

最小二乘直线拟合方法及其在同位素地质年代学中的应用研究

范文博¹⁾, 王文龙²⁾, 陈岳龙¹⁾, 李大鹏¹⁾

1) 中国地质大学(北京)地球科学与资源学院, 北京, 100083; 2) 北京理工大学数学学院, 北京, 100081

内容提要:最小二乘法(least squares method)是直线拟合常用的方法,在自然科学和社会科学内被广泛应用,尤其在同位素地质年代学领域更是必不可少。仅考虑 x 或 y 误差的普通最小二乘法(normal/ordinary least squares)广为人知,但事实上最小二乘法并非如此简单,尤其是在同时考虑 x 、 y 误差(乃至误差相关)并采用加权处理时,其数学处理方法会变得十分复杂,而此时普通最小二乘法显得极不合理。本文在前人研究的基础上,结合同位素地质年龄计算的需要,对直线拟合的最小二乘法进行了系统地总结研究,详细介绍了普通最小二乘法、加权普通最小二乘法(normal/ordinary weighting least squares)、标准加权最小二乘法(standard weighting model)、标准独立加权最小二乘法(standard independent weighting model)、独立加权最小二乘法(independent weighting model)及误差相关最小二乘法(error correlated independent weighting model)的数学原理及相关变量的计算过程。在此基础上,进一步阐述了 $MSWD$ (mean squared weighted deviation)这个同位素地质年代学中经常使用的参数的数学意义,以及 $MSWD$ 对计算结果的评判意义。准确理解这些数学方法,对于我们合理选择同位素地质年龄相关参数的计算方法,科学评价直线拟合结果并探讨其地质意义至关重要。我们的研究同时有助于拓展和完善数学领域最小二乘法的基本理论,并可用于其他领域相似的研究之中。

关键词:最小二乘法;直线拟合;同位素年代学; $MSWD$;迭代法

同位素地质年代学研究是确定地质体绝对年龄的主要手段,包括 Rb-Sr、Sm-Nd、Re-Os、U-Th-Pb、K-Ar(Ar-Ar)等测年方法,在当前地质学领域内得到了广泛应用。在同位素地质年龄的计算中,常常需要根据一组样品的同位素测试结果进行直线拟合,最常见的是等时线拟合,主要有 Rb-Sr、Sm-Nd、K-Ar、Re-Os、Pb-Pb 等等时线。除此之外, U-Pb 不一致年龄的计算,以及利用 ^{26}Al - ^{26}Mg 、 ^{129}I - ^{129}Xe 等绝灭了的放射性(extinct radioactivity)同位素方法研究陨石的相对年龄时,也要用到直线拟合。事实上,在地球科学内的诸多领域,都用到了直线拟合。

直线拟合常用的是最小二乘法,其在自然科学和社会科学领域内得到了广泛的应用。其所要解决的问题是:根据两个理论上存在线性关系的变量的几组测试数据点(多数情况下,这些数据点不见得恰好位于理论直线之上),寻找这两个变量之间的线性函数关系。当这些测试数据点与拟合直线的“偏差平方和”最小时,得到最佳拟合直线,即为所

求。通常情况下,直线拟合多数采用最简单的普通最小二乘法,这也是诸多数学书籍中介绍最多、并被多数人所熟知的直线拟合方法(例如中国科学院数学研究所数理统计组,1974;张启锐,1988;周复恭和黄运成,1989;张小蒂,1991;周纪芄,1993;吴翊等,1995;李庆阳,2001;同济大学应用数学系,2002;盛聚等,2008;周惟公,2008;王黎明等,2008;褚宝增和王翠香,2010;何晓群和刘文卿,2011)。这种方法隐含地假设变量 x 不存在误差,并将所有误差都归于变量 y (类似,可将所有误差都归于变量 x),即便在 x 和 y 同为测量值的情况下,此时普通最小二乘法在确定未知参数及估计其真实误差方面,失去了准确性(Borcherds and Sheth, 1995; Sheth et al., 1996)。

仅考虑 x 或 y 误差的普通最小二乘法广为人知,但实际上最小二乘法并非如此简单,尤其是在同时考虑 x 、 y 误差(乃至误差相关)并采用加权处理时,其数学处理方法会变得十分复杂,而这对于许多

注:本文为国家自然科学基金资助项目(编号40973010)的成果。

收稿日期:2012-11-15;改回日期:2013-06-25;责任编辑:章雨旭。

作者简介:范文博,男,1987年生。在读硕士研究生。目前主要从事地球化学、岩石学的学习和研究。Email: fans07cugb@126.com。

应用同位素地质年代学方法进行地质学研究的人来说,都十分陌生。之所以会如此,主要有两方面的原因:① 数学处理方法的复杂性以及相关资料的相对缺乏;② 采用已有的软件(Isoplot)而“无需”思考其数学原理。

虽然同位素地质年代学方法被广泛应用,但地质学研究者多热衷于年龄结果本身的地质意义,同时由于同时跨越了数学和地质学两个学科,因此很少有人同时结合地质需要、对最小二乘法的数学原理作系统深入的总结介绍。如前所述,许多数学书籍也只涉及最简单的普通最小二乘法,何晓群和刘文卿(2011)、费业泰(2010)对考虑加权的最小二乘法也只是做了简单介绍。李志昌等(2004)介绍了York(1966,1969)、McIntyre等(1966)算法的一部分,而更多的同位素书籍(韩吟文和马振东,2003;陈骏和王鹤年,2004;陈岳龙等,2005;陈道公等,2009),则忽略具体算法而只强调MSWD(见下文介绍)对等时线拟合结果评价的用处。事实上,想要真正了解MSWD对拟合结果的评判意义,就必须对最小二乘法本身有所把握。国外一些学者在前人[如Deming,1943;据York(1966),Borcherds and Sheth(1995)所述]研究的基础上,对最小二乘法的数学算法及相关问题作了进一步的研究和探索(如McIntyre et al.,1966;York,1966,1967,1969;Brooks et al.,1972;Titterton and Halliday,1979;Borcherds and Sheth,1995;Sheth et al.,1996;Mahon,1996;Bruzzone and Moreno,1998),取得了突破性进展。但这些研究各有新的尝试和侧重,同时难免存在着一些分歧、甚至错误,因此想要系统了解最小二乘法并非易事。Ludwig(2008)开发的Excel加载宏Isoplot,提供了同位素年龄计算的一个很好的平台,其以York(1969)、Titterton and Halliday(1979)、Brooks等(1972)等的算法(Ludwig,2008)为基础,但未对具体数学方法作详细介绍。这样的话,虽然Isoplot被广泛使用,但相关同位素年龄的计算方法,对多数地质学研究者来说仍然是一个“谜”。正如Allegre(2008)所述,如今相关测试方法已经有了很大的进步,而理论模型——或许是统计模型应该跟上分析方法前进的步伐,这点十分重要,但事实并非如此,这使得我们不得不思考如何对待通过自动化的分析方法已经积累的大量数据。可见,将最小二乘法这个在同位素地质年代学内普遍应用的统计方法“显性化”,是非常有意义的,这也有助于解决实际研究中存在的一些困扰,例

如如何理解MSWD对年龄结果的评价及参考意义。

鉴于此,以下将在前人研究的基础上,结合同位素地质年龄计算的相关知识,对直线拟合的最小二乘法作详细总结和研究,并进一步在此基础上讨论同位素地质年代学应用的相关问题。

1 相关测年原理简介

首先以Rb-Sr等时线法为例,介绍应用最小二乘法所要解决的地质问题。核素 ^{87}Rb 通过一次 β^- 衰变变成稳定同位素 ^{87}Sr ,根据放射性衰变定律,进一步推演得到(Faure,1986):

$$n(^{87}\text{Sr}) = n(^{87}\text{Sr})_0 + n(^{87}\text{Rb})(e^{\lambda t} - 1)$$

式中, λ 为 ^{87}Rb 的衰变常数, t 为从同位素体系封闭到现在的时间,亦即所求的地质年龄。 $n(^{87}\text{Sr})$ 、 $n(^{87}\text{Sr})_0$ 和 $n(^{87}\text{Rb})$ 均为物质的量。

两边同时除以 $n(^{86}\text{Sr})$,得(Faure,1986):

$$\frac{n(^{87}\text{Sr})}{n(^{86}\text{Sr})} = \left[\frac{n(^{87}\text{Sr})}{n(^{86}\text{Sr})} \right]_0 + \left[\frac{n(^{87}\text{Rb})}{n(^{86}\text{Sr})} \right] (e^{\lambda t} - 1)$$

此式为Rb-Sr同位素定年的基础,其中 $\frac{n(^{87}\text{Rb})}{n(^{86}\text{Sr})}$ 、

$\frac{n(^{87}\text{Sr})}{n(^{86}\text{Sr})}$ 为现今样品的实测同位素含量比值,

$\left[\frac{n(^{87}\text{Sr})}{n(^{86}\text{Sr})} \right]_0$ 为距今 t 时刻时这两种同位素的初始比

值。由于 $\left[\frac{n(^{87}\text{Sr})}{n(^{86}\text{Sr})} \right]_0$ 未知,因此不能直接计算年龄。

对于一套同源、同时、封闭的样品,其 $\frac{n(^{87}\text{Sr})}{n(^{86}\text{Sr})}$ 初始值

均一,也就是不同样品间的 $\left[\frac{n(^{87}\text{Sr})}{n(^{86}\text{Sr})} \right]_0$ 一致,而其距今 t 时刻时的 ^{87}Rb 含量可能不同,从而导致放射性

成因累积的 ^{87}Sr 含量不同。这样的话,这一套样品内的不同样品之间,

$\frac{n(^{87}\text{Rb})}{n(^{86}\text{Sr})}$ 、 $\frac{n(^{87}\text{Sr})}{n(^{86}\text{Sr})}$ 就会有所差异。

这时, t 、 λ 为常数,上式便可看作直线方程:

$$y = a + xb \quad (\text{Faure, 1986})$$

这里 $b = e^{\lambda t} - 1$, $y = \frac{n(^{87}\text{Rb})}{n(^{86}\text{Sr})}$, $x = \frac{n(^{87}\text{Sr})}{n(^{86}\text{Sr})}$ 。

对于上述一组样品,根据其 $\left[\frac{n(^{87}\text{Rb})}{n(^{86}\text{Sr})}, \frac{n(^{87}\text{Sr})}{n(^{86}\text{Sr})} \right]$ 测试值及其误差,运用最小二

乘法拟合直线,就能够得到其斜率与截距。根据斜率 $b = e^{\lambda t} - 1$,就可得到所关注的年龄 t ,而其截距

则为 $\left[\frac{n(^{87}\text{Sr})}{n(^{86}\text{Sr})} \right]_0$,可用于同位素地球化学示踪。

除 Rb-Sr 同位素体系之外,Sm-Nd、K-Ar、Re-Os 同位素体系定年也常用到等时线法,其同位素定年方程分别为 (Faure, 1986) :

$$\begin{aligned} \frac{n(^{143}\text{Nd})}{n(^{144}\text{Nd})} &= \left[\frac{n(^{143}\text{Nd})}{n(^{144}\text{Nd})} \right]_0 + \left[\frac{n(^{147}\text{Sm})}{n(^{144}\text{Nd})} \right] (e^{\lambda t} - 1) \\ \frac{n(^{187}\text{Os})}{n(^{188}\text{Os})} &= \left[\frac{n(^{187}\text{Os})}{n(^{188}\text{Os})} \right]_0 + \left[\frac{n(^{187}\text{Re})}{n(^{188}\text{Os})} \right] (e^{\lambda t} - 1) \\ \frac{n(^{40}\text{Ar})}{n(^{36}\text{Ar})} &= \left[\frac{n(^{40}\text{Ar})}{n(^{36}\text{Ar})} \right]_0 + \left[\frac{\lambda_e}{\lambda} \cdot \frac{n(^{40}\text{K})}{n(^{36}\text{Ar})} \right] (e^{\lambda t} - 1) \end{aligned}$$

对于 U-Pb 同位素体系,最小二乘法直线拟合主要用于 Pb-Pb 等时线拟合以及 U-Pb 谐和图中的不一致线拟合。前者的斜(率)截(距)式直线方程为 (Rosholt et al., 1973; Faure, 1986; 陈道公, 2009) :

$$\begin{aligned} \frac{n(^{207}\text{Pb})}{n(^{204}\text{Pb})} &= \left\{ \left[\frac{n(^{207}\text{Pb})}{n(^{204}\text{Pb})} \right]_0 - m \left[\frac{n(^{206}\text{Pb})}{n(^{204}\text{Pb})} \right]_0 \right\} + \\ &\frac{n(^{206}\text{Pb})}{n(^{204}\text{Pb})} \cdot m, \text{ 其中 } m = \frac{1}{137.8} \cdot \frac{e^{\lambda_{235}t} - 1}{e^{\lambda_{238}t} - 1} \end{aligned}$$

后者的点斜式直线方程为 (Wetherill, 1956; Allegre, 2008) :

$$\frac{\frac{n(^{206}\text{Pb})}{n(^{238}\text{U})} - (e^{\lambda_{238}T_1} - 1)}{\frac{n(^{207}\text{Pb})}{n(^{235}\text{U})} - (e^{\lambda_{235}T_1} - 1)} = \frac{e^{\lambda_{238}t} - e^{\lambda_{238}T_1}}{e^{\lambda_{235}t} - e^{\lambda_{235}T_1}}$$

在应用绝灭了放射性同位素对陨石进行相对

年龄测定时,同样需要首先用到直线拟合,如 $^{26}\text{Al}-^{26}\text{Mg}$ 、 $^{129}\text{I}-^{129}\text{Xe}$ 陨石测年,其对应直线方程分别为 (Allegre, 2008) :

$$\begin{aligned} \frac{n(^{26}\text{Mg})}{n(^{24}\text{Mg})} &= \left[\frac{n(^{26}\text{Mg})}{n(^{24}\text{Mg})} \right]_0 + \frac{n(^{27}\text{Al})}{n(^{24}\text{Mg})} \cdot \left[\frac{n(^{26}\text{Al})}{n(^{24}\text{Al})} \right]_i \\ \frac{n(^{129}\text{Xe})}{n(^{130}\text{Xe})} &= \left[\frac{n(^{129}\text{Xe})}{n(^{130}\text{Xe})} \right]_0 + \\ &k \cdot \frac{n(^{128}\text{Xe})}{n(^{130}\text{Xe})} \cdot \left[\frac{n(^{129}\text{I})}{n(^{127}\text{I})} \right]_i \end{aligned}$$

综上可知,这里所要解决的问题是,根据样品的一组同位素比值测试数据(往往带有实验误差) $(x_i, y_i), i = 1, 2, 3, \dots, n$, 选择最优的最小二乘法处理模型,拟合最佳直线并计算斜率、截距及其误差,并进一步根据拟合结果,尽可能地甄别、评判其是否有真正的地质年代意义。

2 直线的最小二乘法拟合

现给出样品的一组同位素比值测试数据点 $(x_i, y_i), i = 1, 2, 3, \dots, n (n > 2)$, 据此求自变量 x 和因变量 y 之间的函数关系 $L: y = a + xb$ 。其中 a 、 b 为待定参数,当拟合直线与数据点之间的(加权)偏差平方和 S 最小时,得到最佳拟合直线。关于偏差平方和,根据 York (1969), 有两种等价的最一般形式:

$$S = \sum_{i=1}^n \left[\frac{w(x_i) (x_i - X_i)^2 - 2\rho_i \cdot \sqrt{w(x_i)w(y_i)} \cdot (x_i - X_i)(y_i - Y_i) + w(y_i) \cdot (y_i - Y_i)^2}{(1 - \rho_i^2)} \right] \tag{1}$$

$$S = \sum_{i=1}^n w_i (y_i - a - bx_i)^2 \tag{2}$$

这里 (x_i, y_i) 是第 i 个测试值, (X_i, Y_i) 是对应的调整值(图 1), 即最佳拟合直线上对应的点。

$w(x_i)$ 、 $w(y_i)$ 分别是第 i 个样品 x 、 y 测试值的权,通常用对应测试值的不确定度(取标准差)的平方倒数给出,即 $w(x_i) = \frac{1}{\sigma^2(x_i)}$ 、 $w(y_i) = \frac{1}{\sigma^2(y_i)}$ (McIntyre et al., 1966; York, 1966, 1967, 1969; Brooks et al., 1972; Titterton and Halliday, 1979; G. 福尔, 1983; Borchers and Sheth, 1995; Sheth et al., 1996; 费业泰, 2010), $\sigma(x_i)$ 、 $\sigma(y_i)$ 为第 i 个样品给定的 x_i 、 y_i 的标准差, w_i 为其有效权:

$$w_i = \frac{w(x_i)w(y_i)}{b^2w(y_i) + w(x_i) - 2b\rho_i \sqrt{w(x_i)w(y_i)}} \text{ 或}$$

$$w_i = \frac{1}{\sigma^2(y_i) + b^2\sigma^2(x_i) - 2b\rho_i\sigma(x_i)\sigma(y_i)} \tag{3}$$

ρ_i 为第 i 个点 x 和 y 变量的误差相关系数(error correlation, 详见误差相关最小二乘法部分), 对于 U-Th-Pb 测年, 由于纵横坐标变量 x_i 、 y_i 是同时同步测试的, 往往需要考虑误差相关系数, 而对于其它测年方法, 在实际应用中多不考虑这个参数。

当 $\rho_i = 0$ 时, x 、 y 的误差不相关, 这时得到 York (1966)、Borchers 和 Sheth (1995) 中所提到的 Deming (1943) 所定义的偏差平方和的公式, 即与 (1) 式对应的 (4) 式, 以及 McIntyre 等 (1966)、Wendt (1969) [转引自 Brooks, 1972]、Borchers 和 Sheth (1995)、Sheth 等 (1996)、何晓群和刘文卿 (2011) 所用的与 (2) 对应的 (5) 式,

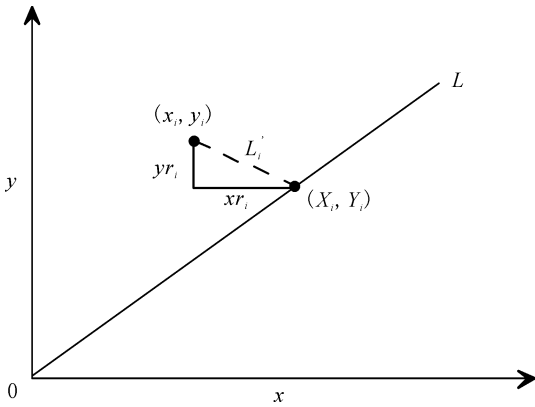


图 1 最小二乘拟合中,相关量之间的一般几何关系
Fig. 1 Illustration of relationship between various variables

in least squares fitting of a straight line

L 为最佳拟合直线, L' 为调整线, (x_i, y_i) 为测试点, (X_i, Y_i) 为

对应调整点, y_{ri} 表示 y 方向的残差, x_{ri} 表示 x 方向的残差

L is the best straight line, L' is the adjustment line, (x_i, y_i) is the data point while (X_i, Y_i) is the adjustment point corresponding. y_{ri} denotes residual in y direction, and x_{ri} in x direction

$$S = \sum_{i=1}^n [w(x_i) (x_i - X_i)^2 + w(y_i) (y_i - Y_i)^2] \quad (4)$$

$$S = \sum_{i=1}^n w_i (y_i - a - bx_i)^2 \quad (5)$$

$$w_i = \frac{w(x_i)w(y_i)}{w(x_i) + b^2w(y_i)} \text{ 或 } w_i = \frac{1}{\sigma^2(y_i) + b^2\sigma^2(x_i)} \quad (6)$$

图 1 给出了拟合直线与相关量之间的一般几何关系。 L 为最佳拟合直线 $y = a + xb$, 连结测试点 (x_i, y_i) 与 (X_i, Y_i) 的直线 L'_i 为调整线, 其斜率为每个点的 y 方向残差 (y_i -residual, 以下用 y_{ri} 表示) 与 x 方向残差 (x_i -residual, 以下用 x_{ri} 表示) 之比, 即:

$$L'_i \text{ 的斜率} = \frac{y_{ri}}{x_{ri}} = \frac{y_i - Y_i}{x_i - X_i} \quad (7)$$

可以看出, 随着拟合直线 (即参数 a, b) 的变化, S 也会相应地发生变化, 当 S 取最小值时, 所有点沿各自调整线回归到拟合直线上的距离之和最短, 这时就得到了最佳拟合直线。 方便起见, 这里定义最佳拟合直线所对应的 S 的最小值为 S^* 。 以下从最简单的普通最小二乘法开始, 逐步深入, 分别介绍基于不同假设所得到的最小二乘直线拟合方法。 实际应用中, 则需要根据数据本身特征, 选择最适合的最小二乘法进行直线拟合。

2.1 普通最小二乘法 (normal/ordinary least squares)

顾名思义, 这是介绍并使用最多的最小二乘拟合方法 (例如吴翊等, 1995; 同济大学应用数学系, 2002; 李庆阳, 2001; 盛聚等, 2008; 王黎明等, 2008; 褚宝增和王翠香, 2010; 何晓群和刘文卿, 2011), 在 Isoplot 被广泛使用之前, 等时线拟合多用这种方法 [如申浩澈 (1983) 所述]。 假设 x_i 不存在误差 ($x_i = X_i$), y_i 存在误差 ($y_i \neq Y_i$), 令 $w_i = w(y_i) = 1$ (可认为 $w(x_i) = \infty$), 这时所有数据点沿与 y 轴平行的调整线调整到最佳拟合直线 (图 2a),

$$S(a, b) = \sum_{i=1}^n [y_i - (a + bx_i)]^2 \quad (8)$$

对于固定的一组数据, (8) 式是 S 关于 a, b 的二元函数, 根据多元函数的极值条件 (同济大学应用数学系, 2002), 分别对 S 求关于 a, b 的偏导数, 并令其为 0, 得:

$$\frac{\partial S}{\partial a} = -2 \sum_{i=1}^n (y_i - a - bx_i) = 0,$$

$$\frac{\partial S}{\partial b} = -2 \sum_{i=1}^n x_i (y_i - a - bx_i) = 0$$

进一步变换得到所谓的“法方程”或“正规方程”,

$$na + b \sum_{i=1}^n x_i = \sum_{i=1}^n y_i \quad (9)$$

$$a \sum_{i=1}^n x_i + b \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i y_i \quad (10)$$

据 (9)、(10), 解得 (吴翊等, 1995; 王黎明等, 2008; 盛聚等, 2008; 何晓群和刘文卿, 2011):

$$b = \frac{n \sum_{i=1}^n x_i y_i - (\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}$$

$$= \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (11)$$

$$a = \bar{y} - b\bar{x} \quad (12)$$

这里 $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ 。 由 (9) 或 (12) 可知, 最佳拟合直线过所有点的重心, 这点对以下其它最小二乘法也都成立。

与之类似, 也可将所有的误差都归于 x_i ($x_i \neq X_i, y_i = Y_i$), 令 $w(x_i) = 1, w_i = w(x_i)/b^2 = 1/b^2$ (可认为 $w(y_i) = \infty$), 进行处理, 如图 2b。

对于一些研究对象, 普通最小二乘法假设一个

变量不存在误差,这是成立的。而对于同位素地质年代学,由于所有 x_i, y_i 都是测试值,这点显然不成立。笔者认为,对于 Rb-Sr 等时线法,考虑到 $n(^{87}\text{Rb})/n(^{86}\text{Sr})$ 的测试误差往往比 $n(^{87}\text{Sr})/n(^{86}\text{Sr})$ 的误差大许多,因此采用将所有误差归于 $n(^{87}\text{Rb})/n(^{86}\text{Sr})$ 的(加权)普通最小二乘法,或许是可以接受的。Sm-Nd 等时线法类似。

2.2 加权普通最小二乘法

(normal/ordinary weighting least squares)

这是对普通最小二乘法的改进,与前者不同的是其给予每个点不同的权重。假设 x_i 不存在误差($x_i = X_i$), y_i 存在误差($y_i \neq Y_i$),这时 $w_i = w(y_i) = 1/\sigma^2(y_i)$ (可认为 $w(x_i) = \infty$),所有调整线都平行于 y 轴(图 2a),

$$S(a, b) = \sum_{i=1}^n \frac{1}{\sigma^2(y_i)} [y_i - (a + bx_i)]^2 \quad (13)$$

同理,可以根据 $\partial S/\partial a = 0, \partial S/\partial b = 0$,解得(Bruzzone and Moreno, 1998):

$$\begin{cases} b = \frac{SS_{xy} - S_x S_y}{SS_{xx} - S_x^2} \\ a = \bar{y} - b\bar{x} = \frac{S_{xx} S_y - S_x S_{xy}}{SS_{xx} - S_x^2} \end{cases} \quad (14)$$

其中: $S = \sum_{i=1}^n w_i, S_x = \sum_{i=1}^n w_i x_i, S_y = \sum_{i=1}^n w_i y_i,$

$$S_{xx} = \sum_{i=1}^n w_i x_i^2, S_{yy} = \sum_{i=1}^n w_i y_i^2, S_{xy} =$$

$$\sum_{i=1}^n w_i x_i y_i, \bar{x} = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i}, \bar{y} = \frac{\sum_{i=1}^n w_i y_i}{\sum_{i=1}^n w_i}, w_i =$$

$$w(y_i) = \frac{1}{\sigma^2(y_i)}.$$

对(14)式中 b 的表达式进一步变换可得(15)

$$b = \frac{\sum (y_i - \bar{y})^2 - \sum (x_i - \bar{x})^2 + \sqrt{[\sum (y_i - \bar{y})^2 - \sum (x_i - \bar{x})^2]^2 + 4[\sum (x_i - \bar{x})(y_i - \bar{y})]^2}}{2 \sum (x_i - \bar{x})(y_i - \bar{y})} \quad (19)$$

就是所求斜率,其中 $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ 。

a 的解法同上,由最佳拟合直线过重心计算得到[(17)式,与(12)式本质相同]。

2.4 标准独立加权最小二乘法

(standard independent weighting model)

Sheth 等(1996)中将 $\sigma(x_1) = \sigma(x_2) = \dots =$

式:

$$b = \frac{\sum_{i=1}^n w_i (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n w_i (x_i - \bar{x})^2} \quad (15)$$

类似,也可将所有的误差都归于 $x_i (x_i \neq X_i, y_i = Y_i)$ 并采用加权处理,令 $w(x_i) = 1/\sigma^2(x_i), w_i = w(x_i)/b^2$ (可认为 $w(y_i) = \infty$),进而确定最佳拟合直线(如图 2b)。

2.3 标准加权最小二乘法

(standard weighting model)

普通最小二乘法以及加权普通最小二乘法的一个显著缺陷是只考虑一个变量的误差,在同位素年代学研究中,两个变量都是测试值,从而不可避免都存在误差,这显然不够精确。

标准加权模型是双误差最小二乘法的一种特殊情形,它同时考虑两个变量的误差,但假设 $w(x_i) = w(y_i) = 1$ (这意味着两个坐标的误差大小接近一致), $w_i = \frac{1}{1 + b^2}$,这时所有调整线都相互平行且垂直于最佳拟合直线(图 2c),这时

$$S = \frac{1}{1 + b^2} \sum_{i=1}^n (y_i - a - bx_i)^2 \quad (16)$$

同上,令 $\partial S/\partial a = 0$,得:

$$a = \frac{1}{n} \left(\sum_{i=1}^n y_i - b \sum_{i=1}^n x_i \right) = \bar{y} - b\bar{x} \quad (17)$$

令 $\partial S/\partial b = 0$,得:

$$\begin{aligned} & \frac{1}{1 + b^2} \sum_{i=1}^n x_i (y_i - a - bx_i) + \\ & \frac{b}{(1 + b^2)^2} \sum_{i=1}^n (y_i - a - bx_i)^2 = 0 \end{aligned} \quad (18)$$

将(17)式代入(18),化简会得到关于 b 的一元二次方程,解之并取其中一个解,得到:

$\sigma(x_i) = \dots = \sigma(x), \sigma(y_1) = \sigma(y_2) = \dots = \sigma(y_i) = \dots = \sigma(y)$ 且 $\sigma(x) \neq \sigma(y)$ 的情形称作标准独立加权,即对于所有的样品, x 变量的误差一致, y 变量的误差也一致,但 x 和 y 的误差不等。这时, $w(x_i) = 1/\sigma^2(x) = w_1, w(y_i) = 1/\sigma^2(y) = w_2, w = \frac{w_1 w_2}{w_1 + b^2 w_2}$,所有调整线都互相平行(图 2d),

$$S = \frac{w_1 w_2}{w_1 + b^2 w_2} \sum_{i=1}^n (y_i - a - b x_i)^2 \tag{20}$$

令 $\partial S/\partial a = 0, \partial S/\partial b = 0$, 同上解得:

$$b = \frac{\sum (y_i - \bar{y})^2 - \frac{w_1}{w_2} \sum (x_i - \bar{x})^2 + \sqrt{\left[\sum (y_i - \bar{y})^2 - \frac{w_1}{w_2} \sum (x_i - \bar{x})^2 \right]^2 + 4 \frac{w_1}{w_2} \left[\sum (x_i - \bar{x})(y_i - \bar{y}) \right]^2}}{2 \sum (x_i - \bar{x})(y_i - \bar{y})} \tag{21}$$

其中 $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$, 其它同 2.3 节。

的情况—— x_i 和 y_i 的不确定度(标准差) 互相成比例, 即 $\sigma(x_i) = c\sigma(y_i), c$ 为常数, 这时 $w(x_i) =$

Bruzzone 和 Moreno (1998) 讨论了一种更一般

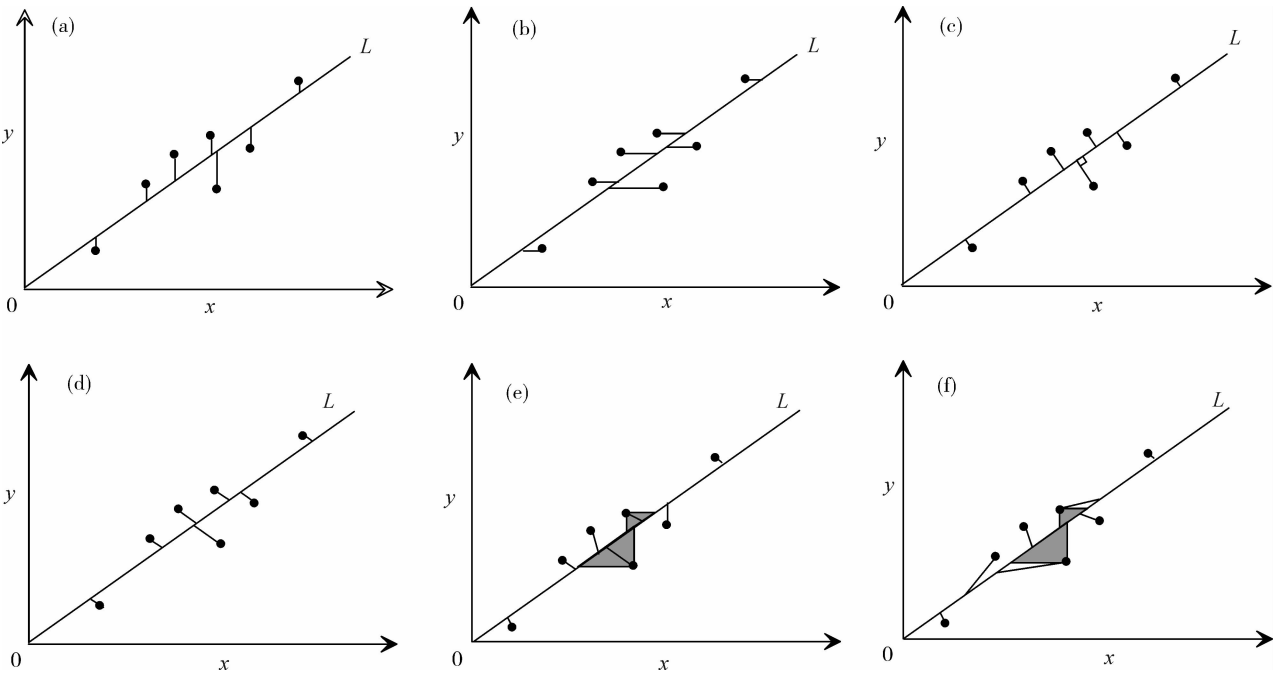


图 2 不同类型的最小二乘拟合方法中,数据点的调整方式:(a) (加权的)普通最小二乘模型,只考虑纵坐标的误差,所有调整线都平行于 y 轴,需要说明的是,对于同一组数据点,加权的普通最小二乘拟合与普通最小二乘拟合所拟合直线可能会不同;(b) (加权的)普通最小二乘模型,只考虑横坐标的误差;(c) 标准加权模型,所有调整线都互相平行且垂直于最佳拟合直线;(d) 标准独立加权模型,所有调整线都互相平行;(e) 独立加权模型,一般情况下,不同数据点的调整线不平行,但限制在三角阴影区内;(f) 误差相关的加权最小二乘拟合模型,同样不同数据点的调整线不再平行,但调整线不再限制在三角阴影区内(参考 York, 1966, 1967, 1969; Borchers and Sheth, 1995; Sheth et al., 1996; Mahon, 1996; Bruzzone and Moreno, 1998,有增改)

Fig. 2 Various types of adjustments in different models of least squares fitting: (a) normal/ordinary least squares and normal/ordinary weighting least squares, only the errors in y direction are considered. It should be noted that weighting least squares may have different results from the not-weighting in this situation; (b) normal/ordinary least squares and normal/ordinary weighting least squares, only the errors in x direction are considered; (c) standard weighting model, all the adjusting lines are parallel to each other and are perpendicular to the best fitting line at the same time; (d) standard independent weighting model, the adjusting lines are parallel to each other; (e) independent weighting model, the adjusting lines are not necessary to be parallel but should be situated within the shadow area; (f) error correlated independent weighting model, the adjusting lines are not necessary to be parallel or be restricted within the shadow area(after York, 1966, 1967, 1969; Borchers and Sheth, 1995; Sheth et al., 1996; Mahon, 1996; Bruzzone and Moreno, 1998)

$\frac{1}{\sigma^2(x_i)} = \frac{1}{c^2\sigma^2(y_i)}, w(y_i) = \frac{1}{\sigma^2(y_i)}$, 所有数据点也都沿着相互平行的调整线调整到最佳拟合直线(图 2d), 且

$$S = \frac{1}{1 + b^2c^2} \sum_{i=1}^n \frac{(y_i - a - bx_i)^2}{\sigma^2(y_i)} \tag{22}$$

S 的最小化给出关于 b 的二次方程, 解之并选取其中一个合适的根作为拟合直线的斜率, 得到 (Bruzzone and Moreno, 1998):

$$\begin{aligned} 2cb &= -\frac{1}{cb^*} + \frac{cb^*}{r^2} + \\ &\frac{1}{cb^*} \cdot \sqrt{\left[1 - \frac{(cb^*)^2}{r^2}\right]^2 + 4(cb^*)^2} \tag{23} \\ r^2 &= \frac{(SS_{xy} - S_xS_y)^2}{(SS_{xx} - S_x^2)(SS_{yy} - S_y^2)} \end{aligned}$$

这里 b^* 指(14) 中的 b , 即对于此组数据, 假若 x_i 没有误差时计算所得的斜率 $S_x, S_x, S_y, S_{xy}, S_{xx}, S_{yy}$ 如 2.2 节中所述。

2.5 独立加权最小二乘法
(independent weighting model)

这是双变量最小二乘拟合的更一般情形, 其同时考虑每个点 x_i, y_i 的不确定度, 并认为 $\sigma(x_i), \sigma(y_i)$ 二者之间不存在确定的数学关系。与上述 2.1~2.4 节的情形不同, 此时调整线一般不再平行, 并被限制在假设 x, y 都不存在误差时的竖直和水平调整线以及最佳直线所包围的三角形内(图 2e), 而且得不到斜率 b 和截距 a 的解析解, 必须使用迭代法, 寻找认为达到我们所需要精度的近似解。下面给出两种处理办法。

2.5.1 独立加权模型解法一

York (1966) 从表达式(4) 着手解决这个问题, 延续其思路, 以下进一步更详细地解析相关过程:

2.5.1.1 步骤 1

$$S = \sum_{i=1}^n [w(x_i)(x_i - X_i)^2 + w(y_i)(y_i - Y_i)^2]$$

S 是关于 X_i, Y_i 的泛函, 根据泛函极值条件及相关运算法则, 当 S 取得极值时, 其一阶变分为 0, 即有:

$$\begin{aligned} \delta_S &= - \sum_{i=1}^n 2[w(x_i)(x_i - X_i)\delta(X_i) + \\ &\sum_{i=1}^n 2[w(y_i)(y_i - Y_i)\delta(Y_i)] = 0 \tag{24} \end{aligned}$$

$$\begin{aligned} &\sum_{i=1}^n w(x_i)(x_i - X_i)\delta(X_i) + \\ &\sum_{i=1}^n w(y_i)(y_i - Y_i)\delta(Y_i) = 0 \tag{25} \end{aligned}$$

这里 $\delta_s, \delta(x_i), \delta(y_i)$ 分别表示 S, x_i, y_i 的一阶变分 [此概念来自“变分法”这一数学分支学科, 而变分法则是研究求泛函极大值和极小值的方法, 简单理解, “一阶变分”类似于高等数学中的微分, 有关其详细的数学定义可参考老大中 (2004)、邵克勇等 (2011)]。

由于 (X_i, Y_i) 位于最佳拟合直线上, 因此有:

$$Y_i = a + bX_i \tag{26}$$

两边同时求变分, 得到:

$$\delta_a + X_i\delta_b + b\delta(X_i) - \delta(Y_i) = 0 \tag{27}$$

对于不同的 i , (27) 式乘以各自的一个未定乘数 λ_i , 并求和, 得到:

$$\begin{aligned} \delta_a \sum_{i=1}^n \lambda_i + \delta_b \sum_{i=1}^n (\lambda_i X_i) + b \sum_{i=1}^n [\lambda_i \delta(X_i)] \\ - \sum_{i=1}^n [\lambda_i \delta(Y_i)] = 0 \tag{28} \end{aligned}$$

(25) + (28), 得:

$$\begin{aligned} &\sum_{i=1}^n \{[w(x_i)(x_i - X_i) + b\lambda_i]\delta(X_i)\} + \\ &\sum_{i=1}^n \{[w(y_i)(y_i - Y_i) - \lambda_i]\delta(Y_i)\} + \delta_a \sum_{i=1}^n \lambda_i + \\ &\delta_b \sum_{i=1}^n (\lambda_i X_i) = 0 \tag{29} \end{aligned}$$

令 $\delta(X_i), \delta(Y_i), \delta_a, \delta_b$ 的系数为 0, 得到:

$$x_i - X_i = \frac{-b\lambda_i}{w(x_i)} \tag{30}$$

$$y_i - Y_i = \frac{\lambda_i}{w(y_i)} \tag{31}$$

$$\sum_{i=1}^n \lambda_i = 0 \tag{32}$$

$$\sum_{i=1}^n (\lambda_i X_i) = 0 \tag{33}$$

将(30)、(31)分别代入(26)中, 分别替换 X_i, Y_i , 得到:

$$y_i - \frac{\lambda_i}{w(y_i)} = a + b\left[x_i + \frac{b\lambda_i}{w(x_i)}\right] \tag{34}$$

由(34)式解出 λ_i , 得到:

$$\lambda_i = w_i(y_i - a - bx_i) \tag{35}$$

w_i 由(6)式给出。

2.5.1.2 步骤 2

现在, (32)、(33)、(35)一共有 $(n+2)$ 个等式, $(n+2)$ 个未知数: a, b 和 λ_i 。根据(32)和(35), 有:

$$\sum_{i=1}^n w_i(y_i - a - bx_i) = 0 \tag{36}$$

根据(33)和(35), 有:

$$\sum_{i=1}^n w_i (y_i - a - bx_i) X_i = 0 \quad (37)$$

将(30)代入(37), 替换 X_i , 得到:

$$\sum_{i=1}^n w_i (y_i - a - bx_i) \left[x_i + \frac{b\lambda_i}{w(x_i)} \right] = 0 \quad (38)$$

进一步化简, 有:

$$\sum_{i=1}^n [w_i (y_i x_i - ax_i - bx_i^2)] + \sum_{i=1}^n \left[w_i (y_i - a - bx_i) \frac{b\lambda_i}{w(x_i)} \right] = 0 \quad (39)$$

将(35)式代入(39)式, 替换 λ_i 得到:

$$\sum_{i=1}^n [w_i (y_i x_i - ax_i - bx_i^2)] + \sum_{i=1}^n \left[w_i^2 (y_i - a - bx_i)^2 \frac{b}{w(x_i)} \right] = 0 \quad (40)$$

这里可以通过(36)、(40)同时解出 a 、 b 。

2.5.1.3 步骤 3

现在, 从等式(36)中解出 a , 得到[与(12)式形式相同]:

$$a = \frac{\sum_{i=1}^n w_i y_i - b \sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i} = \bar{y} - b\bar{x} \quad (41)$$

这里

$$\bar{y} = \frac{\sum_{i=1}^n (w_i y_i)}{\sum_{i=1}^n w_i}, \quad \bar{x} = \frac{\sum_{i=1}^n (w_i x_i)}{\sum_{i=1}^n w_i} \quad (42)$$

将(41)式代入(40)式, 化简即得到 York (1966) 的“最小二乘立方方程”。事实上, 正如 Mahon (1996) 所认识到的, 由于 w_i 中也含有 b , 因此这也只是一个形式上的立方方程:

$$b^3 \sum_{i=1}^n \frac{w_i^2 (x_i - \bar{x})^2}{w(x_i)} - 2b^2 \sum_{i=1}^n \frac{w_i^2 (x_i - \bar{x}) (y_i - \bar{y})}{w(x_i)} - b \left[\sum_{i=1}^n w_i (x_i - \bar{x})^2 - \sum_{i=1}^n \frac{w_i^2 (y_i - \bar{y})^2}{w(x_i)} \right] + \sum_{i=1}^n w_i (x_i - \bar{x}) (y_i - \bar{y}) = 0 \quad (43)$$

(43)式中, 左边第一项系数化为 1, 得:

$$b^3 - 3\alpha b^2 + 3\beta b - \gamma = 0 \quad (44)$$

其中

$$\alpha = \frac{2 \sum_{i=1}^n \frac{w_i^2 (x_i - \bar{x}) (y_i - \bar{y})}{w(x_i)}}{3 \sum_{i=1}^n \frac{w_i^2 (x_i - \bar{x})^2}{w(x_i)}} \quad (45)$$

$$\beta = \frac{\sum_{i=1}^n \frac{w_i^2 (y_i - \bar{y})^2}{w(x_i)} - \sum_{i=1}^n w_i (x_i - \bar{x})^2}{3 \sum_{i=1}^n \frac{w_i^2 (x_i - \bar{x})^2}{w(x_i)}} \quad (46)$$

$$\gamma = - \frac{\sum_{i=1}^n w_i (x_i - \bar{x}) (y_i - \bar{y})}{\sum_{i=1}^n \frac{w_i^2 (x_i - \bar{x})^2}{w(x_i)}} \quad (47)$$

将(44)式看作一元三次方程, 得到其三个实根为:

$$b_{j+1} = \alpha + 2 \sqrt{\alpha^2 - \beta} \cdot \cos \frac{1}{3}(\varphi + 2\pi j),$$

$$j = 0, 1, 2 \text{ (York, 1966)}$$

$$\text{这里 } \cos \varphi = \frac{\alpha^2 - \frac{3}{2}\alpha\beta + \frac{1}{2}\gamma}{\sqrt{(\alpha^2 - \beta)^3}}$$

York (1966) 指出 b_3 是我们所需的根, 因此:

$$b = \alpha + 2 \sqrt{\alpha^2 - \beta} \cdot \cos \frac{1}{3}(\varphi + 6\pi) \quad (48)$$

由于 α 、 β 、 γ 都含有 w_i , 而 w_i 中也含有 b , 因此并不能直接从(48)式中解出所需的斜率 b 。这时就需要用到方程求根的迭代法 (参见李庆扬, 2001), 用计算机编程计算 (参见彭国伦, 2002)。简要说明其原理:

① 首先给出一个 b 的近似值 $b_{(1)}$, 代入(6) 计算 w_i 并将所得代入(45)、(46)、(47) 计算 α 、 β 、 γ , 进一步将所得代入(48) 得到 $b_{(2)}$;

② 将 $b_{(2)}$ 代入(6) 计算 w_i 并将所得代入(45)、(46)、(47) 计算 α 、 β 、 γ , 进一步将所得代入(48) 得到 $b_{(3)}$;

③ 依次重复以上过程, 直到 $b_{(k+1)} - b_{(k)} < \varepsilon$ 。这里 ε 是一个我们可以接受的、使得 $b_{(k+1)}$ 与 b_k 接近几乎完全相等的数值, 取决于我们所要求的斜率 b 的精度。 $b_{(k+1)}$ 即为所求斜率 b 。

这时, a 同样由(12)式 (也就是(41)式) 解出。

将(30)、(31)代入(7)式, 得到一个极有意义的表达式 (York, 1967),

$$L_i' \text{ 的斜率} = - \frac{1}{b} \cdot \frac{w(x_i)}{w(y_i)} \quad (49)$$

据此可以确定调整线的方向, 可以看出, 调整线的斜

率与最佳拟合直线的斜率 b 符号相反,亦即调整线位于图 2 所示的三角阴影区内。此等式对于 2.1 节至 2.5 节都成立。

2.5.2 独立加权模型解法二

其它学者 (McIntyre et al, 1966; Wendt, 1969, 转引自 Brooks, 1972; Borchers and Sheth, 1995 等) 则从公式 (5) 着手, 根据 S 取最小值时的条件, 得出斜率和截距的计算办法。下面我们尝试推导,

$$\begin{aligned} S &= \sum_{i=1}^n w_i (y_i - a - bx_i)^2 \\ &= \sum_{i=1}^n \frac{w(x_i)w(y_i)}{w(x_i) + b^2 w(y_i)} (y_i - a - bx_i)^2 \end{aligned}$$

令 $\partial S / \partial a = 0$, 得:

$$\begin{aligned} \sum_{i=1}^n w_i (y_i - a - bx_i) &= 0 \Rightarrow \\ a &= \frac{\sum_{i=1}^n w_i y_i - b \sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i} = \bar{y} - b\bar{x} \end{aligned} \quad (50)$$

令 $\partial S / \partial b = 0$, 得:

$$\begin{aligned} \sum_{i=1}^n w_i x_i (y_i - a - bx_i) + \\ b \sum_{i=1}^n w(y_i) w^2_i (y_i - a - bx_i)^2 = 0 \end{aligned} \quad (51)$$

将 (50) 代入 (51), 化简得到:

$$\begin{aligned} \sum_{i=1}^n [w(y_i) w^2_i (x_i - \bar{x})^2] b^3 - \\ \sum_{i=1}^n 2[w(y_i) w^2_i (x_i - \bar{x}) (y_i - \bar{y})] b^2 + \\ \sum_{i=1}^n [w(y_i) w^2_i (y_i - \bar{y})^2 - w_i x_i (x_i - \bar{x})] b + \\ \sum_{i=1}^n w_i x_i (y_i - \bar{y}) = 0 \end{aligned} \quad (52)$$

同样可将 (52) 式看作关于 b 的三次方程, 用上述类似的迭代法处理即可。

Borchers 和 Sheth (1995) 并未先将 (50) 式代入 (51), 而是将解三次方程的复杂性转移到迭代过程中去。根据 (51) 化简得到一个二次方程:

$$Ab^2 + Bb + C = 0 \quad (53)$$

其中

$$\begin{aligned} A &= a \sum w^2(y_i)w(x_i)x_i - \sum w^2(y_i)w(x_i)x_i y_i \\ B &= a^2 \sum w^2(y_i)w(x_i) - 2a \sum w^2(y_i)w(x_i)y_i \\ &\quad + \sum w^2(y_i)w(x_i)y_i^2 - \sum w^2(x_i)w(y_i)x_i^2 \end{aligned}$$

$C = \sum w^2(x_i)w(y_i)x_i y_i - a \sum w^2(x_i)w(y_i)x_i$
这时开始用迭代法:

① 用普通最小二乘法得到 b 的初始值 b_0 , 代入 (50) 计算 a , 用这个 a 代入 (51) 式, 计算得到 b 的两个根。② 将 b 的两个根及 ① 中 a 代入 (5) 分别求 S 。选择较小的 S 对应的 b 作为新的 b 值, 记为 b_1 。③ b_1 再代入 (51) 式, 计算得到 b 的两个根, 与 ② 类似处理, 得到 b_2 。依次重复, 得到 $b_{(j)}$, 直到 $b_{(k+1)} - b_{(k)} < \varepsilon_0 \varepsilon$ 与 2.5.1 节中相同。这时的 $b_{(k+1)}$ 与对应的 a 即为所求。

2.5.3 独立加权模型与其它最小二乘模型的关系

可以认为 2.1 ~ 2.4 节所述的四种模型是 2.5 节中的独立加权最小二乘模型的特殊情形, 虽然如此, 但由于前四种情形下, 能够得到斜率 b 的解析解, 而且实践中会有与之类似的问题, 因此详细讨论是有必要的。而独立加权最小二乘模型则无解析解, 需用迭代法进行数值计算。

2.1 ~ 2.4 节以及 2.5.2 节都是用偏导数为 0 作为取得极值的条件进行推导的, 因此其解是相通的; 而 2.5.1 节则采用变分法得到, 给定 2.5.1 节中 (43) 式不同的条件, 进一步演算, 则可得到公式 (11)、(15)、(20)、(22)。

2.6 误差相关最小二乘法 (error correlated independent weighting model)

York (1969) 将独立加权模型进一步复杂化, 提出了 (1) (2) 两种计算公式, 并引进误差相关 (系数) ρ_i 误差相关这个参数, 即 $\sigma(x_i)$ 和 $\sigma(y_i)$ 之间的线性相关系数, $|\rho_i| \leq 1$, 当其取 0 时, $\sigma(x_i)$ 和 $\sigma(y_i)$ 相互独立; 当其绝对值为 1 时, 说明其完全线性相关, +1 意味着正相关, -1 则负相关 (Ludwig, 2008; 梁晋文, 2001; 费业泰, 2010; 褚宝增和王翠香, 2010)。通常所见的 U-Pb 谐和图中, 误差椭圆的方位也就是误差相关系数的反映。对于此种最小二乘法, 不仅考虑每个点的误差, 而且考虑每个点纵横坐标误差之间的相关系数, 这时所有测试点的调整线一般不平行, 且不会限制在上述 2.5 节中所述的阴影三角形内 (图 2f)。

与 2.5 节类似, 也可以从 (1) 式着手, 令 S 的一阶变分为 0, 或从 (2) 式着手, 分别对 a 、 b 求偏导数并令其结果为 0, 推演 b 的计算过程。这里不再做详细推导, 其具体原理同上。York (1969) 从 (2) 式着手, 得到的最小二乘平方方程为:

$$b^2 \sum_{i=1}^n w_i^2 \left[\frac{U_i V_i}{w(x_i)} - \frac{\rho_i U_i^2}{\alpha_i} \right] + b \sum_{i=1}^n w_i^2 \left[\frac{U_i^2}{w(y_i)} - \frac{V_i^2}{w(x_i)} \right] - \sum_{i=1}^n w_i^2 \left[\frac{U_i V_i}{w(y_i)} - \frac{\rho_i V_i^2}{\alpha_i} \right] = 0 \quad (54)$$

据此的最佳拟合直线斜率为:

$$b = - \frac{\sum_{i=1}^n w_i^2 \left[\frac{U_i^2}{w(y_i)} - \frac{V_i^2}{w(x_i)} \right] + \sqrt{\sum_{i=1}^n w_i^2 \left[\frac{U_i^2}{w(y_i)} - \frac{V_i^2}{w(x_i)} \right] + 4 \sum_{i=1}^n w_i^2 \left[\frac{U_i V_i}{w(x_i)} - \frac{\rho_i V_i^2}{\alpha_i} \right] \times \sum_{i=1}^n w_i^2 \left[\frac{U_i V_i}{w(y_i)} - \frac{\rho_i V_i^2}{\alpha_i} \right]}}{2 \sum_{i=1}^n w_i^2 \left[\frac{U_i V_i}{w(x_i)} - \frac{\rho_i V_i^2}{\alpha_i} \right]} \quad (55)$$

或

$$b = \frac{\sum_{i=1}^n w_i^2 V_i \left[\frac{U_i}{w(y_i)} + \frac{b V_i}{w(x_i)} - \frac{\rho_i V_i}{\alpha_i} \right]}{\sum_{i=1}^n w_i^2 U_i \left[\frac{U_i}{w(y_i)} + \frac{b V_i}{w(x_i)} - \frac{b \rho_i U_i}{\alpha_i} \right]} \quad (56)$$

其中:

$$U_i = x_i - \bar{x}, V_i = y_i - \bar{y}, \bar{y} = \frac{\sum_{i=1}^n (w_i y_i)}{\sum_{i=1}^n w_i}, \bar{x} = \frac{\sum_{i=1}^n (w_i x_i)}{\sum_{i=1}^n w_i}, \alpha_i = \sqrt{w(x_i) w(y_i)} \quad (57)$$

w_i 由(3)式给出。

x, y 的残差分别为:

$$x_i - X_i = \frac{w_i(a + b x_i - y_i) [\rho_i \alpha_i - b w(y_i)]}{w(x_i) w(y_i)} \quad (58)$$

$$y_i - Y_i = \frac{w_i(a + b x_i - y_i) [w(x_i) - b \rho_i \alpha_i]}{w(x_i) w(y_i)} \quad (59)$$

$$L_i' \text{ 的斜率} = \frac{w(x_i) - b \rho_i \alpha_i}{\rho_i \alpha_i - b w(y_i)} \quad (60)$$

由于 $w_i, \bar{x}, \bar{y}, U_i, V_i$ 中都含有 b , 因此(55) 和(56) 的解答也需要用到迭代法。 a 类似, 可用(12)

式计算得到。

3 相关参数的误差估计

关于最佳拟合直线的斜率 b 和截距 a 的误差(标准差)计算, 较多采用的是误差传递公式(McIntyre et al., 1966; York, 1966, 1969; Brooks et al., 1972; Kullerud, 1991; Mahon, 1996), 而 York (1969) 所给的误差公式则存在错误(Titterton and Halliday, 1979; Mahon, 1996)。Titterton and Halliday (1979) 用最大似然法所讨论的结果与误差传递公式的计算结果相同。Sheth 等(1996)则提出了另一种标准差的计算方法。无论哪种算法, 都是对标准差的近似估计, 这里主要介绍前者。而另一个至关重要的问题是在得到标准差后, 如何确定置信区间。

3.1 标准差的计算

这里以误差相关的最小二乘模型为例介绍, 其它模型可以通过相关参数取特殊值得到, 也可用类似的方法进行推导。记(56)式为 $F(b, x_i, y_i) = 0$, 即:

$$F = b^2 \sum_{i=1}^n w_i^2 \left[\frac{U_i V_i}{w(x_i)} - \frac{\rho_i U_i^2}{\alpha_i} \right] + b \sum_{i=1}^n w_i^2 \left[\frac{U_i^2}{w(y_i)} - \frac{V_i^2}{w(x_i)} \right] - \sum_{i=1}^n w_i^2 \left[\frac{U_i V_i}{w(y_i)} - \frac{\rho_i V_i^2}{\alpha_i} \right] \quad (61)$$

根据误差传递公式(省略泰勒序列的高次项), 有(Titterton and Halliday, 1979; Mahon, 1996; 费业泰, 2010):

$$\sigma_b^2 \cong \sum_{i=1}^n \left[\left(\frac{\partial b}{\partial x_i} \right)^2 \cdot \sigma^2(x_i) + \left(\frac{\partial b}{\partial y_i} \right)^2 \cdot \sigma^2(y_i) + 2 \rho_i \sigma(x_i) \sigma(y_i) \left(\frac{\partial b}{\partial x_i} \right) \cdot \left(\frac{\partial b}{\partial y_i} \right) \right] \quad (62)$$

根据隐函数求导公式(同济大学应用数学系, 2002),

$$\frac{\partial b}{\partial x_i} = - \frac{\frac{\partial F}{\partial x_i}}{\frac{\partial F}{\partial b}}, \frac{\partial b}{\partial y_i} = - \frac{\frac{\partial F}{\partial y_i}}{\frac{\partial F}{\partial b}} \quad (63)$$

以上两式联立,就可得到 Mahon(1996)所用的公式:

$$\sigma_b^2 \cong \frac{\sum_{i=1}^n \left[\left(\frac{\partial F}{\partial x_i} \right)^2 \cdot \sigma^2(x_i) + \left(\frac{\partial F}{\partial y_i} \right)^2 \cdot \sigma^2(y_i) + 2\rho_i \sigma(x_i) \sigma(y_i) \left(\frac{\partial F}{\partial x_i} \right) \left(\frac{\partial F}{\partial y_i} \right) \right]}{\left(\frac{\partial F}{\partial b} \right)^2} \quad (64)$$

或

$$\sigma_b^2 \cong \frac{\sum_{i=1}^n \left[\left(\frac{\partial F}{\partial x_i} \right)^2 \cdot \frac{1}{w(x_i)} + \left(\frac{\partial F}{\partial y_i} \right)^2 \cdot \frac{1}{w(y_i)} + \frac{2\rho_i}{\sqrt{w(x_i)w(y_i)}} \cdot \left(\frac{\partial F}{\partial x_i} \right) \cdot \left(\frac{\partial F}{\partial y_i} \right) \right]}{\left(\frac{\partial F}{\partial b} \right)^2} \quad (65)$$

对于截距 a 的误差,也可采用类似的方法,将(12)式看作关于 x_i, y_i 的函数,有:

$$\sigma_a^2 \cong \sum_{i=1}^n \left[\left(\frac{\partial a}{\partial x_i} \right)^2 \sigma^2(x_i) + \left(\frac{\partial a}{\partial y_i} \right)^2 \sigma^2(y_i) + 2\rho_i \sigma(x_i) \sigma(y_i) \left(\frac{\partial a}{\partial x_i} \right) \left(\frac{\partial a}{\partial y_i} \right) \right] \quad (66)$$

将 $\partial F/\partial x_i, \partial F/\partial y_i, \partial F/\partial b, \partial a/\partial x_i, \partial a/\partial y_i$ 代入(64)(66)式,得到以下两式(Mahon, 1996; Titterton and Halliday, 1979):

$$\sigma_b = \frac{1}{\sum_{i=1}^n [w_i(x_i - \bar{x})]} \quad (67)$$

$$\sigma_a = \frac{\sigma_b^2 \sum_{i=1}^n (w_i x_i^2)}{\sum_{i=1}^n w_i} \quad (68)$$

York(1969)中认为可以将点关于最佳拟合直线的分散程度考虑在内,进行误差估计,得到扩展的不确定度(York, 1966, 1969; Kullerud, 1991),即:

$$\sigma_b \text{ exp} = \sigma_b \sqrt{MSWD},$$
$$\sigma_a \text{ exp} = \sigma_a \sqrt{MSWD} \quad (69)$$

有关 $MSWD$ 的介绍见下。Mahon(1996)认为这种做法是从数学角度来说是不严格的,如果 $MSWD$ 小于 1 的话,就会降低误差。因此在实践中如若用到,一定要指明。

3.2 确定置信区间

通常给标准差(σ_b 和 σ_a ;代表 68% 的置信区间)乘以 2(更准确地说是 1.96,表 1),得到斜率 b 和截距 a 95% 的置信区间。事实上,这样做的前提是用于进行等时线拟合的样品数据点足够多(即 n 较大),且斜率和截距的计算结果服从正态分布。这时,平均值附近的 $2\sigma(1.96\sigma)$ 区间,代表了概率密度曲线中 95% 的面积区域。当拟合数据点较少时,就需要乘以 t -分布的 α (这里取 0.025,对应于 95%

的置信区间)分位点 $t_{n-2}(\alpha = 0.025)$ (参见褚宝增和王翠香等,2010)(Mahon, 1996; Ludwig, 2008)。这时:

$$b = b^* \pm t_{n-2}(\alpha = 0.025) \cdot \sigma_b, a = a^* \pm t_{n-2}(\alpha = 0.025) \cdot \sigma_a$$

(b^* 表示最佳斜率, a^* 同)。

当数据点较少时,斜率的误差会非常大,正因为如此,Ludwig(1998, 2000, 2003, 2008 等)一直强调避免 3 点或 4 点等时线。

4 MSWD

根据 Wendt 和 Carl(1991), $MSWD$ (mean squared weighted deviation)计算方法为:

$$MSWD = \frac{1}{f} \sum (y_i - bx_i - a) \cdot w_i \quad (70)$$

其中 $f = n - k$,为自由度,对于加权且固定截距的线性回归, k 取 1,而对于无限制条件的线性回归, k 取 2(Mahon, 1996), n 表示拟合所用的数据点数目。 w_i 如(3)式或(6)式所示。据此可知,这里:

$$MSWD = \frac{S^*}{f} \quad (71)$$

即最佳拟合直线所对应的偏差平方和(S^*)的平均值。根据以上 3 中的讨论可知, S^* 的大小反映了所有数据点偏离最佳等时线的总距离,当所有数据点严格位于最佳等时线时, S^* 取值为 0,当然 $MSWD$ 也为 0,当然这种情况是绝对理想化的。

Wendt 和 Carl(1991)系统论证了 $MSWD$ 的统计规律,指出当所有数据点对最佳拟合直线的偏离如果仅由分析过程中的随机误差所引起,那么 $MSWD$

的平均值恒为 1, $1\sigma = \pm \sqrt{\frac{2}{f}}$, 95% 的置信区间的上界约为 $1 + 2\sqrt{\frac{2}{f}}$, 其概率密度曲线却是非对称的。表 1 给出了此种情况下, $MSWD$ 的 95% 置信区间的临界值, 其中用 H_{MSWD} 表示 $MSWD$ 的 95% 的置

表 1 不同自由度所对应的 $MSWD$ 临界值
Table 1 The critical $MSWD$ values versus degree of freedom

自由度(f)	平均值	中值	L_{MSWD} (95%)	H_{MSWD} (95%)
1	1.0	0.455	0.001	5.020
2	1.0	0.694	0.025	3.690
3	1.0	0.790	0.072	3.117
4	1.0	0.840	0.121	2.775
5	1.0	0.870	0.166	2.560
6	1.0	0.892	0.207	2.400
7	1.0	0.907	0.241	2.286
8	1.0	0.918	0.273	2.188
9	1.0	0.927	0.300	2.111
10	1.0	0.934	0.325	2.050
11	1.0	0.936	0.347	1.991
12	1.0	0.942	0.367	1.942
13	1.0	0.946	0.385	1.900
14	1.0	0.950	0.402	1.864
15	1.0	0.953	0.417	1.833
16	1.0	0.956	0.432	1.800
17	1.0	0.959	0.445	1.776
18	1.0	0.961	0.457	1.750
19	1.0	0.963	0.469	1.732
20	1.0	0.965	0.480	1.710
21	1.0	0.967	0.490	1.690
22	1.0	0.968	0.500	1.673
23	1.0	0.970	0.509	1.657
24	1.0	0.971	0.517	1.642
25	1.0	0.972	0.524	1.624
30	1.0	0.977	0.560	1.567
40	1.0	0.983	0.610	1.483
50	1.0	0.986	0.648	1.428
60	1.0	0.988	0.675	1.388
70	1.0	0.990	0.697	1.357
80	1.0	0.991	0.715	1.338
90	1.0	0.992	0.729	1.311
100	1.0	0.993	0.742	1.300

注:此表给出了当数据点对最佳拟合直线的偏离仅由测试随机误差所引起时, $MSWD$ 的临界值, 据 Mahon (1996)。另据 Wendt 和 Carl (1991), H_{MSWD} (95%) 约为 $1 + 2\sqrt{\frac{2}{f}}$ 。

Note: This table shows the critical $MSWD$ values versus degree of freedom when the data points are away from the best-fit line only due to the analytical uncertainty, after Mahon (1996)。Wendt and Carl (1991) also point out that the H_{MSWD} (95%) is about $1 + 2\sqrt{\frac{2}{f}}$

信区间的上界, 类似 L_{MSWD} 表示下界。这意味着 $MSWD$ 低于 L_{MSWD} 的情况出现的概率只有 2.5% (除非分析误差估计过高), 而 $MSWD$ 值高于 H_{MSWD} 的情况也只有 2.5% 的概率 (除非有分析误差之外的误差存在, 或者数据点不符合线性条件)。同时, 从表 1 可以看出, H_{MSWD} 并非一个固定的值, 而且 $MSWD$ 的中值与平均值 1 也不一致, 当自由度增大时, 中值逐渐靠近平均值 1。

接下来一个重要的问题是在同位素地质年代学中, 如何应用 $MSWD$ 对拟合结果进行解释。通常可能会这样认为, 当测试所得的 $MSWD$ 小于对应自由度的 H_{MSWD} 时, 说明拟合效果较好, 具有地质意义; 反之高于时, 则需要重新考虑是否样品完全符合测年条件。是否如陈骏和王鹤年 (2004) 所述“当 $MSWD > 2.5$ 时, 很可能是一条误差等时线”, 或如韩吟文和马振东 (2003) 所述, “该值越小, 等时线质量越好。当存在地球化学误差时, $MSWD > 1$; 当不存在地球化学误差时, $MSWD \leq 1$ ”。问题似乎并没有这么简单, Kalsbeek (1992)、Wendt (1992) 对此已经做了一些讨论。

首先, 需要明确的是, $MSWD$ 值明显受对实验分析误差估计准确性的影响 (陈道公, 2009; 这点可从公式 (70) 与 (1) 至 (6) 明显地看出), 只有在使用最小二乘法 (或最大似然法) 拟合等时线, 且对数据点的误差估计准确时, $MSWD$ 才可作为检验拟合结果的参数 (Mahon, 1996)。但实际研究工作中, 这点并没有得到重视, 笔者查阅了一些近年来用 Rb-Sr 等时线法测年的文章发现, 许多文章并未交代等时线拟合中相关变量的 (分析) 误差估计, 却使用 $MSWD$ 对拟合结果进行评价, 如李真真等 (2009)、李光来等 (2011)、张臣 (1998)、蓝廷广等 (2011)、祝禧艳等 (2010)、周雄等 (2010)、陈江峰等 (2003)、魏俊浩等 (2003)、毛光周等 (2008)。而给出了误差的文献中, 对分析误差估计的方式依然存在着不明确性。李秋立等 (2006) 明确指出拟合时, $n(^{87}\text{Rb})/n(^{86}\text{Sr})$ 比值采用 2% 误差、Sr 同位素比值采用保守外精度误差 0.05%; 姚海涛等 (2001) 选用 $n(^{87}\text{Rb})/n(^{86}\text{Sr})$ 、 $n(^{87}\text{Sr})/n(^{86}\text{Sr})$ 的分析误差分别为 1.5%、0.02%, 并提及这是“一般的全岩 Rb-Sr 等时线法定年的分析误差”; 杨进辉和周新华 (2000) 则只提及 $n(^{87}\text{Rb})/n(^{86}\text{Sr})$ 的输入误差为 0.5% ~ 2%。

其次, 根据上述 3、4 部分的详细介绍可知, 这里仅考虑了实验分析误差, 且给出了仅由分析过程中的随机误差引起数据点对最佳拟合直线偏离时,

MSWD 的概率分布情况。但事实上,数据点对最佳拟合直线的偏离(甚至靠近),不仅可能来自实验分析误差,而且还有可能来自地质背景、每个样品本身、以及测年方法的选择,但对于除实验分析误差之外的其它误差,目前似乎并没有合理的方法将其定量化(Allegre, 2008),因此在拟合时,也仅考虑了实验测试误差。实际研究中所用样品,更多地是接近而非完全满足同位素测年所需的条件,这就必然降低 *MSWD* 对拟合结果的评判意义。以等时线法为例,测年样品不一定真正具有完全相同的初始同位素比值(见 Zheng, 1989),同位素体系能够在所研究对象内不一定能够完全封闭。这样的话,一方面,95%的置信边界只是表明当数据点的分散仅是由测试过程中的随机误差所引起时,*MSWD* 会有95%的可能性落在这个区间内,但并不排除存在其它因素导致 *MSWD* 落入这个区域的可能性,并且这种可能性有多大是未知的;另一方面,对于 *MSWD* 落入其95%置信区间外的情况,也有可能是可以接受的。

基于此,在应用同位素地质年代学方法对地质体进行年代测定时,需尽量选择符合所用方法基本假设的样品,并且尽可能地准确估计分析测试误差。在此基础上,当 *MSWD* 落入95%的置信区间内时,说明测试结果线性相关程度良好,在尽可能地降低其他因素影响的情况下,所获年龄具有较大的参考意义。而当 *MSWD* 落入95%的置信区间外时,且测试误差估计准确、*MSWD* 不至于过大时,拟合结果仍有一定参考意义。当然 *MSWD* 十分大时,则需要慎重对待测年结果。除此之外,更要结合地质背景,对所得结果进行检验。

5 讨论

在同位素地质年代学研究中,最小二乘法直线拟合是一种常用的数学处理方法。虽然普通最小二乘法广为人知,但由于拟合所用数据点均为测试所得同位素比值,纵横坐标变量均存在误差甚至误差相关,而普通最小二乘法假设一个变量不存在误差的前提条件已经不能成立。因此,需要建立更加符合数据本身特点的拟合方法。

基于此,我们在前人研究的基础上,对最小二乘法进行了系统性地总结与研究。详细介绍了普通最小二乘法、加权普通最小二乘法、标准加权最小二乘法、标准独立加权最小二乘法、独立加权最小二乘法以及误差相关最小二乘法,给出了各自的“最小二乘”原理,以及拟合直线斜率、截距及其误差的计算

方法。在将相关数学计算方法进一步完善并系统化的同时,也阐明了各种最小二乘模型之间的相互关系。误差相关最小二乘法可以看作是最一般的情形,而独立加权、标准独立加权、标准加权、普通最小二乘法则可看作是其一步步“特殊化”的情形。误差相关、独立加权这两种最小二乘法没有数值解,需要借助计算机编程后用迭代法得到其近似解,文中给出了迭代的具体算法;而前四种最小二乘法则具有数值解。

可以看出,不同的拟合模型对应于不同的假设条件,这允许根据数据本身的特点,选择合适的模型进行拟合。如对于 U-Pb 不一致曲线,只能选择误差相关最小二乘法,而对于 Rb-Sr、Sm-Nd、K-Ar、Re-Os 等时线法等,则可选用最一般形式的独立加权最小二乘法。当然,如果数据有符合标准独立加权、标准加权、(普通)最小二乘法的假设条件的,即可采用这些简化的模型,从而简化计算。这里,笔者发现对于 Rb-Sr、Sm-Nd 等时线法,由于 $n(^{87}\text{Rb})/n(^{86}\text{Sr})$ $[n(^{147}\text{Sm})/n(^{144}\text{Nd})]$ 的实验误差比 $n(^{87}\text{Sr})/n(^{86}\text{Sr})$ $[n(^{143}\text{Nd})/n(^{144}\text{Nd})]$ 要大许多,因此采用普通最小二乘法或加权普通最小二乘法,将数据点偏离拟合直线的原因全部归于横坐标 x $[n(^{87}\text{Rb})/n(^{86}\text{Sr})$ 、 $n(^{147}\text{Sm})/n(^{144}\text{Nd})]$,或许是可以接受的。

我们进一步澄清了参数 *MSWD* 对于同位素地质年代测定结果的参考意义。理论上,在拟合直线时,应该考虑包括每个样品的采样、地质背景、方法应用及实验分析所带来的所有误差,但实际上只有实验分析误差是可定量估计的。因此,拟合过程中我们只考虑了实验分析误差。鉴于此,在所选样品尽量符合所用测年方法原理的基础上,如果实验分析误差估计准确,那么当 *MSWD* 落入其95%的概率区间内或不至于太大时,拟合结果具有较大的参考意义,但也不排除错误的测年结果存在的可能性。当 *MSWD* 过于大时,则需慎重。当然,由于地质“未知”世界的复杂性,在评价拟合结果时,也需要结合已有的地质知识进行判别。实际研究中,对于每一个数据点,更需将其看作“一个个同位素比值”,而非“一个个年龄”。

同位素年代学研究方法是建立在一定的假设基础之上的,根据地质样品本身的规律,对这些假设予以统计学修正,应该是今后发展的一个方向。这也有利于以地质学内容为素材,开拓新的方法。我们的研究对于其他自然科学及社会科学领域内,相似

问题的解决也具有一定的参考意义。

6 结 论

本文对最小二乘直线拟合这一同位素地质年代学中常用的统计方法进行了总结,详解了其数学原理与实现过程,实现了从普通最小二乘模型到误差相关最小二乘模型的统一。不同类型的最小二乘模型对应于不同的假设条件,研究者可根据其数据特点进行选择。对于 U-Pb 不一致曲线,只能选择误差相关最小二乘法,而对于 Rb-Sr、Sm-Nd、K-Ar、Re-Os 及其它等时线,除可选用最一般形式的独立加权最小二乘法,也可据数据特点,选择标准独立加权、标准加权、普通最小二乘法等简化模型计算。在此基础上,我们进一步澄清了 *MSWD* 这一参数对拟合结果的评判意义。准确理解这些,对于我们更好地应用相关测年方法,评价测年结果的可靠性至关重要。我们的研究也可用于其他领域内相似问题的研究中。

致谢:本文完成过程中得到了北京离子探针中心宋彪研究员,中国科学院地质与地球物理研究所姜能研究员、储著银副研究员,中国地质大学(北京)诸慧燕老师、苏文博教授、高世臣教授、王翠香副教授,中国军事医学科学院张春茂、上海财经大学马馨瑞及中国地质大学(北京)甘彩虹、薛玉山、杨宽、张志强等的帮助,匿名评审人及章雨旭编辑的认真工作使本文进一步完善,在此一并表示衷心感谢。

参 考 文 献 / References

陈江峰,谢智,郑永飞,龚冰,陈文,周泰禧,张巽.2003.皖东侵入岩矿物 Rb-Sr 和 Ar-Ar 年代学及其与氧同位素平衡之间的关系. *高校地质学报*,9(2):172~184.

陈道公.2009. *地球化学*. 合肥:中国科学技术大学出版社,230~231.

陈骏,王鹤年.2004. *地球化学*. 北京:科学出版社,141~174.

陈岳龙,杨忠芳,赵志丹.2005. *同位素地质年代学与地球化学*. 北京:地质出版社,22~24.

褚宝增,王翠香.2010. *概率统计*. 北京:北京大学出版社,97~199.

费业泰.2010. *误差理论与数据处理*. 北京:机械工业出版社,57~164.

G. 福尔著,潘曙兰,乔广生译,李继亮校.1983. *同位素地质学原理*. 北京:科学出版社,59~72.

韩吟文,马振东.2003. *地球化学*. 北京:地质出版社,222~223.

何晓群和刘文卿.2011. *应用回归分析*. 北京:中国人民大学出版社,1~126.

蓝廷广,范宏瑞,胡芳芳,杨奎锋,王永.2011. 山东微山稀土矿床成因:来自云母 Rb-Sr 年龄激光 Nd 同位素及流体包裹体的证据. *地球化学*,40(5):428~442.

老大中.2004. *变分法基础*. 北京:国防工业出版社,37~98.

梁晋文. *误差理论与数据处理*.2001. 北京:中国计量出版社,67~91.

李光来,华仁民,韦星林,王旭东,黄小娥.2011. 江西中部徐山钨铜矿床单颗粒白云母 Rb-Sr 等时线定年及其地质意义. *地球科学*,36(2):282~288.

李庆扬. *数值分析基础教程*.2001. 北京:高等教育出版社,9~83.

李秋立,陈福坤,王秀丽,李向辉,李潮峰.2006. 超低本底化学流程和单颗粒云母 Rb-Sr 等时线定年. *科学通报*,51(3):321~325.

李志昌,路远发,黄圭成.2004. *放射性同位素地质学方法与进展*. 武汉:中国地质大学出版社,11~62.

李真真,李胜荣,张华锋.2009. 黑龙江东宁县金厂金矿围岩蚀变和成矿年代学特征. *矿床地质*,28(1):83~92.

毛光周,华仁民,高剑峰,龙光明,陆慧娟.2008. 江西金山金矿含金黄铁矿的 Rb-Sr 年龄. *地球学报*,29(5):599~606.

彭国伦.2002. *Fortran 95 程序设计*. 北京:中国电力出版社,435~472.

邵克勇,王婷婷,宋金波.2011. *最优控制理论与应用*. 北京:化学工业出版社,11~42.

盛聚,谢式千,潘承毅.2008. *概率论与数理统计*. 北京:高等教育出版社,291~293

申浩澈.1983. 同位素等时年龄中若干问题的探讨. *吉林地质*,1:69~76.

同济大学应用数学系.2002. *高等数学*. 北京:高等教育出版社,62~137.

王黎明,陈颖,杨楠.2008. *应用回归分析*. 上海:复旦大学出版社,12~37.

魏俊浩,刘丛强,李志德,赵永鑫.2003. 论金矿床成矿年代的确定——以丹东地区成岩成矿 Rb-Sr、U-Pb 同位素年代为例. *地质学报*,77(1):113~119.

吴翊,李永乐,胡庆军.1995. *应用数理统计*. 长沙:国防科技大学出版社,143~157.

杨进辉,周新华.2000. 胶东地区玲珑金矿矿石和载金矿物 Rb-Sr 等时线年龄与成矿时代. *科学通报*,45(14):1547~1552.

姚海涛,郑海飞. 流体包裹体 Rb-Sr 等时线定年的可靠性. *地球化学*,2001,30(6):507~511.

张臣,吴泰然.1998. 内蒙古温都尔庙群变质基性火山岩 Sm-N d、Rb-Sr 同位素年代研究. *地质科学*,33(1):25~30.

张启锐.1988. *实用回归分析*. 北京:地质出版社,14~43.

张小蒂.1991. *应用回归分析*. 杭州:浙江大学出版社,22~110.

中国科学院数学研究所数理统计组.1974. *回归分析方法*. 北京:科学出版社,5~18.

周复恭,黄运成.1989. *应用线性回归分析*. 北京:中国人民大学出版社,1~169.

周纪芄.1993. *回归分析*. 上海:华东师范大学出版社,1~26.

周惟公.2008. *大学物理实验*. 北京:高等教育出版社,20~22.

周雄,温春齐,张学全,曹盛远,吴鹏宇,周 玉,费光春.2010. 西藏邦铺大型钼铜多金属矿床 Rb-Sr 等时线年龄及地质意义. *高校地质学报*,16(4):448~456.

祝禧艳,陈福坤,杨一增,胡蓉.2000. 微区-微量样品 Rb-Sr 同位素分析技术及其应用前景. *科学通报*,26(1):325~332.

Allegre C J. 2008. *Isotope geology*. New York: Cambridge Univeristy press, 58~215.

Borchers P H and Sheth C V. 1995. Least squares fitting of a straight line to a set of data points. *European Journal of Physics*, 16(5):204~210.

BrooksC, Hart S R, Wendt I. 1972. Realistic Use of Two-Error Regression Treatments as Applied to Rubidium-Strontium Data. *Review of Geophysics*, 10(2): 551~577.

Bruzzone H,Moreno C. 1998. When errors in both coordinates make a difference in the fitting of a straight lines by least squares.

- Measurement Science and Technology, 19(12): 2007 ~ 2011.
- Deming W E. 1943. Statistical adjustment of data. New York: John Wiley & Sons, 1 ~ 261.
- Faure G. 1986. Principles of isotope geology. New York: John Wiley & Sons, 1 ~ 271.
- Kalsbeek F. 1992. The statistical distribution of the mean squared weighted deviation-Comment: Isochrons, errorochrons, and the use of MSWD-values. Chemical Geology, 94(3): 241 ~ 242.
- Kullerud L. 1991. On the calculation of isochrons. Chemical Geology, 87(2): 115 ~ 124.
- Ludwig K R. 1998. Using Isoplot/Ex Version 1. 00b. A Geochronological Toolkit for Microsoft Excel. Berkeley Geochronology Center Special Publication No. 1, 1 ~ 38.
- Ludwig K R. 2000. User's Manual for Isoplot/Ex Version 2. 2. A Geochronological Toolkit for Microsoft Excel. Berkeley Geochronology Center Special Publication No. 1a, 1 ~ 50.
- Ludwig K R. 2003. User's Manual for Isoplot 3.00. A Geochronological Toolkit for Microsoft Excel. Berkeley Geochronology Center Special Publication No. 4, 1 ~ 67.
- Ludwig K R. 2008. User's Manual for Isoplot 3.70. A Geochronological Toolkit for Microsoft Excel. Berkeley Geochronology Center Special Publication No. 4, 1 ~ 72.
- Mahon K. 1996. The New "York" Regression: Application of an Improved Statistical Method to Geochemistry. International Geology Review, 38(4): 293 ~ 303.
- Mcintyre G A, Brooks C, Compston W, Turek A. 1966. The Statistical Assessment of Rb-Sr Isochrons. Journal of Geophysical Research, 71(22): 5459 ~ 5468.
- Rosholt J N, Zartman R E, and Nkomo I T. 1973. Lead Isotope Systematics and Uranium Depletion in the Granite Mountains, Wyoming. Geological Society of America Bulletin, 84(3): 989 ~ 1002.
- Sheth C V, Ngwengwez A, Borchers P H. 1996. Least squares fitting of a straight line to a set of data points: II. Parameter Variances. European Journal of Physics, 17(6): 322 ~ 326.
- Titterton D M, Halliday A N. 1979. On the fitting of parallel isochrones and the method of maximum likelihood. Chemical Geology, 26(3 ~ 4): 183 ~ 195.
- Wendt I. 1969. Derivation of the Formula for a Regression Line Based on Least-square Analysis. Hanover, Western Germany: Bundesanstalt für Bodenforschung Internal Rep.
- Wendt I. 1992. The statistical distribution of the mean squared weighted deviation-reply. Chemical Geology, 94(3): 242 ~ 243.
- Wendt I, Carl C. 1991. The statistical distribution of the mean squared weighted deviation. Chemical Geology, 86(4): 275 ~ 285.
- Wetherill G W. 1956. Discordant uranium - lead Ages, I. Transactions - American Geophysical Union. 37(3): 320 ~ 326.
- York D. 1966. Least squares fitting of a straight line. Canadian Journal of Physics, 44: 1079 ~ 1086.
- York D. 1967. The best isochron. Earth and Planetary Science Letters, 2: 479 ~ 482.
- York D. 1969. Least squares fitting of a straight line with correlated errors. Earth and Planetary Science Letters, 5: 320 ~ 324.
- Zheng Y F. 1989. Influences of the nature of the initial Rb-Sr system on isochron validity. Chemical Geology, 80(1): 1 ~ 16.

Methods for Least Squares Fitting of a Straight Line and Their Application in Geochronology

FAN Wenbo¹⁾, WANG Wenlong²⁾, CHEN Yuelong¹⁾, LI Dapeng¹⁾

1) School of Earth Sciences and Resources, China University of Geosciences, Beijing, 100083;

2) School of Mathematics, Beijing Institute of Technology, Beijing, 100081

Abstract: The methods for least squares fitting of a straight line are applied widely in natural sciences and social sciences, and geology is no exception, especially in geochronology. It is the normal/ordinary least squares that is well known by much of the population, where errors in only one coordinate are considered. In fact, The least square linear regressions is much more complex, especially when errors in both coordinates even the error correlations are considered. Based on the previous research, an overview of the methods for least squares fitting and its application in geochronology are provided here. These methods includes normal/ordinary least squares, normal/ordinary weighting least squares, standard weighting model, standard independent weighting model, independent weighting model, and error correlated independent weighting model. What's more, we give a brief study of the parameter of *MSWD*, which is commonly used but makes many people confused in geochronological study. It is very important for us to know all of these algorithm and their underlying assumption exactly, which will of course benefit us a lot in choosing the best approach to calculate the parameters we need and assessing the results in geochronology. Our study will also help improve the basic methods for least squares fitting in mathematics and could be used in other areas.

Key words: methods for least squares fitting; fitting of a straight line; geochronology; *MSWD*; iterative method