不单纯追求理论上的完善，而着重于运算的能力，着重于在一大堆实验资料面前，如何尽快地解算出最佳值及其误差，以及由最佳值组成的函数的误差，并尽力规格化，以便使读者能应用该原理，举一反三地独立地应用于实践的计算为目标.力图贯彻毛主席关于“少而精”的教导，以便使人抓住要点和实质，节省阅读时间。

矩阵是研究有限维空间线性变换的解析工具。在最小二乘法理论中，利用它能迅速地得到一些重要公式，既大大缩小篇幅，也利于抓住实质性公式，便于记忆和应用，还利于掌握国内外文献中处理数据时经常出现的矩阵的符号和运算。鉴于这样的原因，用矩阵运算推证最小二乘法的一整套重要公式，是完全必要的了。

在实践的推动下，在74.12.华北地区角度学习班和75.4华东地区计量学习班印行该资料的基础上，作者又对该资料

作了一些坛删和修改，但缺点和错误仍会存在，想法也未必能达到，敬请指正。

肖明耀　1978年　于

中国计量科学研究院

注：参照75年10月江苏省无线电计量协作组厉办学习班印的资料“矩阵最小二乘法理论及其在计量科学中的应用”，作者于78年再次修改补充而印成这本小册子，目的是为了使其适应面更广，增加计量界、测量界、物理界等等。

目　　录

第一章　基础予备……………………………………(1)

第一节　矩阵与行列式………………………………(1)

1.矩阵的定义，2.矩阵的加法，3.矩阵与标量的乘积，4.矩阵的乘积，5.矩阵的转置，6.行列式的引出，7.行列式的定义，8.行列式的性质，9.线性方程组解的另一表达方法，10.矩阵的逆，11.矩阵的导数，12.矩阵的分块，13.矩阵的秩，14.矩阵的迹。

第二节　其它…………………………………………(26)

1.期望、方差与相关系数，2.方差的传递，3.权及其传递，4.向量的协差阵。

第二章　独立测量的最小二乘法〔参数平差〕……(31)

第一节　参数平差基本公式…………………………(31)

1.一个实际问题，2.线性模型，3.最小方差意义下线性无偏估计及其协差阵，4.$\hat{x}$的线性函数的方差及其相关系数，5.残差平方和，6.单一测量方差，7.$\hat{x}$的线性函

数的随机不确定度，8.非线性函数的线性化，9.不等权测量的单位权化。
第四节　解的最佳性定理……………………………（43）
第五节　回归线的计算……………………………（46）
第六节　不等权测量的处理………………………（52）
第七节　高斯约化法………………………………（60）

第三章　附有条件的最小二乘法〔附条件的参数平差〕……………………………………………（69）
第八节　基本公式…………………………………（69）
第九节　差值等权全组合比对的基本公式…………（73）
第十节　已知$x_1+\cdots+x_t=tw$时全组合比对计算公式……………………………………………（76）
第十一节　已知$x_1=w_1,\cdots,x_s=w_s$〔$s<t$〕时全组合比对计算公式……………………（79）
第十二节　比值全组合比对的计算原理……………（85）
第十三节　条件平差公式…………………………（87）

第四章　非独立测量的最小二乘法〔相关平差〕……………………………………………………（90）
第十四节　问题的提出……………………………（90）
第十五节　予备定理………………………………（93）
第十六节　附条件的参数相关平差…………………（95）
1.基本公式，2.参数相关平差公式，3.条件相关平差公式，4.算例。

第十七节　条件分组相关平差……………………………（108）

1.改化法条件分组，2.顺序法条件分组。

第十八节　条件及条件分组相关平差举例…………（114）

第十九节　带参数的条件相关平差………………………（121）

后语………………………………………………………………（122）

附录1　t一分布表………………………………………………（124）

主要参考资料……………………………………………………（125）

第一章　基础予备

第一节　矩阵与行列式

1.矩阵的定义

有$n\times m$个数，按n行和m列排列起来并用括号括起，则这$n\times m$个数称为矩阵，用符号表示成

$$\begin{pmatrix} a_{11} & a_{12}\cdots a_{1m} \\ a_{21} & a_{22}\cdots a_{2m} \\ \cdots\cdots & \cdots\cdots \\ a_{n1} & a_{n2}\cdots a_{nm} \end{pmatrix} = (a_{ij}) = A(n\times m) = A = \underset{n\times m}{A}$$

A或(a_{ij})称为矩阵，它是$n\times m$个数的简单表示，其中$n\times m$称为矩阵的阶，a_{ij}（$i=1, \cdots, n$；$j=1, \cdots, m$）称为矩阵元素，而a_{ii}称为矩阵的主对角线元素。

当$n=m$时，则

$$A(n\times n) = A_n = A$$

A称为n阶方阵，有时为叙述方便，也统称为矩阵。

当$m=1$（或$n=1$）时，则称为向量，并用小写字母表示，例如

$$a_{\cdot j} = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}$$

$a_{\cdot j}$称为矩阵A的第j列向量，在未特别指明时，小写字

母通常代表列向量，例如

$$x=\begin{pmatrix} x_1 \\ \vdots \\ x_t \end{pmatrix}$$

等等。

下面几个特定矩阵是本文运算中常用的：

$$I=\begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}=(\delta_{ij})$$

式中矩阵元素

$$\delta_{ij}=\begin{cases} 1, & i=j \\ 0, & i\neq j \end{cases}$$

$$O=\begin{pmatrix} 0\cdots 0 \\ \cdots\cdots\cdots \\ 0\cdots 0 \end{pmatrix}$$

$$J=\begin{pmatrix} 1\cdots 1 \\ \cdots\cdots \\ 1\cdots 1 \end{pmatrix}=11'$$

$$1=\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

$$\varepsilon_i=\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}\cdots_i\text{行为}1$$

I称为单位矩阵、么矩阵或恒等矩阵，它的主对角线元素全为1，而非主对角线元素全为零。O称为零矩阵，它的元素全为零。J称为满1矩阵，**1**称为满1向量，它的元素全为1。ε_i称为单位向量，它的第i行元素为1，其余为0，一组ε_i（$i=1, \cdots, n$）称为标准化正交基。

两个同阶矩阵$A=B$，是指其对应元素$a_{ij}=b_{ij}$；反之亦然，即

$$A=B \Leftrightarrow a_{ij}=b_{ij}，对一切 i，j 成立 \qquad (1.1)$$

2.矩阵的加法

定义：$A+B=(a_{ij})+(b_{ij})=(a_{ij}+b_{ij})$

例：

$$\begin{pmatrix} 2 & 3 \\ 1 & -1 \end{pmatrix}+\begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}=\begin{pmatrix} 3 & 4 \\ 3 & 0 \end{pmatrix}$$

A和B行同、列同才有定义，所以可证

$$\left.\begin{array}{l} A+B=B+A \\ (A+B)+C=A+(B+C) \\ A+O=A \end{array}\right\} \qquad (1.2)$$

3.矩阵与标量的乘积

定义：$\alpha A=\alpha(a_{ij})=(\alpha a_{ij})$

例：

$$2\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}=\begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix}$$

所以可证

$$\left.\begin{array}{l} (\alpha+\beta)A=\alpha A+\beta A \\ \alpha(A+B)=\alpha A+\alpha B \\ (\alpha\beta)A=\alpha(\beta A) \end{array}\right\} \qquad (1.3)$$

4.矩阵的乘积

定义：$AB=(a_{ir})(b_{rj})=(\sum\limits_{r}a_{ir}b_{rj})$

例：

$$\begin{pmatrix}-1 & 2 & 3\\ 5 & 1 & 2\end{pmatrix}\begin{pmatrix}1 & 2\\ 4 & 1\\ 3 & 1\end{pmatrix}$$

$$=\begin{pmatrix}(-1)\times1+2\times4+3\times3 & (-1)\times2+2\times1+3\times1\\ 5\times1+1\times4+2\times3 & 5\times2+1\times1+2\times1\end{pmatrix}$$

$$=\begin{pmatrix}16 & 3\\ 15 & 13\end{pmatrix}$$

定义说明，两矩阵乘积AB的第i行j列元素是由A中第i行所有元素与B中第j列所有对应元素的乘积之和构成，故当A的列数与B的行数相同时，AB有定义，此时AB的阶数为A的行数和B的列数之积。

所以可证

$$\left.\begin{array}{l}(AB)C=A(BC)\\ \alpha(AB)=(\alpha A)B\\ (A+B)C=AC+BC\\ A(B+C)=AB+AC,\ A+AB=A(I+B)\\ (A+B)(C+D)=A(C+D)+B(C+D)\\ AI=IA=A\\ AB\neq BA\text{（一般）}\\ A^2=A\cdot A,\ A^3=A\cdot A\cdot A\text{等等，}A\text{为方阵时}\\ A=(a_{ij})\text{时，则}a_{ij}=\varepsilon_i'A\varepsilon_j\end{array}\right\}\qquad(1.4)$$

这里ε_i'为ε_i的转置（见后），可见，矩阵的元素也可以用矩阵表达出。

5. 矩阵的转置

定义：$A'=(a_{ij})'=(a_{ji})$

例

$$\begin{pmatrix}2 & 3 & 1\\ 7 & 8 & 9\end{pmatrix}'=\begin{pmatrix}2 & 7\\ 3 & 8\\ 1 & 9\end{pmatrix}$$

所以可证

$$\left.\begin{aligned}&(A')'=A\\&(AB)'=B'A'\\&(A+B)'=A'+B'\\&(\alpha A)'=\alpha A'\\&(A^{-1})'=(A')^{-1}\end{aligned}\right\}\qquad(1.5)$$

这里A^{-1}为A的逆阵，见后。

$A=A'$的矩阵，称为对称矩阵。

6. 行列式的引出

例如解一个二阶线性方程组

$$\begin{cases}a_{11}x_1+a_{12}x_2=y_1\cdots\cdots①\\ a_{21}x_1+a_{22}x_2=y_2\cdots\cdots②\end{cases}$$

用消去法

$a_{22}\times①-a_{12}\times②$　　得

$$(a_{11}a_{22}-a_{12}a_{21})x_1=y_1a_{22}-y_2a_{12}$$

$a_{11}\times②-a_{21}\times①$　　得

$$(a_{11}a_{22}-a_{12}a_{21})x_2=y_2a_{11}-y_1a_{21}$$

若用符号规定

$$\begin{vmatrix}a & b\\ c & d\end{vmatrix}=ad-cb$$

则当

$$|A|=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}=a_{11}a_{22}-a_{12}a_{21}\neq 0$$

时，x_1、x_2的解为

$$x_1=\begin{vmatrix} y_1 & a_{12} \\ y_2 & a_{22} \end{vmatrix}\div\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}=\frac{1}{|A|}\begin{vmatrix} y_1 & a_{12} \\ y_2 & a_{22} \end{vmatrix}$$

$$x_2=\begin{vmatrix} a_{11} & y_1 \\ a_{21} & y_2 \end{vmatrix}\div\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}=\frac{1}{|A|}\begin{vmatrix} a_{11} & y_1 \\ a_{21} & y_2 \end{vmatrix}$$

推广到任意n阶线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=y_1 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=y_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=y_n \end{cases}\Rightarrow A^x=y \qquad (1.6)$$

则当

$$|A|=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots\cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}\neq 0$$

时，x_1，…，x_n的解为

$$x=\frac{1}{|A|}\begin{vmatrix} a_{11}\cdots a_{1(i-1)} & y_1 & a_{1(i+1)}\cdots a_{1n} \\ a_{21}\cdots a_{2(i-1)} & y_2 & a_{2(i+1)}\cdots a_{2n} \\ \cdots\cdots\cdots & \cdots & \cdots\cdots\cdots \\ a_{n1}\quad a_{n(i-1)} & y_n & a_{n(i+1)}\cdots a_{nn} \end{vmatrix}\ i=1,\ \cdots,\ n \qquad (1.7)$$

这就是有名的克莱姆法则。那么，行列式怎样定义和计算才能使（1.7）符合（1.6）式呢？下面进一步讨论。

7.行列式定义

定义1：有$n\times n$个数排列n行n列，用竖线括起，代表一个数值则称之为行列式，用符号表示成

$$|A|=\begin{vmatrix} a_{11} & a_{12}\cdots a_{1n} \\ a_{21} & a_{22}\cdots a_{2n} \\ \cdots\cdots & \cdots\cdots\cdots \\ a_{n1} & a_{n2}\cdots a_{nn} \end{vmatrix}$$

这是n阶行列式，可见，方阵的行列式才有定义。

定义2：划去n阶行列式的第i行和第j列的所有元素后剩下的元素组成的行列式，并乘以$(-1)^{i+j}$后，则称之为a_{ij}元素的代数余子式，用符号表示为

$$A_{ij}=(-1)^{i+j}\begin{array}{cl} \begin{vmatrix} a_{11} & a_{12}\cdots & \cdots a_{1n} \\ a_{21} & a_{22}\cdots & \cdots a_{2n} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ \hline \cdots\cdots & \cdots\cdots & \cdots\cdots \\ a_{n1} & a_{n2} & a_{nn} \end{vmatrix} & \begin{array}{l} \\ \\ \\ i\text{行} \\ \\ \end{array} \\ \qquad\qquad j\text{列} & \end{array}$$

A_{ij}是$n-1$阶行列式。

定义3：行列式的值按下式计算

$|A|=a_{i1}A_{i1}+a_{i2}A_{i2}+\cdots+a_{in}A_{in}$，按任意第i行

展开

$$=a_{1j}A_{1j}+a_{2j}A_{2j}+\cdots+a_{nj}A_{nj}$$，按任意第 j 列展开

从上可见，$|A|$为n阶，A_{ij}为n−1阶，依次降阶，最后可算得$|A|$，实质上，行列式是一个数值，是用n×n阶元素来表示的一个数值，按照这样定义后，解（1.7）就能满足（1.6）式了。

例：

$$A=\begin{pmatrix}1 & 2 & 3\\ 4 & 5 & 6\\ 7 & 8 & 9\end{pmatrix}$$

求$|A|$。

解：按第1列展开得

$$|A|=1\cdot\begin{vmatrix}5 & 6\\ 8 & 9\end{vmatrix}-4\cdot\begin{vmatrix}2 & 3\\ 8 & 9\end{vmatrix}+7\cdot\begin{vmatrix}2 & 3\\ 5 & 6\end{vmatrix}=0$$

按第2行展开得

$$|A|=-4\cdot\begin{vmatrix}2 & 3\\ 8 & 9\end{vmatrix}+5\cdot\begin{vmatrix}1 & 3\\ 7 & 9\end{vmatrix}-6\cdot\begin{vmatrix}1 & 2\\ 7 & 8\end{vmatrix}=0$$

也可按其它任意行列展开，其结果都是一样的。

定义4：$|A|\neq 0$时，称A为非奇异（非退化）矩阵；反之亦反。很明显，A为非奇异矩阵时，方程组$Ax=y$才有唯一解。

8.行列式的性质

仅从计算的需要归纳如下：

（1）零性质：

①某行（列）的元素全为零时，行列式为零；

②两行（列）的元素全等时，行列式为零；

③两行（列）的元素成比列时，行列式为零；

（2）恒等性：

①行列式与它的转置行列式相等，即$|A|=|A'|$；

②某行（列）的公因子提出后，行列式不变，例

$$\begin{vmatrix} \alpha a_{11} & \alpha a_{12} & \alpha a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \alpha \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

③某行（列）的元素同乘一数加到另一行（列）后，行列式不变，例

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11}+\alpha a_{21} & a_{12}+\alpha a_{22} & a_{13}+\alpha a_{23} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

（3）交换变向性：两行（列）互相交换位置后，行列式改变符号；

（4）分解（或可加）性：某行（列）的各元素是二项之和则可以分解为二行列式之和；反之亦然。

例：

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ b_{21}+c_{21} & b_{22}+c_{22} & b_{23}+c_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ b_{21} & b_{22} & b_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ c_{21} & c_{22} & c_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

（5）代数余子式性质

$$a_{i1}A_{K1}+a_{i2}A_{K2}+\cdots+a_{in}A_{Kn}=\begin{cases} |A|, & \text{当}K=i\text{时} \\ 0, & \text{当}K\neq i\text{时} \end{cases}$$

$$a_{1j}A_{1K}+a_{2j}A_{2K}+\cdots+a_{nj}A_{nK}=\begin{cases} |A|, & \text{当}K=j\text{时} \\ 0, & \text{当}K\neq j\text{时} \end{cases}$$

（6）两个n阶方阵乘积的行列式等于两个n阶行列式的乘积，即

$$|AB|=|A|\cdot|B|$$

下面给出一个利用行列式性质进行简单计算的例子：

例：$\begin{vmatrix} 3 & 1 & -1 & 2 \\ -5 & 1 & 3 & -4 \\ 2 & 0 & 1 & -1 \\ 1 & 0 & 3 & -3 \end{vmatrix}$

$=\begin{vmatrix} 3 & 1 & -1 & 1 \\ -5 & 1 & 3 & -1 \\ 2 & 0 & 1 & 0 \\ 1 & 0 & 3 & 0 \end{vmatrix}$ 3列加到4列上