第 29 卷第 1 期　　辽宁工程技术大学学报（自然科学版）　　2010 年 2 月
Vol.29 No.1　　Journal of Liaoning Technical University（Natural Science）　　Feb. 2010
文章编号：1008-0562(2010)01-0044-04

整体最小二乘法直线拟合

丁克良[1,2]，沈云中[2]，欧吉坤[3]
（1. 北京建筑工程学院 测绘与城市空间信息学院，北京 100044；2. 现代工程测量国家测绘局 重点实验室，上海 200092；3.中国科学院测量与地球物理研究所，湖北 武汉 430077）

摘　要：针对在直线拟合中，因变量选取不同拟合的结果有差异现象，提出采用整体最小二乘法进行直线拟合。文章在分析直线方程特点的基础上，采用 EIV 模型描述直线方程，在解算中根据系数矩阵的特点应用 QR 分解分为将方程两部分，采用了混合最小二乘法求解。理论分析和实际计算结果表明，整体最小二乘法顾及了因变量和自变量的误差。拟合精度高于普通最小二乘法，采用整体最小二乘拟合直线，整体上优于普通最小二乘法。

关键词：直线拟合；普通最小二乘法；整体最小二乘法；EIV 模型
中图分类号：O 411.1　　　**文献标识码**：A

Methods of line-fitting based on total least-squares

DING Keliang[1,2]，SHENG Yunzhong[2]，OU Jikun[3]
(1.Beijing university Civil Engineering And Architecture, School of Geomatics and Urban Information, Beijing 100044,China ;2. Key Laboratory of Advanced Engineering Surveying of SBSM, Shanghai 200092 China ;3. Institute of Geodesy and Geophysics, Chinese Academy of Sciences, Wuhan, 430077 ,China)

Abstract: Line fitting obtained by ordingary least square is often different if the independent variable is defferent, the reason that result in the defference is analysised in the paper. Then the method of line fitting by total least squares is proposed. We describle the line equation with errors-in-variables model, and in the parameter solution the coefficient matrix is divided into two parts by using the QR decomposition. And then the Related parameter can be achieved by the ordianary least squares and total least squares. A conclusion can be got that ling fitting by total least squares is more effective than that of ordianary least square on the whole.

Key words: line fitting; ordinary least squares; total least squares; errors-in-variable model

0　引　言

直线拟合问题在诸多试验和工程实际问题中都会遇到，它可以描述为：对于给定的 m 个测量点 (x_i, y_i)，$(i=1,2\cdots,m)$，寻找一条最佳的拟合直线，使其尽可能通过或靠近这些点。拟合的实质是求直线参数斜率和截距的最佳估计，拟合方法通常采用普通最小二乘法求解拟合参数。该方法简单实用，应用广泛。然而，值得注意的问题是自变量和因变量的选择不同，所得拟合直线是不同的。理论上讲，无论自变量如何选取，最佳的拟合结果只能有一个。近年来，针对这一问题诸多学者进行了研究[1-3]，但是并未给出明确的答案。本文深入分析了因变量选取不同造成拟合结果差异的原因，进而用 EIV[4]模型描述直线拟合方程，采用整体最小二乘法[5-7]计算拟合参数。理论分析和实际计算表明整体最小二乘曲线拟合可以获得最佳的拟合效果。

1　最小二乘法拟合直线

直线方程可表示为

$$y_i = ax_i + b \quad (i=1,2\cdots m) \tag{1}$$

式中，(x_i, y_i) 为测点坐标，a 为直线的斜率，b 为 y 轴的截距，a、b 为待估参数，a_0、b_0 为它们的近似值。令

$$a = a_0 + \delta a$$
$$b = b_0 + \delta b$$

以 y 作为因变量，以 x 为自变量，误差方程为[8]

收稿日期：2007-01-25
基金项目：现代工程测量国家测绘局重点实验基金资助项目（ES-SBSM-(07)-05）；国家自然科学基金资助项目（40771178）
作者简介：丁克良（1968-），男，河南 淮阳人，博士，主要从事现代测量数据误差处理研究，E-mail：hyiding@bucea.edu.cn。本文编校：于永江

$$v_{yi}=\begin{bmatrix}x_i & 1\end{bmatrix}\begin{bmatrix}\delta a\\ \delta b\end{bmatrix}+(a_0x_i+b_0-y_i) \quad (2)$$

误差方程矩阵表达式

$$A\delta X=l+V \quad (3)$$

其中

$$A=\begin{bmatrix}x_1 & 1\\ x_2 & 1\\ \vdots & \vdots\\ x_m & 1\end{bmatrix},\quad l=\begin{bmatrix}a_0x_1+b_0-y_1\\ a_0x_2+b_0-y_2\\ \vdots\\ a_0x_m+b_0-y_m\end{bmatrix},$$

$$V=\begin{bmatrix}v_{y1}\\ v_{y2}\\ \vdots\\ v_{ym}\end{bmatrix},\quad \delta X=\begin{bmatrix}\delta a\\ \delta b\end{bmatrix},$$

按最小二乘准则

$$V^{\mathrm{T}}V=\min$$

即

$$\sum_{i=1}^{m}\|ax_i-b-y_i\|^2=\min \quad (4)$$

其最小二乘解为

$$\delta\hat{X}=(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}l \quad (5)$$

因变量残差

$$V=A\delta\hat{X}-l \quad (6)$$

单位权中误差

$$\hat{\sigma}_0=\sqrt{\frac{V^{\mathrm{T}}V}{m-2}} \quad (7)$$

实际工程中也有采用横坐标 x 为因变量，纵坐标 y 为自变量进行拟合,这时直线方程可表示为

$$x_i=k_1y_i+k_2 \quad (i=1,2,\cdots,m) \quad (8)$$

求解方法与前类似。

以上是直线拟合的通常做法，这种方法的前提是只考虑因变量的误差，误差方程中的系数阵 A 被认为是精确值。实际情况是由于测量误差的存在，横坐标 x、纵坐标 y 都有误差，系数阵 A 同样存在误差。表 1 为同一组数据选取不同因变量的计算结果。从计算结果可以看出，两种方法拟合的参数和观测值中误差均有明显的差别。观测量 x、y 均是含有误差的随机变量，而拟合时只考虑了因变量的误差，即采用普通最小二乘法直线拟合只能顾及其中一种变量的误差。自变量的选取、观测精度、直线的特点等都会影响拟合结果，两种方法拟合结果自然就不一致。

表 1　同一组数据不同变量最小二乘直线拟合比较

Tab.1　least squares line fitting comparsion of defferent variable about the same data

测点坐标/cm	x	5.9	5.4	4.4	4.6	3.5	3.7	2.8	2.8	2.4	1.5
	y	0.0	0.9	1.8	2.6	3.3	4.4	5.2	6.1	6.5	7.4
数学模型		$y=ax+b$				$x=k_1y+k_2$，$a=1/k_1$，$b=-k_2/k_1$					
拟合准则		$\sum_{i=1}^{m}\|(ax_i+b-y_i)\|^2=\min$				$\sum_{i=1}^{m}\|(k_1y_i+k_2-x_i)\|^2=\min$					
待估参数	a	-1.767 1				-1.853 3					
	b	10.358 4				10.677 2					
单位权中误差/cm		0.572				0.316					

2　整体最小二乘法直线拟合

如果考虑到直线方程（1）中自变量 x 的误差，直线的条件方程可表示为

$$y_i+v_{yi}=\hat{a}(x_i+v_{xi})+\hat{b} \quad (i=1,2,\cdots,m) \quad (9)$$

相应误差方程（3）中设计矩阵 A 和观测向量 l 都含有误差，则误差方程可以按 EIV 模型描述[6]

$$(A+E_A)\delta X=l+E_l \quad (10)$$

式中，E_A、E_l 分别表示设计矩阵 A 和观测向量 l 的误差。

注意到设计矩阵 A 中，一列元素为固定值 1，这是一个混合最小二乘问题[7]，因此，应按混合最小二乘法求解。

令

$$A=\begin{bmatrix}A_1 & A_2\end{bmatrix}$$

式中

$$A_1=\begin{bmatrix}1 & 1 & \cdots & 1\end{bmatrix}^{\mathrm{T}},\quad A_2=\begin{bmatrix}x_1 & x_2 & \cdots & x_m\end{bmatrix}^{\mathrm{T}}$$

下面按参数求解、残差计算及精度评定等几个方面阐述其计算过程。

2.1　整体最小二乘参数求解

构造增广矩阵 $C=[A\quad l]$ ，并对其作 QR 三

角分解

$$C = QR \tag{11}$$

式中，Q 为正交阵，R 为上三角阵。则

$$R=Q^{T}C=Q^{T}[A_1 \quad A_2 \quad l]=\begin{bmatrix} R_{11} & R_{12} & R_{1l} \\ 0 & R_{22} & R_{2l} \end{bmatrix}^{1}_{2} \tag{12}$$

可将方程分为两个部分

$$R_{11}\delta b + R_{12}\delta a = R_{1l} \tag{13}$$

$$R_{22}\delta a = R_{2l} \tag{14}$$

这里，R_{11}、R_{12}、R_{22}、R_{1l}、R_{2l}均为标量。在求解参数时，首先采用整体最小二乘法求解方程（14）获得参数δa，然后回代到方程（13）按普通最小二乘法求解参数δb。

构造增广矩阵$C_R = [R_{22} \quad R_{2l}]$，并对其进行奇异值分解

$$C_R = U \Sigma N^{T}$$

式中，$U=\left[\underbrace{U_1}_{2\times1} \quad \underbrace{U_2}_{2\times1}\right]$，$V=\left[\underbrace{N_1}_{2\times1} \quad \underbrace{N_2}_{2\times1}\right]$，

$$\Sigma = diag(\sigma_1, \sigma_2), \quad \sigma_1 > \sigma_2$$

参数δa整体最小二乘解为[7]

$$\delta\hat{a} = (R_{22}^{T}R_{22} - \sigma_2^2)^{-1}R_{22}^{T}R_{2l} \tag{15}$$

回代到方程（13）求得参数$\delta\hat{b}$

$$\delta\hat{b} = R_{11}^{-1}(R_{1l} - R_{12}\delta\hat{a}) \tag{16}$$

也可以按下式求解[7]

$$\begin{bmatrix} \delta\hat{a} \\ \delta\hat{b} \end{bmatrix} = \left(A^{T}A - \sigma_2^2\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right)^{-1} A^{T}l \tag{17}$$

未知数为

$$\hat{X}_{ab} = \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} a_0 + \delta\hat{a} \\ b_0 + \delta\hat{b} \end{bmatrix}$$

2.2 整体最小二乘残差计算和精度评定

应用整体最小二乘法求解未知参数，改正量包括设计矩阵、观测向量及观测值改正数，由于直线拟合方程是线性方程，各改正量之间有内在的一致性。由方程$(A+E_A)\delta X = l + E_l$和直线方程$y = ax + b$可知，设计矩阵改正量$E_A$应是观测值$x$的改正数，观测向量改正数$E_l$是观测值$y$的改正数。$v_{xi}$、$v_{yi}$为整体最小二乘在$x$方向和$y$方向的改正数，$v_{si}$为正交方向的改正数。如图1为普通最小二乘、整体最小二乘几何特性差异，和普通最小二乘法相比，这里采用正交方向的改正来量来估计直线拟合单位权中误差。残差的计算可按下面的方式计算[6]。

增广矩阵C_R的改正量为

$$E_{C_R} = \sigma_2 U_2 N_2^{T} \tag{18}$$

R的改正量为

$$E_R = \begin{bmatrix} 0 & 0 \\ 0 & E_{C_R} \end{bmatrix} \tag{19}$$

由此可计算设计矩阵和观测向量的改正数

$$[E_A \quad E_l] = [E_{A_1} \quad E_{A_2} \quad E_l] = Q^{T}E_R \tag{20}$$

则x方向、y方向的残差为

$$\begin{cases} v_x = E_{A_2} \\ v_y = E_l \end{cases} \tag{21}$$

正交距离残差

$$v_{si} = \sqrt{v_{xi}^2 + v_{yi}^2} \tag{22}$$

从几何意义上来讲，整体最小二乘的求解准则是要求正交距离的残差平方和最小，因此，观测值精度的评定应按照正交距离的残差来计算。

$$F_{tls} = v_s^{T}v_s = v_x^{T}v_x + v_y^{T}v_y \tag{23}$$

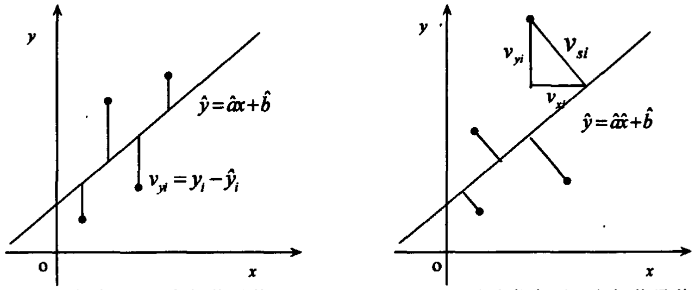
$$\sigma_{tls} = \sqrt{\frac{F_{tls}}{m-2}} \tag{24}$$

由于，整体最小二乘法顾及了x坐标和y坐标的误差，而且从x方向或者y方向改正量以及正交距离残差的改正数来看，都比普通最小二乘残差改正数小，单位权中误差要优于普通最小二乘法的计算值。

3 不同方法直线拟合计算结果比较

以上简要阐述了普通最小二乘法、整体最小二乘法的直线拟合方法，为对其拟合结果进行比较分析，以下以文献[9]给出的直线拟合数据为例，分别采用普通最小二乘和整体最小二乘法进行计算，并从模型、参数估值、单位权中误差等几个方面对各自的拟合结果对比分析。为便于比较，每种方法都选取$y = ax + b$、$x = k_1 y + k_2$两种模型进行对比计

算；对模型 $x=k_1y+k_2$ 的计算结果均换算为模型 $y=ax+b$，以比较各种方法的内在一致性。表 2 给出了两种方法的计算结果比较。



(a) LS(仅考虑以 y坐标的误差)　　(b) TLS同时考虑 x和 y坐标的误差

图 1 普通最小二乘和整体最小二乘几何特性比较

Fig.1 geometrical comparsion between ordinary least squares and total least squares

表 2 整体最小二乘与普通最小二乘直线拟合比较

Tab.2 comparsion between ordinary least squares and total least squares line fitting

项　目		普通最小二乘		整体最小二乘	
因变量		y	x	y	x
自变量		x	y	x	y
直线方程		$\hat{y}=\hat{a}x+\hat{b}$	$\hat{x}=\hat{k}_1y+\hat{k}_2$	$\hat{y}_i=\hat{a}\hat{x}_i+\hat{b}$	$\hat{x}_i=\hat{k}_1\hat{y}_i+\hat{k}_2$
误差模型		$A\delta X=l+V$		$(A+E_A)\delta X=(l+E_l)$	
待估参数	a	0.245 7	0.386 0	0.253 96	0.253 96
	b	3.662 9	2.934 9	3.620 10	3.620 10
单位权中误差/cm		0.485	1.576	0.471 0	0.471 0

从本例看出，对于普通最小二乘法而言，自变量选取不同参数拟合结果差别较大；原因在于拟合时只考虑因变量误差，未考虑自变量的误差。在许多实际的工程应用中，x、y 的选取是相对的，拟合结果好坏的标准应该是依据与实际工程整体接近程度，而不是仅仅考虑在一个方向最为接近。

整体最小二乘法综合考虑了系数阵和观测向量存在的误差，顾及了自变量和因变量的误差，从几何特性方面而言，整体最小二乘准则为 $\min\sum_{i=1}^{m}\left\|(ax_i+b-y_i)/\sqrt{1+a^2}\right\|^2$，其实质为测点到拟合直线的正交距离的平方和最小。

4 结　论

（1）从计算方法来讲，采用普通最小二乘法拟合直线，忽略了自变量的误差，方法简单，但是变量的选取不同，拟合的结果不同，差异的大小与测点的精度有关；拟合的结果只能保持在一个方向的拟合效果最佳。

（2）采用整体最小二乘法拟合直线，同时顾及了自变量和因变量的误差，无论选取哪种自变量模型其结果都是完全一致的，和普通最小二乘相比，采用整体最小二乘拟合直线能够获得最佳的拟合效果。

参考文献：

[1] 彭　放.经典回归直线的非唯一性研究[J].数学的实践与认识,2001(4):440-441.
[2] 李雄军.对 X 和 Y 方向最小二乘线性回归的讨论[J].计量技术.误差与数据处理,2005(1):50-51.
[3] 梁家辉.用最小二乘法进行直线拟合的讨论[J].工程物理,1995.
[4] Björck Å. Numerical Methods for Least Squares Problems. SIAM Publications[M].Philadelphia PA,1996.
[5] Golub G H, Van Loan C F. An analysis of the Total Least Squares problem[J]. SIAM J Numer Anal,1980,17(6):883-893.
[6] Van Huffel S, Vandewalle J.The Total Least Squares Problem Computational Aspects and Analysis[M].SIAM, Philadelphia ,1991.
[7] 丁克良.整体最小二乘法及其在测量数据处理中的若干应用研究[D].汉:中国科学院测量与地球物理研究所,2006.
[8] 武汉大学测绘学院.误差理论与测量平差基础[M].武汉:武汉大学出版社,2003.
[9] Wolf P R, Ghilani C D. Adjust Computations: Statistics and Least Squares in Surveying and GIS[M]. John:Wiley & Sons, Inc 1997.