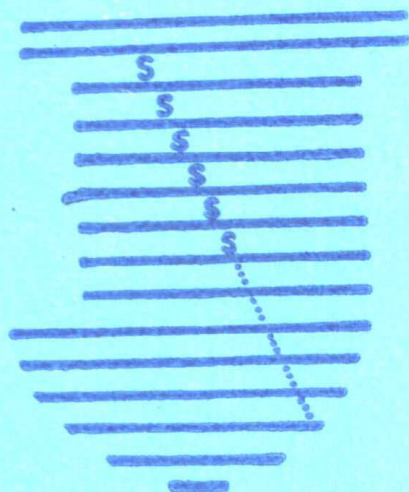


解病态线性方程组的奇异值分解法及其应用

何昌礼 编著



中国地质大学出版社

解病态线性方程组的奇异值分解法及其应用

何昌礼 编著

责任编辑 成金华

中国地质大学出版社出版

(武汉市 喻家山)

中国地质大学出版社印刷厂印刷 湖北省新华书店经销

开本 787×1092 1/32 印张5.5 字数121 千字

1990年5月第1版 1990年5月第1次印刷

印数1—1000册 定价：1.50元

ISBN 7-5625-0395-8/O·21

内 容 提 要

奇异值分解法是反演过程中解病态线性方程组的较为有效的方法,具有克服病态能力强、运算中不放大误差的优点。本书系统地介绍了这方面的理论和技术,国内这种系统介绍还不多见。本书对从事计算机软件、应用数学、地球物理专业和其它相近专业的人员有帮助,亦可以作为有关专业研究生和大学生的教材。

前 言

在地球物理以及其它科学领域中，大约有75%的科学技术问题都以不同方式、不同程度涉及到解反演问题。而反演过程中又经常碰到解线性方程组的问题，解方程方法的优劣直接影响到反演解释的效果。许多事实证明，有时候某个方法效果不好，并不是方法本身的缺陷，而是构造出来的线性方程组病态程度高，用一般的解方程方法不能取得稳定合理的解所致。近20年来，国外普遍采用奇异值分解方法 (Singular Value Decomposition Method) 来求解线性方程组，取得了好的结果。奇异值分解方法，具有克服病态能力强，运算中不放大误差的优点。奇异值分解方法属于广义逆方法中的一种，近年来广义逆理论的发展与完善，促进了求解线性方程组方法理论的发展与完善。奇异值分解方法在我国应用还不普遍，有关资料多为外文文献。

本书系统地介绍了奇异值分解法的理论基础、应用方法与技术，在整理国内外有关文献中的奇异值分解的理论和作法的基础上，加进了作者近年来在这方面的科研成果。本书在写作上力求由浅入深，难点之处辅以例题加以说明。这样，有助于读者系统、深入地学习和掌握，并且有助于提高反演解释能力与效果：

本书承蒙中国地质大学数学教授侯吉占和地球物理教授周国藩审阅和关心，作者在此表示感谢！

目 录

第一章 普通逆	(1)
一、各类线性方程组的建立.....	(1)
二、法方程.....	(6)
三、普通逆矩阵的求法.....	(6)
四、线性方程组求解中的问题.....	(15)
第二章 奇异值分解的基础知识	(20)
一、线性空间.....	(20)
二、欧氏空间.....	(22)
三、向量和矩阵的范数.....	(23)
四、相似变换.....	(26)
五、正交矩阵与正交变换.....	(28)
六、QR分解.....	(56)
第三章 广义逆的一般概念及性质	(60)
一、广义逆 A^- 的一般概念.....	(60)
二、广义逆 A^- 在解线性方程组中的应用.....	(74)
三、广义逆 A^+ 的定义与性质.....	(88)
四、广义逆 A^+ 的计算方法.....	(94)
第四章 奇异值分解法	(105)
一、奇异值分解.....	(105)
二、用奇异值分解法解线性方程组.....	(109)
三、奇异值分解法解线性方程组的辅助信息.....	(111)
四、奇异值分解法与阻尼最小二乘法.....	(116)
五、奇异值分解的算法.....	(120)

六、奇异值分解的计算实例.....	(137)
七、奇异值分解的取解方法.....	(149)
八、奇异值分解的应用.....	(160)
参考文献.....	(168)

第一章 普通逆

一、各类线性方程组的建立

在科技工程领域中,经常要碰到解线性方程组的问题,如震源参数的测定和物探的反演计算问题。

1. 震源参数的测定

一次地震发生,要求准确测定四个基本震源参数 λ 、 φ (震中经纬度)、 h (震源深度)、 T (发生时刻),这也是一个反演问题。

设有 M ($M \geq 4$)台记录仪记录到该次地震,各台观测的到时相应为 t_i^0 ($i=1, 2, \dots, M$)。

根据震源参数 λ 、 φ 、 h 、 T 的初值(任意假定或粗略估算),可以计算各台的理论到时为 $t_i^e(\lambda, \varphi, h, T)$, ($i=1, 2, \dots, M$)。换言之,震源参数决定了理论到 t_i^e ,求出各台观测到时与理论到时残差

$$\varepsilon_1 = t_1^0 - t_1^e(\lambda, \varphi, h, T)$$

$$\varepsilon_2 = t_2^0 - t_2^e(\lambda, \varphi, h, T)$$

$$\vdots$$

$$\varepsilon_M = t_M^0 - t_M^e(\lambda, \varphi, h, T)$$

根据各台到时残差平方和构造一个目标函数

$$\Phi(\lambda, \varphi, h, T) = \sum_{i=1}^M \varepsilon_i^2$$

$$= \sum_{i=1}^M [t_i^0 - t_i^e(\lambda, \varphi, h, T)]^2$$

求目标函数 $\Phi(\lambda, \varphi, h, T) = \min$, 从中解出四个参数 λ, φ, h, T , 就达到反演计算目的。

2. 物探找矿的反演计算问题

重力、磁法、电法的观测异常, 一般反映了地下介质的不均匀性。如果存在矿体, 就可以在地面用相应的仪器测出各点的异常值 $Z_i (i=1, 2 \dots M)$ 。而引起异常的场源可以用一组几何的及物理的参数来描述。例如一个二维有限延伸板状体, 其参数包括板的半宽度 b 、下延深度 l 、板倾角 α 、磁化强度的倾角 i 及磁化强度 J 等5个, 则解向量写为

$$x = (b, l, \alpha, i, J)^T$$

对给定的初始参数值, 可计算各测点的理论异常值 $Z_1^e(x), Z_2^e(x), \dots, Z_M^e(x)$ 。由相应点的实测值与计算值 (也称理论值) 相减, 便可得残差 ε_i 。

$$\varepsilon_i = Z_i - Z_i^e$$

其构造的目标函数为

$$\Phi(x) = \sum_{i=1}^M \varepsilon_i^2 = \sum_{i=1}^M [Z_i - Z_i^e(x)]^2$$

取 $\Phi(x) = \min$, 用一定的数学方法可以求出场源参数向量 $x = (b, l, \alpha, i, J)^T$ 。

综合这两个例子可以看出, 反演中数学模型的建立过程为:

由实际观测得到的 M 个离散的地球物理数据, 以 M 维向量表示

$$y^0 = (y_1^0, y_2^0, \dots, y_M^0)^T \quad (1-1.1)$$

它称为观测值向量。

一个具体的地球物理问题，可以用 N 个离散的（具有相同或不同量纲）物理参数表示反演计算所要求的模型参数，写为 N 维向量

$$X = (x_1, x_2, \dots, x_N)^T \quad (1-1.2)$$

称其为“模型参数”或参数向量。

由给定的“模型参数”，可以计算一组地球物理理论值（也称模型响应），它可用 M 维向量表示

$$Y^c(X) = (y_1^c(X), y_2^c(X), \dots, y_M^c(X))^T \quad (1-1.3)$$

称其为模型理论值向量。

地球物理反演问题的中心问题是用一定的数学或物理方法，寻找一个“模型参数”，使它尽量与真实场源参数符合。换言之，即寻找参数向量 X ，使其计算的理论向量 Y^c 与观测值向量 y^0 尽可能符合。为此，可利用最小二乘法，构造一个 M 维残差向量

$$\varepsilon(X) = (\varepsilon_1(X), \varepsilon_2(X), \dots, \varepsilon_M(X))^T \quad (1-1.4)$$

式中

$$\varepsilon_1(X) = y_1^0 - y_1^c(X)$$

$$\varepsilon_2(X) = y_2^0 - y_2^c(X)$$

$$\vdots \quad \quad \quad \vdots$$

$$\varepsilon_M(X) = y_M^0 - y_M^c(X)$$

令残差向量范数平方极小

$$\|\varepsilon(X)\|^2 = \sum_{i=1}^M \varepsilon_i^2(X) = \min \quad (1-1.5)$$

经过一定的数学变换或采用某些专门的数学方法，可根据

(1-1.5) 式解出 X , 从而使问题得解。

如 (1-1.3) 式中, $Y^e(X)$ 为模型参量的线性函数, 则各分量可以写为

$$\begin{aligned} y_1^e(X) &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1N}x_N \\ y_2^e(X) &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2N}x_N \\ &\vdots \\ y_M^e(X) &= a_{M1}x_1 + a_{M2}x_2 + \cdots + a_{MN}x_N \end{aligned} \quad (1-1.6)$$

式中 a_{ij} 为常系数。这可以直接写为矩阵形式

$$Y = AX \quad (1-1.7)$$

其中 Y 为 M 维向量, X 为 N 维向量, A 为 $M \times N$ 矩阵 (系数矩阵)。

一般说来, 地球物理问题中都是非线性函数。这时, 可在某一初始点 X 处对其进行泰勒展开, 并略去二阶以上高次项。(1-1.3) 式可写成多元函数泰勒展开式

$$\begin{aligned} y_1^e(X) &= y_1^e(x_0) + \sum_{j=1}^N \left(\frac{\partial y_1^e}{\partial x_j} \right)_{x_0} \Delta x_j \\ y_2^e(X) &= y_2^e(x_0) + \sum_{j=1}^N \left(\frac{\partial y_2^e}{\partial x_j} \right)_{x_0} \Delta x_j \\ &\vdots \\ y_M^e(X) &= y_M^e(x_0) + \sum_{j=1}^N \left(\frac{\partial y_M^e}{\partial x_j} \right)_{x_0} \Delta x_j \end{aligned} \quad (1-1.8)$$

写成矩阵形式为

$$Y^e(X) = Y^e(x_0) + A \Delta X \quad (1-1.9)$$

式中 A 称为偏导数矩阵 (或雅可比矩阵), 其元素为

$$a_{ij} = \left. \frac{\partial y_i}{\partial x_j} \right|_{x_0} \quad \begin{pmatrix} i=1, \dots, M \\ j=1, \dots, N \end{pmatrix}$$

写得更具体是

$$A = \begin{bmatrix} \frac{\partial y_1^e}{\partial x_1} & \frac{\partial y_1^e}{\partial x_2} & \cdots & \frac{\partial y_1^e}{\partial x_N} \\ \frac{\partial y_2^e}{\partial x_1} & \frac{\partial y_2^e}{\partial x_2} & \cdots & \frac{\partial y_2^e}{\partial x_N} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial y_M^e}{\partial x_1} & \frac{\partial y_M^e}{\partial x_2} & \cdots & \frac{\partial y_M^e}{\partial x_N} \end{bmatrix} \quad (1-1.10)$$

$M \times N$

ΔX 为参数 X 的改正量

$$\Delta X = (\Delta x_1, \Delta x_2, \cdots, \Delta x_N)^T \quad (1-1.11)$$

由 (1-1.9) 式, 可将残差向量写为

$$\begin{aligned} \varepsilon(X) &= Y^0 - Y^e(X) \\ &= Y^0 - Y^e(x_0) - A \Delta X \end{aligned} \quad (1-1.12)$$

$$\text{令} \quad Y^0 - Y^e(x_0) = B \quad (1-1.13)$$

式中 Y^0 是观测值向量

向量 $B = (B_1, B_2, \cdots, B_M)^T$ 表示观测值与初始模型理论计算值之差, 则 (1-1.12) 式可改写为

$$A \Delta X = B - \varepsilon(X) \quad (1-1.14)$$

(1-1.14) 式与 (1-1.7) 线性关系式是相似的。这说明, 一个非线性问题可以用迭代方法化为线性问题进行计算。在每次迭代时, 都要从新的原点出发, 重新计算雅可比矩阵 A , 直到残差向量范数 $\|\varepsilon\|$ 极小为止。

因此, 今后我们计算中只考虑线性方程组

$$AX = B \quad (1-1.15)$$

的解法。对于反演计算为线性函数情况, 可以直接把 X 作为模型参数, B 作为观测值向量。对于非线性函数情况, 可把 x 作为模型参数的改正量, B 作为观测值与计算值之差。

在地球物理问题中, A 为 $M \times N$ 矩阵, B 为 M 维向量, X 为 N 维向量, 并且通常 $M > N$, 为一矛盾方程组。为了能用所学过的解方程方法来求解方程组, 必须对方程组 (1—1.15) 作适当变换, 下面讨论这类方程组的经典解法。

二、法方程 (正则方程)

我们平常只能对行与列相等的方阵进行求解, 这时可用学过的高斯消元法等任一方法来求解。方程 (1—1.15) 是行多于列, 即方程个数多于变量个数, 通常的作法是将方程 (1—1.15) 两端同时左乘 A^T , 得到

$$A^T A \cdot X = A^T B \quad (1-2.1)$$

即

$$P_{n \times n} X_{n \times 1} = q_{n \times 1} \quad (1-2.2)$$

式中, $P = A^T A$, $q = A^T B$ 。方程 $PX = q$ 被称为法方程或正则方程 (normal equation)。注意到 P 矩阵只取决于基本函数, 而不涉及实测值。由于 P 是对称方阵, 因而存储可以减半, 而且可以证明 P 矩阵是正定的, 不需作特别处理。因此, 采用供正定对称矩阵用的高斯法的变种, 只需花一半的工作量就可以求解。

然而, 使用正则方程求解有一个严重的缺陷, P 矩阵常常有很大的条件数, 因此, 无论用什么方法来解法方程, 数据的误差和计算过程中引入的舍入误差, 在计算法方程系数过程中都会大大地被放大, 继而造成求得的解误差很大。

三、普通逆矩阵的求法

方程 (1—2.2) 的求解可以用下面的方法。

$$X = P^{-1}q \quad (1-3.1)$$

式中, P^{-1} 称为 P 矩阵的逆矩阵, 只要行列式 $|P|$ 不等于零, 此方程就可以解出。

下面介绍有关逆矩阵方面的知识。

定义: 设有一方阵 A , 如果存在另一方阵 B , 且满足关系式

$$AB = BA = I$$

则方阵 B 叫做 A 的逆矩阵, 记为 A^{-1} , I 是和 A 、 B 同阶的单位方阵。当然, 也可以把方阵 A 叫做 B 的逆矩阵, 记作 B^{-1} 。

逆矩阵具以下性质

$$1. \quad (A^{-1})^{-1} = A \quad (1-3.2)$$

这就是说对矩阵 A , 通过两次取逆, 即还原矩阵 A

$$2. \quad (AB)^{-1} = B^{-1}A^{-1} \quad (1-3.3)$$

这实际上是矩阵乘积求逆的运算法则

$$3. \quad (A^{-1})^T = (A^T)^{-1} \quad (1-3.4)$$

即逆矩阵的转置等于转置矩阵的逆矩阵。

4. 如果 A 为对角方阵, 且对角线元素 $a_{11}, a_{22}, \dots, a_{nn}$ 均不为零, 即

$$A = \begin{bmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{bmatrix}$$

则 A 的逆矩阵 A^{-1} 为

$$A^{-1} = \begin{bmatrix} \frac{1}{a_{11}} & & & \\ & \frac{1}{a_{22}} & & \\ & & \ddots & \\ & & & \frac{1}{a_{nn}} \end{bmatrix} \quad (1-3.5)$$

逆矩阵的这些基本性质，以后要经常用到。

下面介绍逆矩阵的求法。

定理：设有一 n 阶方阵 A

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

如果满秩的话，即 $\text{rank}(A) = n$ ，则一定存在它的逆矩阵 A^{-1} ，且有

$$A^{-1} = \frac{1}{|A|} \cdot \begin{bmatrix} |A_{11}| & |A_{21}| & \cdots & |A_{n1}| \\ |A_{12}| & |A_{22}| & \cdots & |A_{n2}| \\ \cdots & \cdots & \cdots & \cdots \\ |A_{1n}| & |A_{2n}| & \cdots & |A_{nn}| \end{bmatrix} \quad (1-3.6)$$

式中，符号 $|A|$ 表示由矩阵元素组成的 n 阶行列式。

$|A_{ij}|$ 表示行列式 $|A|$ 中元素 a_{ij} 的代数余子式。我们把由这些代数余子式按(1-3.6)表中排列成的矩阵，称为 A 的伴

随矩阵，并表示成 $\text{adj}A$ (adj 是英文adjoint的缩写)，即

$$\text{adj}A = \begin{bmatrix} |A_{11}| & |A_{21}| & \cdots & |A_{n1}| \\ |A_{12}| & |A_{22}| & \cdots & |A_{n2}| \\ \cdots & \cdots & \cdots & \cdots \\ |A_{1n}| & |A_{2n}| & \cdots & |A_{nn}| \end{bmatrix}$$

也就是说， A 的伴随矩阵是以 $|A|$ 的代数余子式为元素所组成的矩阵的转置矩阵。

有了伴随矩阵的概念，矩阵 A 的逆矩阵 A^{-1} 可以简单地表示成

$$A^{-1} = \frac{1}{|A|} \text{adj}A$$

例：已知方阵

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & -2 & 0 \\ 2 & 1 & 1 \end{bmatrix}$$

求逆矩阵 A^{-1} 。

解：方法1. 先计算矩阵 A 的行列式之值

$$|A| = \begin{vmatrix} 1 & 2 & 1 \\ 0 & -2 & 0 \\ 2 & 1 & 1 \end{vmatrix} = 2$$

A 的伴随矩阵 $\text{adj}A$ 为

$$\text{adj}A = \begin{bmatrix} \begin{vmatrix} -2 & 0 \\ 1 & 1 \end{vmatrix} & -\begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} & \begin{vmatrix} 2 & 1 \\ -2 & 0 \end{vmatrix} \\ -\begin{vmatrix} 0 & 0 \\ 2 & 1 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 2 & 1 \end{vmatrix} & -\begin{vmatrix} 1 & 1 \\ 2 & 1 \end{vmatrix} \\ \begin{vmatrix} 0 & -2 \\ 2 & 1 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 0 & -2 \end{vmatrix} \end{bmatrix}$$

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

$$A^{-1} = \frac{1}{|A|} \cdot \text{adj} A = \begin{bmatrix} -1 & -\frac{1}{2} & 1 \\ 0 & -\frac{1}{2} & 0 \\ 2 & \frac{3}{2} & -1 \end{bmatrix}$$

检验所求 A^{-1} 是否正确

$$AA^{-1} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & -2 & 0 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} -1 & -\frac{1}{2} & 1 \\ 0 & -\frac{1}{2} & 0 \\ 2 & \frac{3}{2} & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

由此看出, 所求 A^{-1} 的确为 A 的逆矩阵, 因为满足逆矩阵的定义。

方法表明, 这一节的定理实际上也给出了求逆矩阵的方法。但是, 在实际应用中这种方法很不方便, 如果矩阵的阶数很高时, 这种求逆方法甚至是行不通的。下面介绍一种适用于计算机求逆的方法——矩阵分块法求矩阵。

方法2. 分块矩阵求原矩阵的逆

假设方阵 A 可按下面的方式进行分块

$$(A)_{n \times n} = \begin{bmatrix} [A_{11}]_{k \times k} & [A_{12}]_{k \times (n-k)} \\ [A_{21}]_{(n-k) \times k} & [A_{22}]_{(n-k) \times (n-k)} \end{bmatrix} \quad (1-3.7)$$

再设所求的逆矩阵 $A^{-1}=B$, 也按上述方法分成块, 即

$$[B]_{n \times n} = \begin{bmatrix} [B_{11}]_{k \times k} & [B_{12}]_{k \times (n-k)} \\ [B_{21}]_{(n-k) \times k} & [B_{22}]_{(n-k) \times (n-k)} \end{bmatrix} \quad (1-3.8)$$

根据逆矩阵的定义

$$\begin{aligned} & [A]_{n \times n} [B]_{n \times n} \\ &= \begin{bmatrix} [A_{11}]_{k \times k} & [A_{12}]_{k \times (n-k)} \\ [A_{21}]_{(n-k) \times k} & [A_{22}]_{(n-k) \times (n-k)} \end{bmatrix} \\ & \cdot \begin{bmatrix} [B_{11}]_{k \times k} & [B_{12}]_{k \times (n-k)} \\ [B_{21}]_{(n-k) \times k} & [B_{22}]_{(n-k) \times (n-k)} \end{bmatrix} \\ &= \begin{bmatrix} [I]_{k \times k} & \\ & [I]_{(n-k) \times (n-k)} \end{bmatrix} \quad (1-3.9) \end{aligned}$$

由乘法规则, 并考虑到等式两端矩阵对应元素应相等, 则有

$$A_{11}B_{11} + A_{12}B_{21} = I \quad (1-3.10)$$

$$A_{11}B_{12} + A_{12}B_{22} = 0 \quad (1-3.11)$$

$$A_{21}B_{11} + A_{22}B_{21} = 0 \quad (1-3.12)$$

$$A_{21}B_{12} + A_{22}B_{22} = I \quad (1-3.13)$$

如果 A_{11} 的逆矩阵 A_{11}^{-1} 已经求得, 则可从以上四个式子求出逆矩阵 B 的各元素

由 (1-3.11) 式

$$B_{12} = -A_{11}^{-1}A_{12}B_{22} \quad (1-3.14)$$

将 (1-3.14) 式代入 (1-3.13) 式, 有

$$B_{22} = (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} \quad (1-3.15)$$

由 (1-3.10) 式得

$$B_{11} = A_{11}^{-1} - A_{11}^{-1}A_{12}B_{21} \quad (1-3.16)$$

将 (1-3.16) 式代入 (1-3.12) 得