

# 超定线性系统最小二乘解的可信验证

于 茜, 李 喆\*

长春理工大学理学院, 吉林 长春

Email: 1131663641@qq.com, zheli200809@163.com

收稿日期: 2021年4月21日; 录用日期: 2021年5月7日; 发布日期: 2021年5月26日

---

## 摘 要

给定一个系数矩阵为病态矩阵的超定线性系统, 本文利用区间算法和增广矩阵技术, 设计算法输出给定超定线性系统的微小扰动区间系统, 算法保证在输出的区间系统中存在系数阵精确秩亏的超定线性系统, 算法输出该系数阵精确秩亏的超定线性系统的最小二乘解的高精度近似解和其相应的可信误差界。

## 关键词

可信验证, 最小二乘解, 超定线性系统

---

# The Verification of the Least Square Solution of an Overdetermined Linear System

Xi Yu, Zhe Li\*

School of Science, Changchun University of Science and Technology, Changchun Jilin

Email: 1131663641@qq.com, zheli200809@163.com

Received: Apr. 21<sup>st</sup>, 2021; accepted: May 7<sup>th</sup>, 2021; published: May 26<sup>th</sup>, 2021

---

## Abstract

Given an overdetermined linear system with ill-conditioned coefficient matrix, this paper mainly uses interval algorithm and augmented matrix technology to design an algorithm to output a minimally perturbed interval system for the given overdetermined linear system. The algorithm guarantees that there is an overdetermined linear system with a coefficient matrix of exactly rank deficiency in the output interval system. The algorithm also outputs the high-precision approximate solution of the least square solution of the overdetermined linear system and corresponding verified error bound.

\*通讯作者。

## Keywords

### Verification, Least Square Solution, Overdetermined Linear System

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

线性系统是计算科学领域中一种很重要的数学模型。线性系统理论在航空航天、化学化工、机械研究等技术领域中都有着广泛的应用实例。超定线性系统是指方程个数大于变量个数的线性系统。研究表明, 在数字信号处理和其他面向定量的学科中有许多问题最终都归结为超定线性系统求解的计算问题, 最小二乘拟合算法的实质就是对超定线性系统进行求解。

多年来, 对于超定线性系统求解问题的研究始终吸引着国内外的专家学者。Levitani 等人[1]研究了针对特定抽象范数求解超定线性系统最优解的解决方案。Barroda 和 Roberts [2]提出了能够解决线性数据拟合问题的有效算法, 该算法是对线性规划单纯形法的一种改进, 适用于  $l_1$  问题的初步表述。Cadzow [3]提出了超定线性系统最小 1、2 和  $\infty$  范数近似解的解决方案。Demmel 等人[4]研究了超定最小二乘问题精确迭代细化的算法、误差范围和数值结果, 将线性系统优化算法应用于超定线性系统最小二乘问题的增强线性系统公式。Novoselac 和 Pavic [5]研究了  $p$  范数中超定线性方程组的解性质, 从而得到了该问题具有回归、尺度和仿射等变性质。Popa [6]提出了超定线性系统最小二乘解的数值计算的两种迭代算法, 这些算法在 Kaczmarz 迭代算法的基础上, 可以收敛到最小二乘解的序列。Pelckmans 等人[7]对超定线性系统的噪声观测进行了稀疏估计, 提出了一种受高斯噪声干扰的线性方程估计未知参数的方法。Zhijun Zhang 等人[8]为解决时变的超定问题, 提出并设计了一种新颖的变参数收敛微分神经网络, 通过研究表明, 这种神经网络具有超指数收敛速度的最小二乘解。

近年来, 很多学者着手研究超定系统最小二乘解的可信验证。Rump [9]提出了一种不依赖于正规方程且适用于稀疏矩阵的新算法, 该算法用于计算系数阵列满秩的最小二乘问题的最小二乘解和欠定线性系统的最小二范数解。Miyajima 和 Shinya [10]在考虑了所有可能的舍入误差的情况下, 给出了系数阵为列满秩的欠定系统最小二范数解和超定系统最小二乘解的验证算法, 阐明了所获得的误差范围的大小与系数矩阵的条件数相关。Rump [11]提出的快速算法的误差界是标量, 而 Miyajima 提出的算法的误差界是向量。Miyajima [12]总结提出了一种用于矩阵方程  $AXB = C$  的最小二乘解的高效能算法。

本文考虑系数矩阵为病态矩阵的超定线性系统的可信计算。计算给定超定线性系统的微小摄动区间系统, 算法保证在输出的区间系统中存在系数阵精确秩亏的超定线性系统, 并计算该系数阵精确秩亏的超定线性系统的最小二乘解的可信区间解。

## 2. 符号和预备知识

令  $\mathbb{R}$  表示实数集合。设  $A \in \mathbb{R}^{m \times n}$ ,  $A(\cdot)$  表示由矩阵  $A$  的列合并得到的长列向量。对于满足条件  $1 \leq i_1 \leq i_2 \leq m$ ,  $1 \leq j_1 \leq j_2 \leq n$  的正整数  $i_1, i_2, j_1, j_2$ ,  $A_{i_1:i_2, \cdot}$  表示由矩阵  $A$  的第  $i_1$  行到第  $i_2$  行所构成的子矩阵,  $A_{\cdot, j_1:j_2}$  表示由矩阵  $A$  的第  $j_1$  列到第  $j_2$  列所构成的子矩阵,  $A_{i_1:i_2, j_1:j_2}$  表示由矩阵  $A$  的第  $i_1$  行到第  $i_2$  行, 第  $j_1$  列到第  $j_2$  列所构成的子矩阵。令  $O_{m,n}$  表示  $m \times n$  阶零矩阵,  $I_{m \times m}$  表示  $m \times m$  阶单位阵。

令  $\mathbb{IR}$  表示区间集合  $\{[x, \bar{x}] : x, \bar{x} \in \mathbb{R}, x \leq \bar{x}\}$ , 用  $mid[x, \bar{x}]$  表示区间  $[x, \bar{x}]$  的中点, 用  $length$  表示区间长度, 即  $length([x, \hat{x}]) = \bar{x} - x$ 。区间向量  $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n)^T \in \mathbb{IR}^n$  是分量为区间的  $n$  维向量, 令  $int(\mathbf{X})$  表示它的内部。区间矩阵  $\mathbf{A}$  是分量为区间的矩阵。给定区间矩阵  $\mathbf{A} \in \mathbb{IR}^{m \times n}$ , 若区间矩阵  $\mathbf{A}$  包含的所有实矩阵都是非奇异的, 则称区间矩阵  $\mathbf{A}$  非奇异。给定区间矩阵  $\mathbf{A} \in \mathbb{IR}^{m \times n}$ , 若区间矩阵  $\mathbf{A}$  包含的所有实矩阵都是满秩的, 则称区间矩阵  $\mathbf{A}$  满秩。对于  $\mathbf{A} \in \mathbb{IR}^{m \times n}$ ,  $mid(\mathbf{A})$  是由矩阵  $\mathbf{A}$  中每个分量中点对应的值构成的实矩阵, 令  $length(\mathbf{A}) = \max\{length(\mathbf{A}_{i,j}), 1 \leq i \leq m, 1 \leq j \leq n\}$ 。

**定义 1:** 对于给定的阈值  $\delta > 0$ , 若所有与  $A$  的 2-范数距离不超过  $\delta$  的  $m \times n$  阶矩阵的最小秩为  $r$ , 即

$$r = \min_{\substack{\|A-B\|_2 \leq \delta \\ B \in \mathbb{R}^{m \times n}}} \{rank(B)\}$$

则称  $A$  的数值  $\delta$ -秩为  $r$ , 记为  $rank_\delta(A) = r$ 。

**注 1:** 给定矩阵  $A \in \mathbb{R}^{m \times n}$ , 其中  $m > n$ , 对于阈值  $\delta > 0$ , 如果矩阵  $A$  的奇异值  $\sigma_1(A), \dots, \sigma_n(A)$  满足  $\sigma_1(A) \geq \dots \geq \sigma_r(A) > \delta \geq \sigma_{r+1}(A) \geq \dots \geq \sigma_n(A)$ ,

则  $rank_\delta(A) = r$ 。

**引理 1:** 给定矩阵  $A, B \in \mathbb{R}^{n \times n}$ , 若  $\|I_n - BA\|_p \leq \alpha < 1$ , 其中  $p \in \{1, 2, \dots, \infty\}$ , 则  $A, B$  均是非奇异矩阵。

**定理 1:** 对于给定的区间矩阵  $\mathbf{A} \in \mathbb{IR}^{n \times n}$  和区间向量  $\mathbf{b} \in \mathbb{IR}^n$ , 向量  $\mathbf{b}$  满足  $\mathbf{0} \notin \mathbf{b}$ , 若 INTLAB 工具箱中 `verifylss` 函数成功输出区间向量  $\mathbf{X} \subset \mathbb{IR}^n$ , 则区间矩阵  $\mathbf{A}$  非奇异, 且  $\mathbf{X}$  满足条件

$$\{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{b} \quad \forall A \in \mathbf{A}, \forall \mathbf{b} \in \mathbf{b}\} \subseteq \mathbf{X} \quad (1)$$

**定理 2:** 设矩阵  $\mathbf{A} \in \mathbb{IR}^{m \times n}$ ,  $m > n$ , 向量  $\mathbf{b} \in \mathbb{IR}^m$  满足  $\mathbf{0} \notin \mathbf{b}$ , 当输入区间矩阵  $\begin{pmatrix} \mathbf{A} & I_m \\ O_{n,n} & \mathbf{A}^T \end{pmatrix}$  和区间向量  $(\mathbf{b}, O_{n,1})^T$  时, 若 INTLAB 工具箱中 `verifylss` 函数成功地输出区间向量, 则区间矩阵  $\mathbf{A}$  满秩。

**定理 3:** 设  $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$  并且  $f = (f_1, f_2, \dots, f_n)^T$ , 其中  $f_1, \dots, f_n$  为连续可微函数。当  $\tilde{\mathbf{x}} \in \mathbb{R}^n$  且  $\mathbf{X} \in \mathbb{IR}^n$ ,  $\mathbf{0} \in \mathbf{X}$  时, 令  $f'(\tilde{\mathbf{x}} + \mathbf{X})$  表示系统在区间  $\tilde{\mathbf{x}} + \mathbf{X}$  上的雅可比矩阵。设区间矩阵  $\mathbf{M} \in \mathbb{IR}^{n \times n}$  满足条件  $f'(\tilde{\mathbf{x}} + \mathbf{X}) \subseteq \mathbf{M}$ 。若对矩阵  $R \in \mathbb{R}^{n \times n}$ , 有

$$-Rf(\tilde{\mathbf{x}}) + (I_n - RM)\mathbf{X} \subseteq int(\mathbf{X}), \quad (2)$$

则存在唯一的  $\hat{\mathbf{x}} \in \tilde{\mathbf{x}} + \mathbf{X}$  使得  $f(\hat{\mathbf{x}}) = \mathbf{0}$ 。

**注 2:** INTLAB 工具箱中 `verifylss` 函数利用定理 3 来验证非线性系统的解。

### 3. 主要结论

设  $A \in \mathbb{R}^{m \times n}$ , 其中  $m > n$ , 令  $A = U\Sigma V^T$  表示  $A$  的奇异值分解。若  $rank_\delta(A) = r$ , 则  $U_{:,r+1:m}$  和  $V_{:,r+1:m}$  的列向量分别张成矩阵  $A$  的近似  $\delta$ -左零空间和近似  $\delta$ -右零空间。定义变量矩阵  $E$

$$E = \begin{pmatrix} e_{1,1} & e_{1,2} & \dots & e_{1,n} \\ e_{2,1} & e_{2,2} & \dots & e_{2,n} \\ \dots & \dots & \dots & \dots \\ e_{m,1} & e_{m,2} & \dots & e_{m,n} \end{pmatrix} \quad (3)$$

为了简便起见, 用符号  $\mathbf{e}$  表示向量  $E(\cdot)$ , 用符号  $A(\mathbf{e})$  表示矩阵  $A + E$ 。

**引理 2:** 令  $A \in \mathbb{R}^{m \times n}$ , 其中  $m > n$ , 且  $rank_\delta(A) = r$ , 则矩阵

$$\begin{pmatrix} A(\mathbf{e}) & U_{:,r+1:m} \\ V_{:,r+1:n}^T & O_{n-r,n-r} \end{pmatrix}$$

当  $\mathbf{e} = \mathbf{0}$  时非奇异。

**引理 3:** 令  $A \in \mathbb{R}^{m \times n}$ ,  $m > n$ , 且  $\text{rank}_\delta(A) = r$ , 则矩阵  $H(\mathbf{e})$

$$H(\mathbf{e}) = \begin{pmatrix} A(\mathbf{e}) & U_{:,r+1:n} & -I_m & O_{m,n-r} \\ V_{:,r+1:n}^T & O_{n-r,n-r} & O_{n-r,m} & -I_{n-r} \\ O_{n,n} & O_{n,n-r} & A(\mathbf{e})^T & V_{:,r+1:n}^T \\ O_{n-r,n} & O_{n-r,n-r} & U_{:,r+1:n}^T & O_{n-r,n-r} \end{pmatrix} \quad (4)$$

当  $\mathbf{e} = \mathbf{0}$  时非奇异。

引入如下线性系统

$$\begin{pmatrix} A(\mathbf{e}) & U_{:,r+1:n} & -I_m & O_{m,n-r} \\ V_{:,r+1:n}^T & O_{n-r,n-r} & O_{n-r,m} & -I_{n-r} \\ O_{n,n} & O_{n,n-r} & A(\mathbf{e})^T & V_{:,r+1:n}^T \\ O_{n-r,n} & O_{n-r,n-r} & U_{:,r+1:n}^T & O_{n-r,n-r} \end{pmatrix} \begin{pmatrix} C(\mathbf{e}) \\ F(\mathbf{e}) \\ Y(\mathbf{e}) \\ Z(\mathbf{e}) \end{pmatrix} = \begin{pmatrix} O_{m,n-r} \\ I_{n-r} \\ O_{n,n-r} \\ O_{n-r,n-r} \end{pmatrix}. \quad (5)$$

其中矩阵  $C(\mathbf{e})$ ,  $F(\mathbf{e})$ ,  $Y(\mathbf{e})$  和  $Z(\mathbf{e})$  的行维数分别是  $n$ ,  $n-r$ ,  $m$  和  $n-r$ 。

定义矩阵

$$\tilde{A}(\mathbf{e}) = A(\mathbf{e}) - Y(\mathbf{e})(C^T(\mathbf{e})C(\mathbf{e}))^{-1}C^T(\mathbf{e}). \quad (6)$$

**命题 1:** 假设向量  $\hat{\mathbf{e}} \in \mathbb{R}^{n^2}$  满足如下条件:

- (1)  $\|\hat{\mathbf{e}}\|_\infty < \frac{1}{(m+n)\sqrt{m+n}\|H(\mathbf{0})^{-1}\|_2}$ ,
- (2)  $F(\hat{\mathbf{e}}) = O_{n-r,n-r}$ ,
- (3) 矩阵  $Z(\hat{\mathbf{e}}) + I_{n-r}$  非奇异,
- (4) 矩阵

$$\begin{pmatrix} \tilde{A}(\hat{\mathbf{e}}) \\ C(\hat{\mathbf{e}})^T \end{pmatrix}. \quad (7)$$

满秩, 则  $\text{rank}(\tilde{A}(\hat{\mathbf{e}})) = r$ 。

**证明:** 易知

$$\begin{aligned} \|I_{m+n} - H(\mathbf{0})^{-1}H(\hat{\mathbf{e}})\|_2 &\leq \|H(\mathbf{0})^{-1}\|_2 \|H(\mathbf{0}) - H(\hat{\mathbf{e}})\|_2 \\ &\leq \sqrt{m+n} \|H(\mathbf{0})^{-1}\|_2 \|H(\mathbf{0}) - H(\hat{\mathbf{e}})\|_\infty \\ &\leq (m+n)\sqrt{m+n} \|H(\mathbf{0})^{-1}\|_2 \|\hat{\mathbf{e}}\|_\infty < 1. \end{aligned}$$

故矩阵  $H(\hat{\mathbf{e}})$  非奇异。若  $F(\hat{\mathbf{e}}) = O_{n-r,n-r}$ , 则  $A(\hat{\mathbf{e}})C(\hat{\mathbf{e}}) = Y(\hat{\mathbf{e}})$ 。由(6)式可得  $\tilde{A}(\hat{\mathbf{e}})C(\hat{\mathbf{e}}) = O_{n-r,m}$ , 断言  $C(\hat{\mathbf{e}})$  满秩。事实上, 如果  $C(\hat{\mathbf{e}})$  的列向量线性相关, 则存在一个非零向量  $\boldsymbol{\alpha} \in \mathbb{R}^{m-r}$  使得  $C(\hat{\mathbf{e}})\boldsymbol{\alpha} = \mathbf{0}$ 。由于  $V_{:,r+1:n}^T C(\hat{\mathbf{e}}) - Z(\hat{\mathbf{e}}) = I_{n-r}$ , 所以  $Z(\hat{\mathbf{e}})\boldsymbol{\alpha} + \boldsymbol{\alpha} = \mathbf{0}$ , 这与矩阵  $Z(\hat{\mathbf{e}}) + I_{n-r}$  非奇异相矛盾。假设  $\text{rank}(\tilde{A}(\hat{\mathbf{e}})) < r$ , 则存在一个向量  $\boldsymbol{\beta} \in \text{null}(\tilde{A}(\hat{\mathbf{e}})^T)$  使得矩阵  $(C(\hat{\mathbf{e}})\boldsymbol{\beta})$  满秩。定义非零向量

$$\boldsymbol{\gamma} = (C(\hat{\boldsymbol{e}}) \boldsymbol{\beta}) \begin{pmatrix} (C(\hat{\boldsymbol{e}})^T C(\hat{\boldsymbol{e}}))^{-1} C(\hat{\boldsymbol{e}})^T \boldsymbol{\beta} \\ -1 \end{pmatrix},$$

通过计算可知

$$\begin{pmatrix} \tilde{A}(\hat{\boldsymbol{e}}) \\ C(\hat{\boldsymbol{e}})^T \end{pmatrix} \boldsymbol{\gamma} = 0.$$

上式与矩阵(7)满秩相矛盾。 □

由引理 3 可知, 存在零向量的一个邻域, 使得对这个邻域内的任何向量  $\boldsymbol{e}$ , 矩阵  $H(\boldsymbol{e})$  均是非奇异的。因此, 在这个邻域内, 系统(5)的解的每一项对每个变量均可导。对系统(5)的两边关于变量求导, 可得如下的线性系统

$$H(\boldsymbol{e}) \begin{pmatrix} \frac{\partial C(\boldsymbol{e})}{\partial e_{i,j}} \\ \frac{\partial F(\boldsymbol{e})}{\partial e_{i,j}} \\ \frac{\partial Y(\boldsymbol{e})}{\partial e_{i,j}} \\ \frac{\partial Z(\boldsymbol{e})}{\partial e_{i,j}} \end{pmatrix} = \begin{pmatrix} O_{i-1,n-r} \\ -C(\boldsymbol{e})_{j,:} \\ O_{m+n-r-i+j-1,n-r} \\ -Y(\boldsymbol{e})_{i,:} \\ O_{2n-r-j,n-r} \end{pmatrix}, \quad 1 \leq i \leq m, 1 \leq j \leq n, \quad (8)$$

$$H(\boldsymbol{e}) \begin{pmatrix} \frac{\partial^2 C(\boldsymbol{e})}{\partial e_{i_1,j_1} \partial e_{i_2,j_2}} \\ \frac{\partial^2 F(\boldsymbol{e})}{\partial e_{i_1,j_1} \partial e_{i_2,j_2}} \\ \frac{\partial^2 Y(\boldsymbol{e})}{\partial e_{i_1,j_1} \partial e_{i_2,j_2}} \\ \frac{\partial^2 Z(\boldsymbol{e})}{\partial e_{i_1,j_1} \partial e_{i_2,j_2}} \end{pmatrix} = \begin{pmatrix} O_{i_2-1,n-r} \\ \left( \frac{\partial C(\boldsymbol{e})}{\partial e_{i_1,j_1}} \right)_{j_2,:} \\ O_{m+n-r-i_2+j_2-1,n-r} \\ \left( \frac{\partial Y(\boldsymbol{e})}{\partial e_{i_1,j_1}} \right)_{i_2,:} \\ O_{2n-r-j_2,n-r} \end{pmatrix} + \begin{pmatrix} O_{i_1-1,n-r} \\ \left( \frac{\partial C(\boldsymbol{e})}{\partial e_{i_2,j_2}} \right)_{j_1,:} \\ O_{m+n-r-i_1+j_1-1,n-r} \\ \left( \frac{\partial Y(\boldsymbol{e})}{\partial e_{i_2,j_2}} \right)_{i_1,:} \\ O_{2n-r-j_1,n-r} \end{pmatrix}, \quad 1 \leq i \leq m, 1 \leq j \leq n. \quad (9)$$

考虑如下约束优化问题

$$\begin{aligned} \min \quad & \frac{1}{2} \|\boldsymbol{e}\|_2^2 \\ \text{s.t.} \quad & F(\boldsymbol{e}) = O_{n-r,n-r} \end{aligned} \quad (10)$$

引入拉格朗日乘子

$$\boldsymbol{\lambda} = (\lambda_{1,1}, \lambda_{1,2}, \dots, \lambda_{n-r,1}, \lambda_{n-r,2}, \dots, \lambda_{n-r,2}, \dots, \lambda_{1,n-r}, \lambda_{2,n-r}, \dots, \lambda_{n-r,n-r})^T$$

定义拉格朗日函数

$$g(\boldsymbol{e}, \boldsymbol{\lambda}) = \frac{1}{2} \|\boldsymbol{e}\|_2^2 + \sum_{j=1}^{n-r} \sum_{i=1}^{n-r} \lambda_{i,j} F(\boldsymbol{e})_{i,j}. \quad (11)$$

经计算可知  $g(\boldsymbol{e}, \boldsymbol{\lambda})$  的梯度为

$$g'(\mathbf{e}, \boldsymbol{\lambda}) = \begin{pmatrix} \mathbf{e} + \sum_{i=1}^{n-r} \sum_{j=1}^{n-r} \lambda_{i,j} F'(\mathbf{e})_{i,j} \\ F(\mathbf{e}) \end{pmatrix}. \quad (12)$$

$g(\mathbf{e}, \boldsymbol{\lambda})$  的海森矩阵为

$$g''(\mathbf{e}, \boldsymbol{\lambda}) = \begin{pmatrix} I_{mm} + \sum_{i=1}^{n-r} \sum_{j=1}^{n-r} \lambda_{i,j} F''(\mathbf{e})_{i,j} & \mathbf{F}'(\mathbf{e})(\cdot)^T \\ \mathbf{F}'(\mathbf{e})(\cdot) & O_{(n-r)^2, (n-r)^2} \end{pmatrix}. \quad (13)$$

**注 3:** 数值部分利用牛顿迭代法

$$\begin{pmatrix} \tilde{\mathbf{e}} \\ \tilde{\boldsymbol{\lambda}} \end{pmatrix} \leftarrow \begin{pmatrix} \tilde{\mathbf{e}} \\ \tilde{\boldsymbol{\lambda}} \end{pmatrix} - g''(\tilde{\mathbf{e}}, \tilde{\boldsymbol{\lambda}})^{-1} g'(\tilde{\mathbf{e}}, \tilde{\boldsymbol{\lambda}}) \quad (14)$$

计算  $g'(\mathbf{e}, \boldsymbol{\lambda}) = 0$  的高精度近似解  $(\tilde{\mathbf{e}}, \tilde{\boldsymbol{\lambda}})$ , 其中  $g'(\tilde{\mathbf{e}}, \tilde{\boldsymbol{\lambda}})$ ,  $g''(\tilde{\mathbf{e}}, \tilde{\boldsymbol{\lambda}})$  可由求系统(8) (9)当  $\mathbf{e} = \tilde{\mathbf{e}}$  时的解得到。

**注 4:** 验证部分利用区间算法计算系统  $g'(\mathbf{e}, \boldsymbol{\lambda}) = \mathbf{0}$  近似解的可信误差界。对于任何  $\tilde{\mathbf{E}} \in \mathbb{IR}^{mm}$ ,  $\tilde{\boldsymbol{\Lambda}} \in \mathbb{IR}^{(n-r)(n-r)}$ , 利用 `verifylss` 函数求解当  $\mathbf{e} = \tilde{\mathbf{E}}$  时线性方程组(5) (8) (9)的解, 从而得到  $g'(\tilde{\mathbf{E}}, \tilde{\boldsymbol{\Lambda}})$  和  $g''(\tilde{\mathbf{E}}, \tilde{\boldsymbol{\Lambda}})$ 。利用定理 3 和 `verifylss` 函数的代码, 验证部分计算区间向量  $\hat{\mathbf{E}}$  和  $\hat{\boldsymbol{\Lambda}}$ , 其分别包含实向量  $\hat{\mathbf{e}}$  和  $\hat{\boldsymbol{\lambda}}$ , 使得  $g'(\hat{\mathbf{e}}, \hat{\boldsymbol{\lambda}}) = \mathbf{0}$ 。

#### 算法 1

输入  $A \in \mathbb{R}^{m \times n}$ ,  $m > n$ ,  $\mathbf{b} \in \mathbb{R}^m$ ,  $\delta \in \mathbb{R}^+$ ;

输出  $r \in \mathbb{N}$ ,  $\tilde{A}(\hat{\mathbf{e}}) \in \mathbb{IR}^{m \times n}$  或者 “Failure”。

步骤 1 计算满足条件(1)的正数  $r$ 。

步骤 2 计算系统(11)的近似解  $(\tilde{\mathbf{e}}, \tilde{\boldsymbol{\lambda}})$ 。

步骤 3 计算包含系统(11)精确解的区间向量  $(\hat{\mathbf{E}}, \hat{\boldsymbol{\Lambda}})$ 。

步骤 4 利用(6)式计算区间矩阵  $\tilde{A}(\hat{\mathbf{E}})$ 。

步骤 5 利用定理 2 验证矩阵  $Z(\hat{\mathbf{E}}) + I_{n-r}$  是否非奇异, 如果失败, 则返回 “Failure”。

步骤 6 利用定理 3 验证矩阵  $\begin{pmatrix} \tilde{A}(\hat{\mathbf{E}}) \\ C(\hat{\mathbf{E}})^T \end{pmatrix}$  是否秩亏, 如果失败, 则返回 “Failure”。 □

若算法 1 成功地输出区间向量  $\hat{\mathbf{E}}$ , 则由命题 5 可知, 存在  $\hat{\mathbf{e}} \in \hat{\mathbf{E}}$ , 使得  $\text{rank}(\tilde{A}(\hat{\mathbf{e}})) = r$ 。因此, 存在满足条件  $1 \leq i_1 < i_2 < \dots < i_r \leq n$  的正整数序列  $i_1, i_2, \dots, i_r$ , 使得区间矩阵  $A(\hat{\mathbf{E}})_{:,i_1, i_2, \dots, i_r}$  满秩, 则考虑如下区间线性系统

$$\begin{pmatrix} \tilde{A}(\hat{\mathbf{E}})_{:,i_1, i_2, \dots, i_r} & -I_m \\ O_{r,r} & \tilde{A}(\hat{\mathbf{E}})^T_{:,i_1, i_2, \dots, i_r} \end{pmatrix} \begin{pmatrix} \mathbf{y}(\hat{\mathbf{E}}) \\ \boldsymbol{\omega}(\hat{\mathbf{E}}) \end{pmatrix} = \begin{pmatrix} \mathbf{b} \\ \mathbf{0} \end{pmatrix}. \quad (15)$$

**命题 2:** 设  $r$  维向量  $\mathbf{x}(\mathbf{e})$  向量满足如下条件

$$\mathbf{x}(\mathbf{e})_{:,k} = \begin{cases} \mathbf{y}(\mathbf{e})_{:,j}, & \text{当 } k = i_j, 1 \leq j \leq r; \\ 0, & \text{其他情况.} \end{cases} \quad (16)$$

若向量  $\hat{\mathbf{e}} \in R^{m \times n}$  满足  $\text{rank}(\tilde{A}(\hat{\mathbf{e}})) = r$ , 则  $\mathbf{x}(\hat{\mathbf{e}})$  是超定系统

$$\tilde{A}(\hat{e})x = b$$

的最小二乘解。

**证明:** 假设存在向量  $\tilde{x} \in \mathbb{R}^n$ , 使得

$$\|\tilde{A}(\hat{e})\tilde{x} - b\|_2 \leq \|\tilde{A}(\hat{e})x(\hat{e}) - b\|_2. \tag{17}$$

根据  $\text{rank}(\tilde{A}(\hat{e})) = r$ , 可知  $\hat{A}(\hat{e})_{i_1, i_2, \dots, i_r}$  满秩。因此, 存在实向量  $\alpha^{(1)}, \alpha^{(2)}, \dots, \alpha^{(n-r)} \in \mathbb{R}^r$ , 使得

$$\tilde{A}(\hat{e})_{j_1, j_2, \dots, j_{n-r}} = \tilde{A}(\hat{e})_{i_1, i_2, \dots, i_r} (\alpha^{(1)} \alpha^{(2)} \dots \alpha^{(n-r)}), \tag{18}$$

其中  $\{j_1, j_2, \dots, j_{n-r}\} = \{1, 2, \dots, n\} \setminus \{i_1, i_2, \dots, i_r\}$ 。显然, 区间系统(15)的解包含着下述系统

$$\begin{pmatrix} \tilde{A}(\hat{e})_{i_1, i_2, \dots, i_r} & -I_m \\ O_{r, r} & \tilde{A}(\hat{e})_{i_1, i_2, \dots, i_r}^T \end{pmatrix} \begin{pmatrix} y(\hat{e}) \\ \omega(\hat{e}) \end{pmatrix} = \begin{pmatrix} b \\ 0 \end{pmatrix}$$

的解。由(17) (18)式可知,

$$\|\tilde{A}(\hat{e})\tilde{x} - b\|_2 = \|\tilde{A}(\hat{e})_{i_1, i_2, \dots, i_r} ((\alpha^{(1)} \alpha^{(2)} \dots \alpha^{(n-r)})\tilde{x}_{j_1, j_2, \dots, j_r} + \tilde{x}_{i_1, i_2, \dots, i_r}) - b\|_2 \leq \|\tilde{A}(\hat{e})_{i_1, i_2, \dots, i_r} y(\hat{e}) - b\|_2,$$

而  $y(\hat{e})$  是系数矩阵  $\tilde{A}(\hat{e})_{i_1, i_2, \dots, i_r}$  和右端向量  $b$  的超定线性系统的唯一最小二乘解, 故

$$y(\hat{e}) = \tilde{x}_{i_1, i_2, \dots, i_r} + (\alpha^{(1)} \alpha^{(2)} \dots \alpha^{(n-r)})\tilde{x}_{j_1, j_2, \dots, j_{n-r}}.$$

因此  $\|\tilde{A}(\hat{e})\tilde{x} - b\|_2 = \|\tilde{A}(\hat{e})x(\hat{e}) - b\|_2$ . □

算法 1

输入  $A \in \mathbb{R}^{m \times n}$ ,  $m \leq n$ ;  $b \in \mathbb{R}^m$ ,  $\delta \in \mathbb{R}^+$ ;

输出  $r \in \mathbb{N}$ ,  $\hat{A} \in \mathbb{I}\mathbb{R}^{m \times n}$ ,  $\hat{X} \in \mathbb{I}\mathbb{R}^n$ 。

步骤 1 执行算法 1。

步骤 2 利用 `verifylss` 函数解线性方程组(15)计算  $y(\hat{E})$ ,  $\omega(\hat{E})$ 。

步骤 3 确定指标集合  $i_1, i_2, \dots, i_r$ , 使得  $\hat{A}(\hat{E})_{i_1, i_2, \dots, i_r}$  满秩。

步骤 4 利用(15) (16)式计算区间向量  $\hat{x}(\hat{E})$ 。

步骤 5 令  $\hat{A} \leftarrow A(\hat{E})$ ,  $\hat{X} \leftarrow \hat{x}(\hat{E})$ . □

由命题 1 和命题 2 可得如下定理。

**定理 4:** 若算法 1 成功输出  $r \in \mathbb{N}$ ,  $\hat{A} \in \mathbb{I}\mathbb{R}^{m \times n}$ ,  $\hat{X} \in \mathbb{I}\mathbb{R}^n$ , 则存在实矩阵  $\hat{A} \in \hat{A}$ , 实向量  $\hat{x} \in \hat{X}$  使得  $\hat{x}$  是超定线性系统  $\hat{A}x = b$  的最小二乘解, 且  $\text{rank}(\hat{A}) = r$ 。

### 4. 应用实例

**例 1** 给定超定线性系统  $Ax = b$ , 其中,

$$A = \begin{bmatrix} 0.6406 & -0.7615 & 0.0742 & -1.4399 & 0.5767 \\ -0.5122 & 0.9371 & -0.9758 & -0.0093 & -0.0319 \\ -0.4263 & 0.1856 & -0.2452 & 0.2856 & -0.4537 \\ 0.1684 & 0.0526 & -0.4056 & -0.8077 & 0.3922 \\ 0.4236 & -0.5423 & 0.3510 & -0.4947 & 0.2687 \\ -0.8300 & -0.4306 & -0.9047 & -1.0113 & -1.0731 \end{bmatrix}$$

$$\mathbf{b} = (2.0790 \quad -0.8294 \quad -0.6716 \quad 0.7956 \quad 1.0071 \quad 0.3704)^T$$

利用算法 1 可得  $\text{mid}(\mathbf{X}) = (-1.8809 \quad 0 \quad 0 \quad -0.9969 \quad 0 \quad 1.2501)^T$ ,

$$\text{mid}(\hat{\mathbf{A}}) = \begin{bmatrix} 0.6405 & -0.7614 & 0.0742 & -1.4398 & 0.5767 \\ -0.5122 & 0.9370 & -0.9757 & -0.0092 & -0.0319 \\ -0.2463 & 0.1855 & -0.2451 & 0.2855 & -0.4537 \\ 0.1683 & 0.0526 & -0.4055 & -0.8076 & 0.3922 \\ 0.4236 & -0.5423 & -0.9047 & -0.4947 & 0.2687 \\ -0.8299 & -0.4305 & -0.2765 & -1.0113 & -1.0731 \end{bmatrix}$$

$$\text{rad}(\hat{\mathbf{A}}) = 3.8858\text{e}-16, \quad \text{rad}(\hat{\mathbf{X}}) = 2.9976\text{e}-15, \quad \text{length}(A - \hat{\mathbf{A}}) = 2.7967\text{e}-05。$$

**Table 1.** The output results of Example 2

**表 1.** 例 2 的输出结果

(m, n, r)	(10, 5, 4)	(10, 9, 7)	(12, 8, 7)	(12, 10, 9)	(15, 10, 9)	(15, 14, 13)	(18, 16, 15)	(20, 15, 14)
$\text{rad}(\hat{\mathbf{A}})$	8.8818e-16	1.0270e-15	9.9920e-16	8.8818e-16	1.5543e-15	1.9984e-15	3.4972e-15	2.4424e-15
$\text{rad}(\hat{\mathbf{X}})$	1.2684e-14	2.7245e-13	3.7090e-12	1.1363e-13	9.5146e-13	1.1391e-13	4.6038e-13	6.2050e-12
$\text{length}(A - \hat{\mathbf{A}})$	0.0100	0.0172	2.9505e-05	8.4993e-06	0.0108	0.0198	0.0164	0.0089

**例 2** 系数矩阵为接近秩亏随机矩阵的超定线性系统  $\mathbf{Ax} = \mathbf{b}$ , 表给出了不同的  $(m, n, r)$  取值, 算法 1 的输出结果见表 1。

## 参考文献

- [1] Levitan, M.L. and Lynn, R.Y.S. (1976) An Overdetermined Linear System. *Journal of Approximation Theory*, **18**, 264-277. [https://doi.org/10.1016/0021-9045\(76\)90020-4](https://doi.org/10.1016/0021-9045(76)90020-4)
- [2] Barroda, I. and Roberts, F.D.K. (1974) Solution of an Overdetermined System of Equations in the  $l_1$  Norm [F<sub>4</sub>]. *Communications of the ACM*, **17**, 319. <https://doi.org/10.1145/355616.361024>
- [3] Cadzow, J.A. (2002) Minimum  $l_1$ ,  $l_2$ , and  $l_\infty$  Norm Approximate Solutions to an Overdetermined System of Linear Equations. *Digital Signal Processing*, **12**, 524-560. <https://doi.org/10.1006/dspr.2001.0409>
- [4] Demmel, J., Hida, Y., Riedy, E.J., et al. (2009) Extra-Precise Iterative Refinement for Overdetermined Least Squares Problems. *ACM Transactions on Mathematical Software*, **35**, 1-32. <https://doi.org/10.1145/1462173.1462177>
- [5] Novoselac, V. and Pavic, Z. (2019) Optimal Solution Properties of an Overdetermined System of Linear Equations. *European Journal of Pure and Applied Mathematics*, **12**, 1360-1370. <https://doi.org/10.29020/nybg.ejpam.v12i4.3517>
- [6] Popa, C. (1995) Least-Squares Solution of Overdetermined Inconsistent Linear Systems Using Kaczmarz's Relaxation. *International Journal of Computer Mathematics*, **55**, 79-89. <https://doi.org/10.1080/00207169508804364>
- [7] Dai, L. and Pelckmans, K. (2014) Sparse Estimation from Noisy Observations of an Overdetermined Linear System. *Automatica*, **50**, 2845-2851. <https://doi.org/10.1016/j.automatica.2014.08.018>
- [8] Zhang, Z., Zheng, L., Qiu, T., et al. (2020) Varying-Parameter Convergent-Differential Neural Solution to Time-Varying Overdetermined System of Linear Equations. *IEEE Transactions on Automatic Control*, **65**, 874-881. <https://doi.org/10.1109/TAC.2019.2921681>
- [9] Rump, S.M. (2012) Verified Bounds for Least Squares Problems and Underdetermined Linear Systems. *SIAM Journal on Matrix Analysis & Applications*, **33**, 130-148. <https://doi.org/10.1137/110840248>
- [10] Miyajima, S. (2014) Componentwise Enclosure for Solutions of Least Squares Problems and Underdetermined Systems. *Linear Algebra & Its Applications*, **444**, 28-41. <https://doi.org/10.1016/j.laa.2013.11.044>
- [11] Rump, S.M. (2014) Improved Componentwise Verified Error Bounds for Least Squares Problems and Underdeter-



- mined Linear Systems. *Numerical Algorithms*, **66**, 309-322. <https://doi.org/10.1007/s11075-013-9735-6>
- [12] Miyajima, S. (2015) Fast Enclosure for the Minimum Norm Least Squares Solution of the Matrix Equation  $AXB = C$ . *Numerical Linear Algebra with Applications*, **22**, 548-563. <https://doi.org/10.1002/nla.1971>