第 33 卷 第 4 期
2004 年 11 月

测 绘 学 报
ACTA GEODAETICA et CARTOGRAPHICA SINICA

Vol. 33, No. 4
Nov. , 2004

文章编号: 1001-1595(2004) 04-0293-06 中图分类号: P207 文献标识码: A

L_q 估计的渐近方差-协方差矩阵及其特点

彭军还[1,2], 陶本藻[3], 黄 珹[1], 叶叔华[1]

(1. 中国科学院 上海天文台, 上海 200030; 2. 桂林工学院 土木工程系, 广西 桂林 541004; 3. 武汉大学 测绘学院, 湖北 武汉 430079)

The Asymptotic Variance-covariance Matrix in L_q-norm Estimate and Its Characteristic

PENG Jun-huan[1,2], TAO Ben-zao[3], HUANG Cheng[1], YE Shu-hua[1]

(1. *Shanghai Astronomical Observatory, CAS, Shanghai* 200030, *China;* 2. *Deartment of Civil Engineering, Guilin Institute of Technology, Guilin* 541004, *China;* 3. *School of Geodesy and Geomatics, Wuhan University, Wuhan* 430079, *China)*

Abstract: For independent and heteroscedastic errors generated by increasing from the independent, identically distributed errors, according to the Bahadur-type linear representation of M-estimate of unknowns, this paper derives the Bahadur-type linear representation of the basic vector including the observational vector, the residual vector, the estimated vector of the unknowns and the adjusted observational vector. The asymptotic variance-covariance matrix of the basic vector for statistical analysis is further derived from the law of variance propagation and determined by the three nuisance parameters. The third nuisance parameter is defined first in this paper. For L_q-norm estimate, the three nuisance parameters and the corresponding variance-covariance matrix are derived respectively from errors being normally distributed and errors being distributed in L_q-norm function. For the Least Squares estimate or L_2-norm estimate, residuals are respectively independent of the estimator of the unknown parameters and the adjusted observations, statistically; the property is irrelative to the error distribution. For L_q-norm estimate with errors being normally distributed, the covariance matrices between the residual vector and the estimated vector of the unknown parameters, as well as the adjusted observational vector are not zero. However, for L_q-norm estimate with errors being distributed in q-norm function, it is the corresponding maximal likelihood estimate, the covariance matrices between the residual vector and the estimated vector of the unknown parameters, as well as the adjusted observational vector are zero. The derived forms and conclusions can be used in statistical analysis.

Key words: Bahadur-type linear representation; basic vector; variance-covariance matrix; nuisance parameters; L_q-norm estimate

摘 要: 针对由独立同分布误差膨胀而成的独立不等精度误差, 根据未知参数的 M 估计的 Bahadur 型线性表达式, 本文导出了由观测量、残差向量、参数估计量和观测量平差向量组成的基本向量的 Bahadur 型表达式。进一步地, 根据方差传播定律导出了 M 估计的基本向量的渐近方差-协方差矩阵, 该矩阵由 3 个多余参数决定, 第三多余参数由本文定义。对 L_q 范估计,

收稿日期: 2003-06-03; 修回日期: 2004-05-12
基金项目: 国家 973 项目(G1998040703); 国家自然科学基金重点项目(19833030); 国家自然科学基金项目(NSFC-10073017); 广西青年自然科学基金资助(9912008)

作者简介: 彭军还(1964-), 男, 重庆人, 博士, 教授, 从事量数据处理理论及其在大地测量、天体测量以及空间飞行器精密定轨中的应用研究。

分别计算了误差分别为正态分布和 q 范分布时的 3 个多余参数, 以及相应的基本向量的方差协方差矩阵。对最小二乘估计, 残差向量与参数估计量和观测量的平差向量统计独立, 相应的协方差矩阵为零, 这| 性质与误差分布无关。对正态分布的 L_q 估计, 残差向量与参数估计量和观测量平差向量的协方差不为零; 而对 q 范分布的 L_q 估计, 即是相应的极大似然估计, 残差向量与参数估计量和观测量平差向量的协方差为零。文中所得公式和结论可用于统计分析。

关键词: Bahadur 型线性表达式; 基本向量; 方差协方差矩阵; 多余参数; L_q 范估计

1 引 言

考虑观测方程

$$\boldsymbol{V}= \boldsymbol{A}\hat{\boldsymbol{X}}- \boldsymbol{L} \tag{1}$$

和随机模型

$$\boldsymbol{D}_{LL}= \sigma^2\boldsymbol{P}^{-1} \tag{2}$$

式中, $\boldsymbol{A}$ 是设计矩阵; $\boldsymbol{L}$ 是观测量, 其平差值为 $\hat{\boldsymbol{L}}$; $\boldsymbol{V}$ 是残差或改正数向量; $\hat{\boldsymbol{X}}$ 是未知参数 $\boldsymbol{X}$ 的平差值; $\boldsymbol{P}$ 是观测量的权阵, 是对角阵, 权为 $p_i = \sigma^2/\sigma_i^2$; σ^2 是单位权方差; $\boldsymbol{D}_{LL}$是观测量的方差矩阵。未知参数的最小二乘解 $\hat{\boldsymbol{X}} = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{PA})^{-1}\cdot\boldsymbol{A}^{\mathrm{T}}\boldsymbol{PL}$ 是观测量的线性函数, 这使得基本随机向量 $\boldsymbol{Z}= (\boldsymbol{L}^{\mathrm{T}}, \hat{\boldsymbol{X}}^{\mathrm{T}}, \boldsymbol{V}^{\mathrm{T}}, \hat{\boldsymbol{L}}^{\mathrm{T}})^{\mathrm{T}}$ 能够表达成观测量(或者观测误差)的线性函数, 从而使基本随机向量的协方差能够求出并以显式表示

$$\boldsymbol{D}_{ZZ}= \begin{bmatrix} \boldsymbol{D}_{LL} & \boldsymbol{D}_{L\hat{X}} & \boldsymbol{D}_{LV} & \boldsymbol{D}_{L\hat{L}} \\ \boldsymbol{D}_{\hat{X}L} & \boldsymbol{D}_{\hat{X}\hat{X}} & \boldsymbol{D}_{\hat{X}V} & \boldsymbol{D}_{\hat{X}\hat{L}} \\ \boldsymbol{D}_{VL} & \boldsymbol{D}_{V\hat{X}} & \boldsymbol{D}_{VV} & \boldsymbol{D}_{V\hat{L}} \\ \boldsymbol{D}_{\hat{L}L} & \boldsymbol{D}_{\hat{L}\hat{X}} & \boldsymbol{D}_{\hat{L}V} & \boldsymbol{D}_{\hat{L}\hat{L}} \end{bmatrix} =$$

$$\sigma^2\begin{bmatrix} \boldsymbol{P}^{-1} & \boldsymbol{AN}^{-1} & \boldsymbol{AN}^{-1}\boldsymbol{A}^{\mathrm{T}}- \boldsymbol{P}^{-1} & \boldsymbol{AN}^{-1}\boldsymbol{A}^{\mathrm{T}} \\ \boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}} & \boldsymbol{N}^{-1} & 0 & \boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}} \\ \boldsymbol{AN}^{-1}\boldsymbol{A}^{\mathrm{T}}- \boldsymbol{P}^{-1} & 0 & \boldsymbol{P}^{-1}- \boldsymbol{AN}^{-1}\boldsymbol{A}^{\mathrm{T}} & 0 \\ \boldsymbol{AN}^{-1}\boldsymbol{A}^{\mathrm{T}} & \boldsymbol{AN}^{-1} & 0 & \boldsymbol{AN}^{-1}\boldsymbol{A}^{\mathrm{T}} \end{bmatrix} \tag{3}$$

式中, $\boldsymbol{N}= \boldsymbol{A}^{\mathrm{T}}\boldsymbol{PA}$。可是, 对具有非线性特征的 M 估计[1], 基本随机向量(除观测量本身外)不能表示成观测量的线性形式, 因此谋求一定条件下 M 估计的某种线性表示并求出这种表示下的基本随机向量的协方差矩阵是有重要意义的。

在测量数据处理理论中, 对非线性函数的处理就是在某点展开成拉格朗日级数, 去掉高次项, 用线性项来逼近非线性的原函数。前提条件是高次项可以看成无穷小。对非线性的条件方程和误差方程, 展开点是观测值(条件平差)或由观测值求得的参数近似值(间接平差), 由于有严格的规范限制和精密仪器保证, 所得到的观测值是真值的很好的近似, 从计算误差角度考虑, 对非线性的条件方程和误差方程展开后的高次项完全可以忽略不计。M 估计的解方程为[1]

$$\sum_{i=1}^{n} \psi(p_i^{1/2}(\boldsymbol{a}_i^{\mathrm{T}}\hat{\boldsymbol{X}} - L_i))\boldsymbol{a}_i p_i^{1/2} = 0 \tag{4}$$

其解式为

$$\hat{\boldsymbol{X}}= (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{WA})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{WL} \tag{5}$$

式中, $\psi(x)= \partial\rho(x)/\partial x$, $\rho(x)$ 是 M 估计的目标函数; $\boldsymbol{a}_i$ 是设计矩阵 $\boldsymbol{A}$ 的第 i 个行向量; p_i 是第 i 个观测 L_i 的权; $\boldsymbol{W}$ 是等价权阵, 是一对角阵, 对角元素为

$$W_i= p_i\psi(p_i^{1/2}V_i)/(p_i^{1/2}V_i) \tag{6}$$

V_i 是第 i 个残差, 由 M 估计解得。

可见, 由 M 估计得到的解 $\hat{\boldsymbol{X}}$ 就是观测量 $\boldsymbol{L}$ 的非线性函数, 关系复杂, 无法展开成 $\boldsymbol{L}$ 的直接线性函数。退而求其次, 如能得到 $\hat{\boldsymbol{X}}$ 关于 $\boldsymbol{L}$(或观测误差)的某种间接线性函数、根据这种间接线性函数进一步能得到整个基本随机向量的某种线性表达式, 从而得到基本随机向量的协方差矩阵, 那也是令人欣慰的[2]。所幸的是统计学家对 M-估计的线性表示从概率意义上已取得了前期重要成果。

该领域的最早工作是中位数的线性表示[3]

$$\sqrt{n}(\hat{X} - X) = (2\sqrt{n}f(0))^{-1}\sum_{i=1}^{n}\mathrm{sign}(\Delta_i) + R_n \tag{7}$$

式中, $\hat{X}$ 是位置参数的中位数估计; X 是期望; f 是误差分布; Δ_i 是独立同分布的观测真误差, 密度为 $f(\Delta_i, \sigma^2)$, σ^2 是方差; sign 是符号函数; $R_n= O(n^{-1/2})$ 随样本容量增大而趋于无穷小。可见线性表示的逼近精度与样本容量有关。如果将上式改写成等价形式

$$\hat{X} - X = (2nf(0))^{-1}\sum_{i=1}^{n}\mathrm{sign}(\Delta_i) + n^{-1/2}R_n \tag{8}$$

则 $n^{-1/2}R_n$ 在样本趋于无穷大时是 R_n 的高阶无穷小。Bahadur 得到了独立同分布的最小绝对残

差的线性表示[4]

$$(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{1/2}(\hat{\boldsymbol{X}}-\boldsymbol{X})=$$
$$(2f(0))^{-1}(\boldsymbol{A}^{\mathrm{T}}A)^{-1/2}\sum_{i=1}^{n}\mathrm{sign}(\Delta_i)\boldsymbol{a}_i+R_n \quad (9)$$

当观测误差独立同分布时, 陈希孺得到了更一般的 M 估计的线性表示[2]

$$(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{1/2}(\hat{\boldsymbol{X}}-\boldsymbol{X})=$$
$$\lambda_1^{-1}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1/2}\sum_{i=1}^{n}\psi(\Delta_i)\boldsymbol{a}_i+R_n \quad (10)$$

式中, $\lambda_1=\mathrm{E}(\psi'(\Delta_i))=\int\psi'(\Delta_i)f(\Delta_i,\sigma^2)\mathrm{d}\Delta_i\neq 0$ 称为第一多余参数(the first nuisance parameter)。$(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{1/2}(\hat{\boldsymbol{X}}-\boldsymbol{X})$ 服从渐近正态分布[2]

$$(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{1/2}(\hat{\boldsymbol{X}}-\boldsymbol{X})\sim N(0,\lambda_1^{-2}\lambda_2\boldsymbol{I}) \quad (11)$$

式中, $\lambda_2=\mathrm{E}(\psi^2(\Delta_i))=\int\psi^2(\Delta_i)f(\Delta_i,\sigma^2)\mathrm{d}\Delta_i$ 是第二多余参数(the second nuisance parameter), 并且有

$$\lambda_1^{-2}\lambda_2=\mathrm{E}(IF^2(\Delta_i)) \quad (12)$$

式中, $IF(\Delta_i)=\psi(\Delta_i)/\mathrm{E}(\psi'(\Delta_i))$ 与独立同分布时误差对 1 维位置参数的影响函数[1]相同。由式(11)得参数的 M 估计的渐近方差矩阵为

$$\boldsymbol{D}_{\hat{X}\hat{X}}=\lambda_1^{-2}\lambda_2(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1} \quad (13)$$

式(10)的等价形式为

$$\hat{\boldsymbol{X}}=\boldsymbol{X}+\lambda_1^{-1}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\sum_{i=1}^{n}\psi(\Delta_i)\boldsymbol{a}_i+(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1/2}R_n=$$
$$\boldsymbol{X}+\lambda_1^{-1}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\psi+(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1/2}R_n \quad (14)$$

式中, $\psi=(\psi(\Delta_1),\psi(\Delta_2),\cdots,\psi(\Delta_n))^{\mathrm{T}}$。根据文献[2], 由 $\lim_{n\to\infty}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}=0$, 可得 $\lim_{n\to\infty}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1/2}=0$, 因此, $(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1/2}R_n$ 是 R_n 的高阶无穷小。因为 $\boldsymbol{X}$ 是常量, $(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1/2}R_n$ 是可忽略的无穷小, 可以忽略。在式(14)中应用方差传播定律同样可得到式(13)。

对测量数据处理而言, 仅仅得到参数的 M 估计的方差矩阵是不够的, 我们还希望得到整个基本向量 $\boldsymbol{Z}=(\boldsymbol{L}^{\mathrm{T}},\hat{\boldsymbol{X}}^{\mathrm{T}},\boldsymbol{V}^{\mathrm{T}},\hat{\boldsymbol{L}}^{\mathrm{T}})^{\mathrm{T}}$ 的方差协方差矩阵。它不仅具有理论意义, 其中的一些方差矩阵还具有实际用途, 比如, 残差的方差矩阵是构造基于 M 残差的巴尔达型统计量的前提。针对由独立同分布误差膨胀而成的独立不等精度误差, 本文在于获取基本向量的方差协方差矩阵。

2　独立不等精度 M 估计的线性表示

本节用 $\bar{\Delta}_1,\bar{\Delta}_2,\cdots,\bar{\Delta}_n$ 表示前面的具有方差 σ^2 的独立同分布误差, 密度为 $f(\bar{\Delta}_i,\sigma^2)$。$\lambda_1=\mathrm{E}(\psi'(\bar{\Delta}_i))$ 和 $\lambda_2=\mathrm{E}(\psi^2(\bar{\Delta}_i))$ 与第一节的多余参数完全等价; $\Delta_1,\Delta_2,\cdots,\Delta_n$ 表示分别具有方差 σ_i^2 的独立不等精度误差, 密度为 $g_i(\Delta_i,\sigma_i^2)$, 同时设该误差由独立同分布误差膨胀生成, 具有确定的函数关系 $\Delta_i={}_i^{-1/2}\bar{\Delta}_i$ 或 $\bar{\Delta}_i={}_i^{1/2}\Delta_i$, ${}_i=\sigma^2/\sigma_i^2$。

对 $\Delta_i={}_i^{-1/2}\bar{\Delta}_i$, 根据已知随机变量密度求随机变量函数的概率密度的定理[10], $g_i(\Delta_i,\sigma_i^2)$ 可以表示为

$$g_i(\Delta_i,\sigma_i^2)=f(\bar{\Delta}_i,\sigma^2)\big|_{\bar{\Delta}_i={}_i^{1/2}\Delta_i}\left|\frac{\mathrm{d}\bar{\Delta}_i}{\mathrm{d}\Delta_i}\right|=$$
$$f({}_i^{1/2}\Delta_i,\sigma^2)\ {}_i^{1/2}$$

同理, 对 $\bar{\Delta}_i={}_i^{1/2}\Delta_i$, 有

$$f(\bar{\Delta}_i,\sigma^2)=g_i(\Delta_i,\sigma_i^2)\big|_{\Delta_i={}_i^{-1/2}\bar{\Delta}_i}\left|\frac{\mathrm{d}\Delta_i}{\mathrm{d}\bar{\Delta}_i}\right|=$$
$$g_i({}_i^{-1/2}\bar{\Delta}_i,\sigma_i^2)\ {}_i^{-1/2}$$

在上述前提下, 于式(14)中, 分别用 $\boldsymbol{P}^{1/2}\boldsymbol{A}$, ${}_i^{1/2}\boldsymbol{a}_i$ 和 ${}_i^{1/2}\Delta_i$ 代替 $\boldsymbol{A}$, $\boldsymbol{a}_i$ 和 Δ_i 即可得到未知参数 M 估计的线性表示

$$\hat{\boldsymbol{X}}=\boldsymbol{X}+\lambda_1^{-1}\boldsymbol{N}^{-1}\sum_{i=1}^{n}\psi({}_i^{1/2}\Delta_i)\ {}_i^{1/2}\boldsymbol{a}_i+\boldsymbol{N}^{-1/2}R_n=$$
$$\boldsymbol{X}+\lambda_1^{-1}\boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}^{1/2}\psi+\boldsymbol{N}^{-1/2}R_n \quad (15)$$

式中, $\lambda_1=\mathrm{E}(\psi'({}_i^{1/2}\Delta_i))$,

$\psi=(\psi({}_1^{1/2}\Delta_1),\psi({}_2^{1/2}\Delta_2),\cdots,\psi({}_n^{1/2}\Delta_n))^{\mathrm{T}}$

$\boldsymbol{N}^{-1/2}R_n$ 与 R_n 同样是高阶无穷小。剩余项 R_n 的无穷小形式在于它的纯数学分析意义, 在实用中可以忽略不计, 而应用中最需要的是线性项部分。

$$\lambda_1=\mathrm{E}(\psi'({}_i^{1/2}\Delta_i))=$$
$$\int\psi'({}_i^{1/2}\Delta_i)g_i(\Delta_i,\sigma_i^2)\mathrm{d}\Delta_i=$$
$$\int\psi'(\bar{\Delta}_i)g_i({}_i^{-1/2}\bar{\Delta}_i,\sigma_i^2)\ {}_i^{-1/2}\mathrm{d}\bar{\Delta}_i=$$
$$\int\psi'(\bar{\Delta}_i)f(\bar{\Delta}_i,\sigma^2)\mathrm{d}\bar{\Delta}_i$$

可以看出, 与第一节的相应多余参数完全等价。对 λ_2 和后面导出的 λ_3 具有同样的性质; 对由独立同分布的误差膨胀而成的独立不等精度的误差进行等精度变换后的多余参数与独立同分布时的多余参数相同, 与不等精度的误差无关。

由式(15), 根据方差传播定律得不等精度观测的参数的 M 估计的方差矩阵为

$$\boldsymbol{D}_{\hat{X}\hat{X}}=\lambda_1^{-2}\lambda_2\boldsymbol{N}^{-1}$$

式中, $\lambda_2=E(\psi^2({}_i^{1/2}\Delta_i))$。据式(15), 有

$$\boldsymbol{V}=\boldsymbol{A}\hat{\boldsymbol{X}}-\boldsymbol{L}=$$

$$\lambda_1^{-1} \boldsymbol{A} \boldsymbol{N}^{-1} \boldsymbol{A}^{\mathrm{T}} \boldsymbol{P}^{1/2} \phi - \Delta + \boldsymbol{A} \boldsymbol{N}^{-1/2} R_n$$

$$\hat{\boldsymbol{L}} = \lambda_1^{-1} \boldsymbol{A} \boldsymbol{N}^{-1} \boldsymbol{A}^{\mathrm{T}} \boldsymbol{P}^{1/2} \phi + \boldsymbol{A} \boldsymbol{X} + \boldsymbol{A} \boldsymbol{N}^{-1/2} R_n$$

基本向量 $\boldsymbol{Z} = (\boldsymbol{L}^{\mathrm{T}}, \hat{\boldsymbol{X}}^{\mathrm{T}}, \boldsymbol{V}^{\mathrm{T}}, \hat{\boldsymbol{L}}^{\mathrm{T}})^{\mathrm{T}}$ 有如下线性表示

$$\begin{bmatrix} \boldsymbol{L} \\ \hat{\boldsymbol{X}} \\ \boldsymbol{V} \\ \hat{\boldsymbol{L}} \end{bmatrix} = \begin{bmatrix} 0 \\ \lambda_1^{-1} \boldsymbol{N}^{-1} \boldsymbol{A}^{\mathrm{T}} \boldsymbol{P}^{1/2} \\ \lambda_1^{-1} \boldsymbol{A} \boldsymbol{N}^{-1} \boldsymbol{A}^{\mathrm{T}} \boldsymbol{P}^{1/2} \\ \lambda_1^{-1} \boldsymbol{A} \boldsymbol{N}^{-1} \boldsymbol{A}^{\mathrm{T}} \boldsymbol{P}^{1/2} \end{bmatrix} \phi + \begin{bmatrix} \boldsymbol{I} \\ 0 \\ -\boldsymbol{I} \\ 0 \end{bmatrix} \Delta + \begin{bmatrix} \boldsymbol{A} \\ -\boldsymbol{I} \\ 0 \\ \boldsymbol{A} \end{bmatrix} \boldsymbol{X} + \begin{bmatrix} 0 \\ \boldsymbol{N}^{-1/2} \\ \boldsymbol{A} \boldsymbol{N}^{-1/2} \\ \boldsymbol{A} \boldsymbol{N}^{-1/2} \end{bmatrix} R_n \tag{16}$$

线性表达式的特点:

1. 无穷小量 R_n 与样本大小有关;
2. λ_1 与误差分布及其单位权方差有关。

3　基本向量的协方差矩阵

由于数学期望 $\boldsymbol{X}$ 是常量, R_n 是关于样本大小的无穷小量, 在计算协方差矩阵时可以忽略。

由于误差向量 Δ 中各误差彼此独立, $\phi(x)$ 是奇函数, 所以

$$\mathrm{E}(\phi({}_i^{1/2}\Delta_i)) = 0$$

$$\mathrm{E}(\phi({}_i^{1/2}\Delta_i)\phi({}_j^{1/2}\Delta_j)) = 0,\ i \neq j$$

令 $\lambda_3 = \mathrm{E}(\phi({}_i^{1/2}\Delta_i)({}_i^{1/2}\Delta_i))$, 我们称为第三多余参数(the third nuisance parameter), 则

$$D_\phi = \mathrm{E}((\phi - \mathrm{E}(\phi))(\phi - \mathrm{E}(\phi))^{\mathrm{T}}) = \lambda_2 \boldsymbol{I} \tag{17}$$

$$D_{\phi\Delta} = \mathrm{E}((\phi - \mathrm{E}(\phi))(\Delta - \mathrm{E}(\Delta))^{\mathrm{T}}) = \lambda_3 \boldsymbol{P}^{-1/2} \tag{18}$$

根据协方差传播定律, 有

$$\boldsymbol{D}_{ZZ} = \begin{bmatrix} 0 \\ \lambda_1^{-1} \boldsymbol{N}^{-1} \boldsymbol{A}^{\mathrm{T}} \boldsymbol{P}^{1/2} \\ \lambda_1^{-1} \boldsymbol{A} \boldsymbol{N}^{-1} \boldsymbol{A}^{\mathrm{T}} \boldsymbol{P}^{1/2} \\ \lambda_1^{-1} \boldsymbol{A} \boldsymbol{N}^{-1} \boldsymbol{A}^{\mathrm{T}} \boldsymbol{P}^{1/2} \end{bmatrix} \boldsymbol{D}_\phi \begin{bmatrix} 0 \\ \lambda_1^{-1} \boldsymbol{N}^{-1} \boldsymbol{A}^{\mathrm{T}} \boldsymbol{P}^{1/2} \\ \lambda_1^{-1} \boldsymbol{A} \boldsymbol{N}^{-1} \boldsymbol{A}^{\mathrm{T}} \boldsymbol{P}^{1/2} \\ \lambda_1^{-1} \boldsymbol{A} \boldsymbol{N}^{-1} \boldsymbol{A}^{\mathrm{T}} \boldsymbol{P}^{1/2} \end{bmatrix}^{\mathrm{T}} + \begin{bmatrix} 0 \\ \lambda_1^{-1} \boldsymbol{N}^{-1} \boldsymbol{A}^{\mathrm{T}} \boldsymbol{P}^{1/2} \\ \lambda_1^{-1} \boldsymbol{A} \boldsymbol{N}^{-1} \boldsymbol{A}^{\mathrm{T}} \boldsymbol{P}^{1/2} \\ \lambda_1^{-1} \boldsymbol{A} \boldsymbol{N}^{-1} \boldsymbol{A}^{\mathrm{T}} \boldsymbol{P}^{1/2} \end{bmatrix} \boldsymbol{D}_{\phi\Delta} \begin{bmatrix} \boldsymbol{I} \\ 0 \\ -\boldsymbol{I} \\ 0 \end{bmatrix} +$$

$$\begin{bmatrix} \boldsymbol{I} \\ 0 \\ -\boldsymbol{I} \\ 0 \end{bmatrix} \boldsymbol{D}_{\phi\Delta} \begin{bmatrix} 0 \\ \lambda_1^{-1} \boldsymbol{N}^{-1} \boldsymbol{A}^{\mathrm{T}} \boldsymbol{P}^{1/2} \\ \lambda_1^{-1} \boldsymbol{A} \boldsymbol{N}^{-1} \boldsymbol{A}^{\mathrm{T}} \boldsymbol{P}^{1/2} \\ \lambda_1^{-1} \boldsymbol{A} \boldsymbol{N}^{-1} \boldsymbol{A}^{\mathrm{T}} \boldsymbol{P}^{1/2} \end{bmatrix}^{\mathrm{T}} + \begin{bmatrix} \boldsymbol{I} \\ 0 \\ -\boldsymbol{I} \\ 0 \end{bmatrix} \boldsymbol{D}_\Delta \begin{bmatrix} \boldsymbol{I} \\ 0 \\ -\boldsymbol{I} \\ 0 \end{bmatrix}^{\mathrm{T}} =$$

$$\begin{bmatrix} \sigma^2 \boldsymbol{P}^{-1} & \lambda_1^{-1}\lambda_3 \boldsymbol{A}\boldsymbol{N}^{-1} & \lambda_1^{-1}\lambda_3 \boldsymbol{A}\boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}} - \sigma^2 \boldsymbol{P}^{-1} & \lambda_1^{-1}\lambda_3 \boldsymbol{A}\boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}} \\ \lambda_1^{-1}\lambda_3 \boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}} & \lambda_1^{-2}\lambda_2 \boldsymbol{N}^{-1} & (\lambda_1^{-2}\lambda_2 - \lambda_1^{-1}\lambda_3) \boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}} & \lambda_1^{-2}\lambda_2 \boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}} \\ \lambda_1^{-1}\lambda_3 \boldsymbol{A}\boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}} - \sigma^2 \boldsymbol{P}^{-1} & (\lambda_1^{-2}\lambda_2 - \lambda_1^{-1}\lambda_3) \boldsymbol{A}\boldsymbol{N}^{-1} & (\lambda_1^{-2}\lambda_2 - 2\lambda_1^{-1}\lambda_3) \boldsymbol{A}\boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}} + \sigma^2 \boldsymbol{P}^{-1} & (\lambda_1^{-2}\lambda_2 - \lambda_1^{-1}\lambda_3) \boldsymbol{A}\boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}} \\ \lambda_1^{-1}\lambda_3 \boldsymbol{A}\boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}} & \lambda_1^{-2}\lambda_2 \boldsymbol{A}\boldsymbol{N}^{-1} & (\lambda_1^{-2}\lambda_2^{\mathrm{T}} - \lambda_1^{-1}\lambda_3) \boldsymbol{A}\boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}} & \lambda_1^{-2}\lambda_2 \boldsymbol{A}\boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}} \end{bmatrix} \tag{19}$$

协方差矩阵特点:

在基本随机向量的协方差矩阵中, 除观测量(或观测误差)的协方差矩阵以外, 有如下特点。

1. 协方差矩阵由 3 个多余参数 λ_1, λ_2, λ_3 设计矩阵和权矩阵决定。
2. 3 个多余参数 λ_1, λ_2 和 λ_3 取决于 M 准则的 $\phi(x)$、误差分布和单位权方差。
3. 协方差矩阵依样本大小渐近成立。
4. 除最小二乘法以外, 一般情况下 $\boldsymbol{D}_{V\hat{X}}$, $\boldsymbol{D}_{V\hat{L}}$ 不再为零, 说明残差与参数平差量和观测量平差量不再独立。其原因在于, 在一般的 M 估计中, 观测值的平差量不再是观测量在设计矩阵列向量所张成的空间的正交投影, 即残差向量与观测量平差量不再垂直。

4　$L_q(1 \leqslant q \leqslant 2)$ 估计的多余参数与协方差矩阵

一旦给定 M 估计准则以及误差分布密度, 我们即可得到多余参数的具体表达式, 进一步地, 就得到式(19)的具体形式。因此, 我们只需求出多余参数的具体形式即可。独立不等精度观测的 $L_q(1 \leqslant q \leqslant 2)$ 范估计准则为

$$\sum_{i=1}^{n} (\sqrt{\ {}_i}\ |\Delta_i|)^q \Big|_{\Delta_i = V_i} = \min \tag{20}$$

由于 ${}_i^{1/2}\Delta$ 就是将不等精度化为等精度, 具有相同的方差 σ^2。在计算多余参数时可以采用具有方差 σ^2 的误差 Δ 来定义相应的函数, 即

$$\rho(\Delta)=q^{-1}|\Delta|^{q}$$
$$\psi(\Delta)=|\Delta|^{q-1}\mathrm{sign}(\Delta)$$
$$\psi'(\Delta)=(q-1)|\Delta|^{q-2}+|\Delta|^{q-1}\mathrm{sign}'(\Delta)$$

4.1 误差为正态分布时的多余参数与协方差矩阵

具有方差 σ^2 的误差正态分布为

$$f(\Delta)=(2\pi)^{-1/2}\sigma^{-1}\exp(-\Delta^2(2\sigma^2)^{-1})$$

$$\lambda_1=E(\psi'(\Delta))=$$
$$\int_{-\infty}^{\infty}((q-1)|\Delta|^{q-2}+|\Delta|^{q-1}\mathrm{sign}'(\Delta))\cdot$$
$$(2\pi)^{-1/2}\sigma^{-1}\exp(-\Delta^2(2\sigma^2)^{-1})\,d\Delta=$$
$$2\int_{0}^{\infty}(q-1)\Delta^{q-2}(2\pi)^{-1/2}\sigma^{-1}\exp(-\Delta^2(2\sigma^2)^{-1}d\Delta+$$
$$2|0|^{q-1}(2\pi)^{-1/2}\sigma^{-1}$$

令 $t=\Delta^2(2\sigma^2)^{-1}$, 将 $\Delta=2^{1/2}\sigma t^{1/2}$ 和 $d\Delta=2^{-1/2}\sigma t^{-1/2}dt$ 代入 λ_1 中并整理得

$$\lambda_1=\sigma^{q-2}2^{(q-2)/2}\pi^{-1/2}(q-1)\Gamma(\frac{q-1}{2})+$$
$$2|0|^{q-1}(2\pi)^{-1/2}\sigma^{-1}=$$
$$\begin{cases}2(2\pi)^{-1/2}\sigma^{-1} & q=1\\ \sigma^{q-2}2^{(q-2)/2}\pi^{-1/2}(q-1)\Gamma(\frac{q-1}{2}) & 2\geqslant q>1\end{cases}$$

同理得

$$\lambda_2=E(\psi^2(\Delta))=\sigma^{2(q-1)}2^{q-1}\pi^{-1/2}\Gamma(\frac{2q-1}{2})$$
$$\lambda_3=E(\psi(\Delta)\Delta)=\sigma^{q}2^{q/2}\pi^{-1/2}\Gamma(\frac{q+1}{2})$$

将 $q=1$ 时的 $\lambda_1^{-2}\lambda_2=2^{-1}\pi\sigma^2$ 和 $\lambda_1^{-1}\lambda_3=\sigma^2$ 代入式(19)即得 L_1 估计的基本向量的方差协方差矩阵

$$\boldsymbol{D}_{ZZ}=\sigma^2\begin{bmatrix}\boldsymbol{P}^{-1} & \boldsymbol{A}\boldsymbol{N}^{-1} & \boldsymbol{A}\boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}}-\boldsymbol{P}^{-1} & \boldsymbol{A}\boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}}\\ \boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}} & 2^{-1}\pi\boldsymbol{N}^{-1} & (2^{-1}\pi-1)\boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}} & 2^{-1}\pi\boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}}\\ \boldsymbol{A}\boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}}-\boldsymbol{P}^{-1} & (2^{-1}\pi-1)\boldsymbol{A}\boldsymbol{N}^{-1} & (2^{-1}\pi-2)\boldsymbol{A}\boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{P}^{-1} & (2^{-1}\pi-1)\boldsymbol{A}\boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}}\\ \boldsymbol{A}\boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}} & 2^{-1}\pi\boldsymbol{A}\boldsymbol{N}^{-1} & (2^{-1}\pi-1)\boldsymbol{A}\boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}} & 2^{-1}\pi\boldsymbol{A}\boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}}\end{bmatrix}\quad(21)$$

可以看出, 该式与最小二乘估计的式(3)不同, 其中没有零分块矩阵, 说明式中的 $\boldsymbol{D}_{VL}$, $\boldsymbol{D}_{VX}$ 及其转置不再为零。

将 $q=2$ 时的 $\lambda_1^{-2}\lambda_2=\sigma^2$ 和 $\lambda_1^{-1}\lambda_3=\sigma^2$ 代入式(19)即得 L_2 估计(最小二乘估计)的基本向量的方差协方差矩阵, 与式(3)完全相同。

当 $2>q>1$ 时, $\lambda_1^{-2}\lambda_2$ 和 $\lambda_1^{-1}\lambda_3$ 分别为:

$$\lambda_1^{-2}\lambda_2=\sigma^2 2\pi^{1/2}(q-1)^{-2}\Gamma(\frac{2q-1}{2})\Gamma(\frac{q-1}{2})^{-2}$$
$$\lambda_1^{-1}\lambda_3=\sigma^2 2(q-1)^{-1}\Gamma(\frac{q+1}{2})\Gamma(\frac{q-1}{2})^{-1}=\sigma^2$$

将它们代入式(19)即得相应的基本向量的方差协方差矩阵(省略, 不写出)。由于这里的 $\lambda_1^{-2}\lambda_2$, $\lambda_1^{-1}\lambda_3$ 和 $\lambda_1^{-2}\lambda_2-\lambda_1^{-1}\lambda_3$ 不都为零, 协方差阵中的 $\boldsymbol{D}_{VL}$, $\boldsymbol{D}_{VX}$ 及其转置也不为零。

4.2 误差为 L_q 范分布时的多余参数与协方差矩阵

此时的 L_q 估计即是 L_q 范分布的极大似然估计。具有方差 σ^2 的 L_q 范分布的密度函数为[7]

$$f(\Delta)=\frac{q^{1-q^{-1}}}{2k\sigma\Gamma(q^{-1})}\exp(-q^{-1}\left|\frac{\Delta}{k\sigma}\right|^{q})\qquad(22)$$

式中, $k=(q^{2q^{-1}}\Gamma(3q^{-1})/\Gamma(q^{-1}))^{1/2}$。

多余参数的计算方法与 4.1 节相同, 有如下结果:

$$\lambda_1=E(\psi'(\Delta))=\int_{-\infty}^{\infty}((q-1)|\Delta|^{q-2}+$$
$$|\Delta|^{q-1}\mathrm{sign}'(\Delta))\frac{q^{1-q^{-1}}}{2k\sigma\Gamma(q^{-1})}\exp(-q^{-1}\left|\frac{\Delta}{k\sigma}\right|^{q})d\Delta=$$
$$(q-1)(k\sigma)^{q-2}q^{1-2q^{-1}}\frac{\Gamma(1-q^{-1})}{\Gamma(q^{-1})}+$$
$$2|0|^{q-1}\frac{q^{1-q^{-1}}}{2k\sigma\Gamma(q^{-1})}=$$
$$\begin{cases}\dfrac{q^{1-q^{-1}}}{k\sigma\Gamma(q^{-1})}, & q=1\\ q^{2(1-q^{-1})}(k\sigma)^{q-2}\dfrac{\Gamma(2-q^{-1})}{\Gamma(q^{-1})}, & 2\geqslant q>1\end{cases}$$

$$\lambda_2=E(\psi^2(\Delta))=$$
$$\int_{-\infty}^{\infty}|\Delta|^{2(q-1)}\frac{q^{1-q^{-1}}}{2k\sigma\Gamma(q^{-1})}\exp(-\ ^{-1}\left|\frac{\Delta}{k\sigma}\right|^{q})d\Delta=$$
$$q^{2(1-q^{-1})}(k\sigma)^{2(q-1)}\frac{\Gamma(2-q^{-1})}{\Gamma(q^{-1})}$$

$$\lambda_3=E(\psi(\Delta)\Delta)=$$
$$\int_{-\infty}^{\infty}|\Delta|^{q-1}\mathrm{sign}(\Delta)\Delta\frac{q^{1-q^{-1}}}{2k\sigma\Gamma(q^{-1})}\cdot$$
$$\exp(-\ ^{-1}\left|\frac{\Delta}{k\sigma}\right|^{q})d\Delta=(k\sigma)^{q}$$

将 λ_1, λ_2, λ_3 代入式(19)并整理得

$$
\boldsymbol{D}_{ZZ}=k_q\sigma^2\begin{bmatrix} k_q^{-1}\boldsymbol{P}^{-1} & \boldsymbol{A}\boldsymbol{N}^{-1} & \boldsymbol{A}\boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}}-k_q^{-1}\boldsymbol{P}^{-1} & \boldsymbol{A}\boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}} \\ \boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}} & \boldsymbol{N}^{-1} & 0 & \boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}} \\ \boldsymbol{A}\boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}}-k_q^{-1}\boldsymbol{P}^{-1} & 0 & -\boldsymbol{A}\boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}}+k_q^{-1}\boldsymbol{P}^{-1} & 0 \\ \boldsymbol{A}\boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}} & \boldsymbol{A}\boldsymbol{N}^{-1} & 0 & \boldsymbol{A}\boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}} \end{bmatrix} \tag{23}
$$

式中, $k_q=q^{-2}\dfrac{\Gamma^2(q^{-1})}{\Gamma(3q^{-1})\Gamma(2-q^{-1})}$。式(23)即是观测误差服从 q-范分布的L_q 估计的基本向量的方差协方差矩阵, 并且式(23)中有 4 个零分块矩阵, 分别与 $\boldsymbol{D}_{VL}$, $\boldsymbol{D}_{VX}$ 及其转置对应。可见, 作为极大似然估计的 L_q 估计, 其残差向量与参数估计量和观测量平差量的协方差为零, 与最小二乘估计具有同样的规律。

5 讨论与结论

由统计学家在误差独立同分布条件下得到的线性表达式与未知参数估计量的渐近方差矩阵不适用于污染分布。因此, 本文导出的结果只适用于由独立同分布误差膨胀而成的独立不等精度误差。

在基本向量的方差协方差矩阵中, 除观测量平差量的方差协方差矩阵仅具有理论意义外, 其余的方差协方差矩阵不仅具有理论意义, 而且具有实用价值。比如, 我们可以根据它们构造方差分量估计公式, 还可以根据残差的方差矩阵构造残差检验统计量。

基本向量的方差协方差矩阵由依赖于误差分布的 3 个多余参数惟一确定, 其中第三多余参数由本文首次定义。对 M 估计的一种, 即 L_q 范估计, 导出了多余参数与相应基本向量的方差协方差矩阵。

1. 当误差取 q 范分布时, L_q 范估计是极大似然估计, $\boldsymbol{D}_{VL}$, $\boldsymbol{D}_{VX}$ 及其转置为零;

2. 误差取正态分布时, 除最小二乘估计外, 其余 L_q 范估计的 $\boldsymbol{D}_{VL}$, $\boldsymbol{D}_{VX}$ 及其转置不为零。

参考文献:

[1] HUBER P J. Robust Statistics [M]. New York: Wiley, 1981.

[2] CHEN Xi-ru. M-methods in Linear Model[M]. Shanghai: Shanghai Scientific and Technical Publishing House, 1995. (in Chinese)

[3] BAHADUR R R. A Note on Quantiles in Large Samples[J]. Annals of Mathematics and Statistics, 1966, (37): 577-580.

[4] BABU G J. Strong Representations for LAD Estimates in Linear Models[J]. Probability Theory Related Fields, 1989, 83(F): 547-558.

[5] YU Zhong-cou. The Basis of Surveying Adjustment [M]. Wuhan: Press of Wuhan Technical University of Surveying and Mapping, 2001. 93-98. (in Chinese)

[6] CHEN Xi-ru. Introduction to Mathematics and Statistics[M]. Beijing: Science Press, 1997. 452-454. (in Chinese)

[7] SUN Hai-yan. The Uniform Mode of Error Distribution[J]. Journal of Wuhan Technical University of Surveying and Mapping, 1995. (in Chinese)

[8] PENG Jun-huan. Research on Non-linear M-estimates and Its Application[D]. Wuhan: Wuhan University, 2003. (in chinese)

[9] PENG Jun-huan, XIE Zhi-yin. Robustified $L\infty$ Norm Estimate[J]. Survey Review, 2002, (4).

[10] SHENG Yong-huan. The Practical Manual of Mathematics[M]. Beijing: Science Press, 1992. 468.

[11] YANG Y. Estimator of Covariance Matrix at Robust Estimation Based on Influence Functions [J]. Zeitschrift fuer Vermessungswesen, 1997, 122(4): 166-174.

[12] XU P. On Robust Estimation with Correlated Observation[J]. Bulletin Geodesique, 1989, (63): 237-252.

[13] YANG Y. Robust Estimation and Its Applications [M]. Beijing: Bayi Publishing House.

[14] RAO C R. Linear Models, Least Squares and Alternatives[M]. New York: Springer-Verlag.

[15] HUANG You-cai. The Data Snoop and Resistant Estimate[M]. Beijing: Publishing House of Surveying and Mapping, 1989. 365-366.