DOI:10.13203/j.whugis1990.02.012

第15卷 第2期 武汉测绘科技大学学报 Vol.15 No.2
1990年 6月 Journal of Wuhan Technical University of Surveying and Mapping Jun. 1990

最小二乘估计残差的统计性质和应用

张 方 仁

摘 要

最小二乘估计求得的观测值残差，有多种形式的定义，它们具有不同的分布。本文论述它们的概率分布和性质并讨论其用途。编算了多余观测分量r_i和标准化残差e_i的临界值表。此外，应用线性变换求得了不相关残差$W=U^TV$和$S=D_W^{-\frac{1}{2}}W$。建议用S作粗差检验，本文编制了相应于这种检验的$\max\limits_{i\leq n-t}|S_i|$的临界值表，最后计算了一个实测的测边网，用求得的各种残差作了统计检验。

【关键词】 最小二乘估计残差；统计分布；多余观测分量；临界值；标准化残差

1 引 言

设平差的函数模型（误差方程）为

$$V=B\cdot X-l;\qquad R(B)=t \tag{1}$$

上式中$B=[b_{(1)}\cdots b_{(t)}]=[b_1\cdots b_n]^T$，$b_{(i)}$是$n$维向量，$b_i$是$t$维行向量，$l=[l_1\cdots l_n]^T$是观测值向量，定义残差$V$。

$$V=B\hat{X}-l=-(I-BN^{-1}B^T)l=-(I-H)l=-Q_Vl \tag{2}$$

$N=B^TB$，$H=(h_{ij})$为帽子矩阵，Q_V是V的协因数阵，V为分析测量误差和评定精度等提供了丰富的信息。残差的种类很多，其用处也不尽相同，本文主要研究它们的统计性质和应用。

2 普通残差

（2）式所定义的残差称为“普通残差”。因为$\mathrm{Var}(v_i)=\sigma^2(1-h_{ii})$，故当$h_{ii}\to 1$时，$\mathrm{Var}(v_i)\to 0$。此时第$i$个误差方程对求参数估值所起的作用大，在观测和平差时应特别注意。这样的点称为强权点。根据多元统计分析理论可导出

收稿日期：1988—06—29

$$F_i^* = \frac{n-t}{t-1} \cdot \frac{h_{ii} - \frac{1}{n}}{1 - h_{ii}} \sim F_{(t-1,\ n-t)} \tag{3}$$

当$F_i^* > F_{\alpha;\ (t-1,\ n-t)}$时，就认为是在显著水平α下的强权点。

用r_i表示第i个观测值的多余观测分量，则

$$r_i = (Q_V P)_{ii} = (I-H)_{ii} = 1 - h_{ii} \tag{4}$$

上式是相应于等权的情形，对于不等权观测，亦有类似的公式。

$$V = -(I - B(B^T PB)B^T P)l = -(I - BN^{-1}B^T P)l$$

$$= -Q^{-\frac{1}{2}}(I - P^{-\frac{1}{2}}BN^{-1}B^T P^{-\frac{1}{2}})P^{-\frac{1}{2}}l = -Q^{-\frac{1}{2}}(I - H')P^{-\frac{1}{2}}l \tag{5}$$

Q是观测值l的协因数阵，$H' = (h'_{ij})$是帽子矩阵。

因 $$Q_V = Q^{-\frac{1}{2}}(I - H')Q^{-\frac{1}{2}} \tag{6}$$

故 $$r_i = (Q_V P)_{ii} = (Q^{-\frac{1}{2}}(I - H')Q^{\frac{1}{2}}P)_{ii} = (I - H')_{ii} = 1 - h'_{ii} \tag{7}$$

由（3）、（4）、（7）式不难求得r_i的临界值。当

$$\frac{n-t}{t-1} \cdot \frac{h_{ii} - \frac{1}{n}}{1 - h_{ii}} = \frac{n-t}{t-1} \cdot \frac{1 - r_i - \frac{1}{n}}{r_i} \leqslant F_{\alpha;\ (t-1,\ n-t)}$$

时 $$r_i \geqslant \frac{(n-1(n-t)}{n[n-t+F_\alpha \cdot (t-1]} = r_\alpha \tag{8}$$

当 $$\frac{n-t}{t-1} \cdot \frac{h_{ii} - \frac{1}{n}}{1 - h_{ii}} = \frac{n-t}{t-1} \cdot \frac{1 - r_i - \frac{1}{n}}{r_i} \geqslant F_{1-\alpha;\ (t-1,\ n-t)}$$

时 $$r_i \leqslant \frac{(n-1)(n-t)}{n[n-t+F_{1-\alpha} \cdot (t-1)]} = r_{1-\alpha} \tag{9}$$

因此得到r_i的允许范围为

$$r_\alpha \leqslant r_i \leqslant r_{1-\alpha} \qquad (i = 1,\ \cdots,\ n) \tag{10}$$

若各误差方程能满足（10）式的要求，当然是最理想的。但因为控制网的多余观测数是一定的，因此当r_i满足（10）式的单侧要求时，亦可认为基本上达到了可靠性指标的要求。

作者根据（8）、（9）式编算了临界值r_α，$r_{1-\alpha}$表。这里仅给出$\alpha = 0.05$的r_α，$r_{1-\alpha}$值（表1），可供应用时参考。

$\alpha = 0.05$　　　　**表 1**

	10	20	30	40	60
10	0.239	0.148	0.107	0.086	0.059
	0.713	0.523	0.408	0.335	0.238
20	0.445	0.313	0.241	0.198	0.144
	0.819	0.663	0.552	0.471	0.364
30	0.567	0.429	0.346	0.291	0.221
	0.868	0.739	0.637	0.558	0.447
40	0.645	0.512	0.428	0.367	0.286
	0.896	0.786	0.695	0.620	0.509

3　标准化残差e′和e

普通残差的方差为$\mathrm{Var}(\nu_i) = \sigma^2(1 - h_{ii})$，由于各个$\nu_i$的方差不相同，因此直接用它们来判定观测值中是否含有粗差显然是不合适的。为此，令

$$e'_i = \frac{\nu_i}{\sigma\sqrt{1 - h_{ii}}} \tag{11}$$

若σ已知，则$e' \sim N(0, 1)$。但σ往往是未知的，一般可用$\hat{\sigma} = [\|B\hat{x} - l\|^2/(n-t)]^{\frac{1}{2}}$代替，则可得

$$e_i = \frac{v_i}{\hat{\sigma}\sqrt{1-h_{ii}}} \tag{12}$$

称e_i为"标准化残差"。关于e_i的分布密度的推证，可参阅文献［1］，这里直接给出结果。

记 $e = [e_1 \cdots e_n]^T$； $e_I = [e_{i_1} \cdots e_{i_m}]^T$

$$D = \mathrm{diag}[1-h_{i_1 i_1}, \cdots, 1-h_{i_m i_m}]; \qquad C_I = D^{-\frac{1}{2}}(I-H_I)D^{-\frac{1}{2}}$$

H_I为H阵中第i_1，…，i_m行和第i_1，…，i_m列交叉处的元素所构成的矩阵，当$\varepsilon \sim N(0, \sigma^2 I)$时，$e_I$的分布密度为

$$f(e_I) = \begin{cases} \dfrac{\Gamma\left(\frac{n-t}{2}\right)|I-H_I|^{-\frac{1}{2}}}{\Gamma\left(\frac{n-m-t}{2}\right)[(n-t)\pi]^{\frac{m}{2}}} \prod\limits_{j=1}^{m}(1-h_{i_j i_j})^{-\frac{1}{2}}\left[1-\dfrac{e_I^T C_I^{-1} e_I}{n-t}\right]^{\frac{n-m-t-2}{2}}, & e_I^T C_I e_I \leqslant n-t \\ 0 & \text{其它} \end{cases} \tag{13}$$

当$m=1$时，就得到了e_i的分布密度。

$$f(e_i) = \begin{cases} \dfrac{\Gamma\left(\frac{n-t}{2}\right)}{\Gamma\left(\frac{n-1-t}{2}\right)\Gamma\left(\frac{1}{2}\right)(n-t)^{\frac{1}{2}}}\left(1-\dfrac{e_i^2}{n-t}\right)^{\frac{n-3-t}{2}}, & |e_i| \leqslant (n-t)^{\frac{1}{2}} \\ 0, & \text{其它} \end{cases} \tag{14}$$

当$m=1$且$t=1$时，就得到τ变量的分布密度

$$\varphi(e_i) = \begin{cases} \dfrac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n-2}{2}\right)\Gamma\left(\frac{1}{2}\right)(n-1)^{\frac{1}{2}}}\left(1-\dfrac{e_i^2}{n-1}\right)^{\frac{n-4}{2}}, & |e_i| \leqslant (n-1)^{\frac{1}{2}} \\ 0, & \text{其它} \end{cases} \tag{15}$$

根据(14)、(15)两式就能得出如下两点推论：

（1）应用分布变换的方法，易证得$\frac{e_i^2}{n-t}$为服从参数$p=\frac{1}{2}$，$q=\frac{n-1-t}{2}$的βeta分布。

（2）用密度公式可证得$E(e_i)=0$，$\mathrm{Var}(e_i)=1$，$\mathrm{Cov}(e_i, e_j) = \frac{h_{ij}}{\sqrt{(1-h_{ii})(1-h_{jj})}}$ $(i \neq j)$。

因为$\mathrm{Cov}(e_i, e_j)$的数值一般都很小，所以常将e_i与e_j看作是不相关的。如再假定$e_i \sim N(0, 1)$，则可对e_i作正态性检验。若先按βeta分布求出$\frac{e_i^2}{n-t}$的临界值，再推算出e_i的临界值（表2给出了$\alpha=0.025$；0.05的临界值），此检验就是严格的。

文献［5］中β分布表载有x值，x是按下列公式计算出的

$$F_x(p,\ q)=\frac{1}{B(p,\ q)}\int_0^x t^{p-1}(1-t)^{q-1}\mathrm{d}t=\alpha \tag{16}$$

上式中 $F_x(p,\ q)$是β变量的分布函数，$\alpha=0.5$；0.25；0.1；0.05；0.025；0.01；0.005；0.0025；0.001等数值，但统计检验需要的是$\alpha=0.995$；0.99；0.975；0.95；0.90等值相应的$1-x$。

因为
$$F_{1-x}(q,\ p)=\frac{1}{B(q,\ p)}\int_0^{1-x} t^{q-1}(1-t)^{p-1}\mathrm{d}t \tag{17}$$

令$t=1-u$，则$1-t=u$，而$\mathrm{d}t=-\mathrm{d}u$。将它们代入（17）式并相应的变换积分限，就得：

$$F_{1-x}(q,\ p)=\frac{1}{B(q,\ p)}\int_1^x \mu^{p-1}(1-\mu)^{q-1}(-\mathrm{d}\mu) \tag{18}$$

考虑到

$$B(q,\ p)=\frac{\Gamma(q)\Gamma(p)}{\Gamma(q+p)}=\frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}=B(p,\ q)$$

则
$$F_{1-x}(q,\ p)=\frac{1}{B(p,\ q)}\int_x^1 t^{p-1}(1-t)^{q-1}\mathrm{d}t \tag{19}$$

所以

$$F_x(p,\ q)+F_{1-x}(q,\ p)=\frac{1}{B(p,\ q)}\left\{\int_0^x t^{p-1}(1-t)^{q-1}\mathrm{d}t+\int_x^1 t^{p-1}(1-t)^{q-1}\mathrm{d}t\right\}=1 \tag{20}$$

根据（20）式再利用β分布表，就可求得相应于$\alpha=0.995$；0.99；0.975；0.95；0.90等的$1-x$值。

4 学生化残差e^*

标准化残差的检验用数表，文献[2]与本文编算了一部分，但仍不能满足不同情况下使用的需要。为此，定义另一种残差。

$$e_i^*=\frac{v_i}{\hat{\sigma}(i)\sqrt{1-h_{ii}}}\qquad (i=1,\ \cdots,\ n) \tag{21}$$

e_i^*称为“学生化残差”，上式中的$\hat{\sigma}(i)$是根据模型（17）求的σ的估值。

$$V=B(i)\cdot X-l \tag{22}$$

用$B(i)$、$l(i)$分别表示从B、l中剔除第i行以后的矩阵和向量。用$\hat{X}(i)$表示由模型（22）求得的X的最小二乘估值。

$$\hat{X}(i)=(B(i)^{T}B(i)^{-1}B(i)^{T}l(i) \tag{23}$$

$$\hat{\sigma}(i)^2=l(i)^{T}[I-B(i)(B(i)^{T}B(i)^{-1}B(i)^{T}]l(i)/(n-1-t) \tag{24}$$

学生化残差的优点表现在：

（1）在观测误差 $\varepsilon \sim N(0, \sigma^2 I)$ 条件下，v_i 和 $\frac{(n-1-t)\hat{\sigma}^2(i)}{\sigma^2}$ 分别服从于正态分布和 $\chi^2_{(n-1-t)}$ 分布，且互相独立。因此 e_i^* 就服从于自由度为 $n-1-t$ 的 t 分布。

（2）当检验和剔除粗差时，若第 i 个观测值中含有粗差，由于计算 $\hat{\sigma}(i)^2$ 时未采用该观测值，从而排除了该粗差的影响，提高了检验功效。

在计算 e_i^* 时不必对每个 i 用(24)式求 $\hat{\sigma}(i)^2$，可用下列的简便计算公式[1]、[3]。

$$\hat{\sigma}(i)^2 = \frac{n-t-e_i^{*2}}{n-1-t}\hat{\sigma}^2 \tag{25}$$

5 不相关残差W

对普通残差 V 进行某种线性变换，可使变换得到的残差 W 的各分量间互不相关。

V 的协因数阵 Q_V 是秩亏的，其秩等于 $n-t$。这说明 V 中只有 $n-t$ 个分量是线性独立的。对 V 作如下的线性变换。

$$W = U^T V \tag{26}$$

称 W 为"不相关残差"。U^T 的行向量间具有正交性。

即：
$$U^T U = I_{n-t} \tag{27}$$

且 U^T 还满足下列条件

$$U^T B = 0 \tag{28}$$

这里的 0 是 $(n-t)\times t$ 的零矩阵，在 $\varepsilon \sim N(0, \sigma^2 I)$ 的假定条件下，$W \sim N(0, \sigma_W^2 I_{n-t})$。

矩阵 U 是确实存在的，对 Q_V 用雅可比方法作变换，使成为

$$Q_W = \widetilde{U} Q_V \widetilde{U}^T$$

Q_W 为一对角阵，对角线上的非零元素为 λ_1，$\cdots\lambda_{n-t}$，它们就是 Q_V 的非零特征值，$\widetilde{U}$ 中的列向量就是 Q_V 的特征向量，取 $\widetilde{U}$ 中相应于特征值为 λ_1，$\cdots$，λ_{n-t} 的特征向量就构成了 U 矩阵。

关于 W 的应用，有以下两点意见：

（1）当 $\varepsilon \sim N(0, \sigma^2 I)$ 时，则 w_i 之间不仅不相关而且是相互独立的。因此可用它作 χ^2 拟合适度检验、偏、峰态检验及夏皮罗检验等。

（2）因 $D_W = \sigma^2 Q_W$，记 $S = D_W^{-\frac{1}{2}} W$，则 $S = [s_1, \cdots, S_{n-t}]^T$ 为标准正态向量，且 s_i 间是互相独立的，故拟用它作为检验粗差的统计量。

取显著水平为 α，现确定临界值 s_α。

$$p(\text{所有的 } |s_i| \leqslant s_\alpha) = p(|s_i| \leqslant s_\alpha)\cdot p(|s_2| \leqslant s_\alpha)\cdots$$
$$p(|s_{n-t}| \leqslant s_\alpha) = p(\max_{i\leqslant n-t} |s_i| \leqslant s_\alpha) = 1-\alpha \tag{30}$$

检验的关键在于求出 $\max\limits_{i\leqslant n-t} |s_i|$ 的概率分布，为此先求 $|s_i|$ 的分布函数 $\phi(x)$ 和分布密度 $\varphi(x)$。

$$\phi(x)=\int_{-x}^{x}f(t)\mathrm{d}t=\int_{-x}^{0}f(t)\mathrm{d}t+\int_{0}^{x}f(t)\mathrm{d}t$$

$$\varphi(x)=\frac{\mathrm{d}\phi(x)}{\mathrm{d}x}=f(x)+f(x)=2f(x)=\sqrt{\frac{2}{\pi}}e^{-\frac{x^2}{2}}$$

若以$F_n(x)$表示$\max\limits_{i\leqslant n-t}|s_i|$的分布函数，则不难证明

$$F_n(x)=p\left(\max_{i\leqslant n-t}|s_i|\leqslant x\right)=[\phi(x)]^{n-t}=\left[\int_0^x\sqrt{\frac{2}{\pi}}e^{-\frac{x^2}{2}}\right]^{n-t} \tag{31}$$

据此编算了$\max\limits_{i\leqslant n-t}|s_i|$的临界值表（表 3 给出$\alpha=0.025$；0.05的临界值）

除以上四种残差外，还有其他的一些残差，因与测绘专业的应用结合得不密切，目前对它们的研究也还很不够，故不作介绍。

6 算　　例

本文对一个测角网（图 1 ）进行了平差计算，算得各种残差，并作了相应的统计检验。该网有 2 个已知点， 8 个未知点，38个方向观测值。

1. 数据

Ⅰ组：是实测的观测数据，观测精度$\sigma=2.00''$。观测数据经过外业检核和各种统计检验，认为观测值中未含有粗差，系统误差亦不显著。

Ⅱ组：在 7 号实测方向值上减去 8″；在20号实测方向值上加上 8″；在30号实测方向值上加上 8″；其余均为实测方向值，其数值与Ⅰ组的相同。

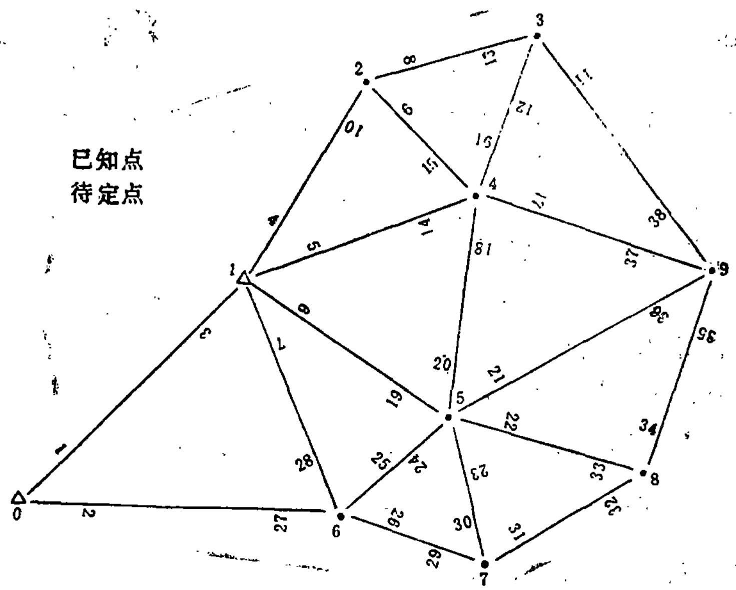

图　1

表 2

α \ $n-1-t$	2	3	4	5	6	7	8	9	10	11	12	13	14
0.025	1.689	1.847	1.941	2.000	2.041	2.070	2.092	2.109	2.123	2.134	2.143	2.151	2.157
0.05	1.645	1.757	1.814	1.898	1.870	1.855	1.896	1.904	1.910	1.916	1.920	1.926	1.926

α \ $n-1-t$	15	16	17	18	19	20	22	24	26	28	30	40	60
0.025	2.163	2.168	2.172	2.176	2.180	2.183	2.188	2.193	2.197	2.200	2.203	2.212	2.223
0.05	1.929	1.931	1.933	1.934	1.936	1.937	1.939	1.941	1.943	1.944	1.945	1.949	1.953

表 3

α \ $n-t$	5	6	7	8	9	10	12	14	16	18	20	30	40
0.025	2.804	2.862	2.911	2.952	2.988	3.020	3.075	3.120	3.160	3.190	3.224	3.338	3.417
0.05	2.569	2.631	2.683	2.727	2.766	2.800	2.857	2.907	2.948	2.984	3.016	3.137	3.220

表 4

	序号	4	7	10	18	28
Ⅰ	v_i	+0″.99	−1″.25	−0″.81	+0″.71	+1″.17
	e_i	+1″.92	−2″.12	−1″.65	+1″.22	+2″.12
	e_i^*	+2″.21	−2″.57	−1″.80	+1″.25	+2″.56
	s_i				+1″.10	

	序号	18	25	30		
Ⅱ	v_i	−3″.00	−0″.16	−2″.67		
	e_i	−2″.44	−0″.12	−2″.13		
	e_i^*	−3″.30	−0″.12	−2″.59		
	s_i		−3″.47			

	序号	4	10	29	34	
Ⅲ	v_i	+5″.85	−4″.77	+1″.36	−4″.36	
	e_i	+2″.59	−2″.22	+0″.62	−2″.09	
	e_i^*	+3″.73	−2″.76	+0″.61	−2″.51	
	s_i			+4″.47		

	序号	1	2	3	27	29
Ⅳ	v_i	+5″.00	−5″.00	−5″.00	+5″.00	−0″.11
	e_i	+2″.92	−2″.92	−2″.92	+2″.92	−0″.06
	e_i^*	+5″.19	−5″.19	−5″.19	+5″.19	−0″.05
	s_i					+6″.27

Ⅲ组：在4号实测方向值上减去15″；在20号实测方向值上减去15″；在34号实测方向值上加上15″；其余均为实测方向值，其数值与Ⅰ组的相同。

Ⅳ组：在3号实测方向值上加上15″；在27号实测方向值上减去15″；在31号实测方向值上减去15″；其余均为实测方向值，其数值均与Ⅰ组的相同。

2．计算

对于四组数据分别计算了v_i、e_i、e_i^*、w_i及s_i，这些数值列在表4中，为了求得上述这些残差，还计算了参数X的估值$\hat{X}$，协因数阵Q_V，Q_V的特征值λ_i和特征向量矩阵U，本文未列出这些数值。

3．统计检验

（1）e_i检验。以$n-1-t=11$，$\alpha=0.05$查表2得临界值$e_{0.05(11)}=1.916$。

Ⅰ组：全部的$|e_i|\leqslant 1.916$。

Ⅱ组：$|e_{18}|=2.44>1.916$；$|e_{30}|=2.13>1.916$；其余的$|e_i|\leqslant 1.916$。

Ⅲ组：$|e_4|=2.59>1.916$；$|e_{10}|=2.22>1.916$；$|e_{34}|=2.09>1.916$；其余的$|e_i|\leqslant 1.916$。

Ⅳ组：$|e_1|=|e_2|=|e_3|=|e_{27}|=2.92>1.916$；其余的$|e_i|\leqslant 1.916$。

（2）e_i^*检验。以$n-1-t=11$，$\alpha=0.05$查t分布表，得临界值$t_{0.025,(11)}=2.201$。

Ⅰ组：全部的$|e_i^*|\leqslant 2.201$。

Ⅱ组：$|e_{18}^*|=3.30>2.201$；$|e_{30}^*|=2.59>2.201$；其余的$|e_i^*|\leqslant 2.201$。

Ⅲ组：$|e_4^*|=3.73>2.201$；$|e_{10}^*|=2.76>2.201$；$|e_{34}^*|=2.51>2.201$；其余的$|e_i^*|\leqslant 2.201$。

Ⅳ组：$|e_1^*|=|e_2^*|=|e_3^*|=|e_{27}^*|=5.19>2.201$；其余的$|e_i^*|\leqslant 2.201$。

（3）s_i验验。该网的观测精度指标$\sigma=2.00''$，根据实测数据算得$\hat{\sigma}=0.96''$。作此项检验时取$\sigma=1.50''$。以$n-t=12$，$\alpha=0.05$查表5得$\max\limits_{i\leqslant 12}|s_i|$的临界值$s_{0.05,(12)}=2.857$。

Ⅰ组：$\max\limits_{i\leqslant 12}|s_i|=1.10<2.857$。　　Ⅱ组：$\max\limits_{i\leqslant 12}|s_i|=3.47>2.857$。

Ⅲ组：$\max\limits_{i\leqslant 12}|s_i|=4.47>2.857$。　　Ⅳ组：$\max\limits_{i\leqslant 12}|s_i|=6.27>2.857$。

7 结 束 语

普通残差一般不能充分反映观测误差的分布情况，绝对值大的残差并不一定出现在误差大的观测方向上。标准化残差和学生化残差的检验结果大致相同。有的观测方向中不含有粗差但还是有绝对值很大的e_i和e_i^*，有的观测方向中含有粗差，但它们的e_i和e_i^*的绝对值也不是很大，不过在较多的情况下，e_i和e_i^*的绝对值大的方向观测值中确含有粗差，因此，对粗差检验来说，e_i和e_i^*检验的总体效果还是较好的，其缺点在于还不能全部精确定位。e_i与e_i^*的检验功效大致相当，但因为计算e_i^*必须先算$\hat{\sigma}(i)$，故似乎作e_i检验要稍方便些。

$\max\limits_{i\leqslant n-t}|s_i|$检验有较好的效果，对于算例中的四种情形均作出了正确的判断。

建议对观测数据最好要进行多种检验，这样就有可能对观测质量作出正确的评价。对于观测误差服从另外的一些分布和观测值中含有系统误差的情形，本文还未用模拟数据进行计算分析，有待今后进一行研究。

参 考 文 献

〔1〕 陈希孺，王松桂．近代回归分析——原理方法及应用．安徽教育出版社，1988．

〔2〕 张方仁，易法楷．残差V和τ变量的概率分布及其应用．武汉测绘学院学报，1985(1)．

〔3〕 王松桂．线性回归诊断．数理统计与管理，1985，6．

〔4〕 张方仁，王 勇．主成分检验和粗差定位．武汉测绘科技大学学报，1986(4)．

〔5〕 Λ Н БОЛЬЩЕБ, Н В CMNPHOB. ТАБЛЧbi MATEMATNЧECKOЙ CTATNCTИKИ. ИЗДATEΛЬCTBO⟨HAYKA⟩, MACKBA 1965.

The Statiscal Properties and Application of Least Squares Estimates of the Residuals

Zhang Fangren

Abstract

The residualt of observations, derived from least squares estimation can be defined in a great variety of ways and they have different probability distributions. This paper deals with their probability distributions, properties and applications. The table for the critical value of the redundancy numbers $r_i=(Q_V P)_{ii}$ and the standar-dized residuals e_i are compiled. In addition, the independent residuals $W=U^T V$ and $S=D_w^{-\frac{1}{2}}W$ are derived by means of linear transformation. It is supposed that S can be used for the detection of gross errors, while the critieal value table carresponding to this test has been given. Finally, the adjustment for a trilateration network has been carried out, and the statisfical tests have been performed by using a variety of the residuals obtained from the adustment.

[Key words] least squares estimates of the residuals; probability distribution; redundancy numbers; critical value; standardized residual; gross errors