非线性函数的协方差传播公式

徐 培 亮

摘 要

本文首先导出非线性函数的协方差传播的一般公式，将此公式简化，可得到一次的和含有二次项的协方差传播公式。然后证明了H.Wolf的含有二次项的传播公式是该公式的一个近似。

【关键词】 非线性；方差协方差传播；应用

众所周知，非线性的观测值函数的精度计算一般是采用只含一次项的误差传播定律。H.Wolf在1961年[1]提出了一个含有二次项的误差传播公式

$$\hat{m}_f^2=\sum_{i=1}^{n}a_i^2m_i^2+\frac{3}{4}\sum_{i=1}^{n}(a_{ii}m_i^2)^2+\frac{1}{2}\sum_{i<j}(a_{ii}a_{jj}+2a_{ij}^2)m_i^2m_j^2$$

其中，a_i，a_{ii}，a_{ij}分别是函数f的关于x_i的一阶导数，f的关于x_i的二阶导数，f的关于x_i、x_j的二阶偏导数，即

$$a_i=\frac{\partial f}{\partial x_i},\qquad a_{ii}=\frac{\partial^2 f}{\partial x_i^2},\qquad a_{ij}=\frac{\partial^2 f}{\partial x_i\partial x_j}$$

下面从方差的一般概念出发，推导一个含有二次项的误差传播公式，近似可以得到含有一次项和二次项的中误差传播公式，并简略地讨论观测值相关时的公式。

一、一般公式

设观测值及其误差为

$$X=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix},\qquad \Delta=X-E(X)=\begin{pmatrix}x_1-E(x_1)\\x_2-E(x_2)\\\vdots\\x_n-E(x_n)\end{pmatrix}=\begin{pmatrix}\Delta_1\\\Delta_2\\\vdots\\\Delta_n\end{pmatrix}$$

且

$$\left.\begin{aligned}&E(X)=\mu_x,\qquad E(\Delta)=0\\&D_x=E(\Delta\Delta^T)\end{aligned}\right\}$$

又设观测值的函数

本文1985年8月收到.

$$Y=\varphi(x_1,\ x_2,\ \cdots x_n)=\varphi(X) \tag{1}$$

现在来求函数Y的方差。

将Y展开成台劳级数为

$$\begin{aligned}Y=&\varphi(E(X))+(\frac{\partial\varphi}{\partial x_1})_E\Delta_1+\cdots\cdots+(\frac{\partial\varphi}{\partial x_n})_E\Delta_n\\&+\frac{1}{2!}\left\{(\frac{\partial^2\varphi}{\partial x_1^2})_E\Delta_1^2+(\frac{\partial^2\varphi}{\partial x_1\partial x_2})_E\Delta_1\Delta_2+\cdots+(\frac{\partial^2\varphi}{\partial x_n^2})_E\Delta_n^2\right\}\\&+\frac{1}{3!}\left\{(\frac{\partial^3\varphi}{\partial x_1^3})_E\Delta_1^3+(\frac{\partial^3\varphi}{\partial x_1^2\partial x_2})_E\Delta_1^2\Delta_2+\cdots\right\}+\cdots\end{aligned} \tag{2}$$

令

$$\left.\begin{aligned}&\alpha^T=\left[\left(\frac{\partial\varphi}{\partial x_1}\right)_E,\ \left(\frac{\partial\varphi}{\partial x_2}\right)_E,\ \cdots\cdots,\ \left(\frac{\partial\varphi}{\partial x_n}\right)_E\right]\\&\beta=\begin{pmatrix}\left(\frac{\partial^2\varphi}{\partial x_1^2}\right)_E & \left(\frac{\partial^2\varphi}{\partial x_1\partial x_2}\right)_E & \cdots\cdots\left(\frac{\partial^2\varphi}{\partial x_1\partial x_n}\right)_E\\ \left(\frac{\partial^2\varphi}{\partial x_2\partial x_1}\right)_E & \left(\frac{\partial^2\varphi}{\partial x_2^2}\right)_E & \cdots\cdots\left(\frac{\partial^2\varphi}{\partial x_2\partial x_n}\right)_E\\ \cdots\cdots & \cdots\cdots & \cdots\cdots\\ \left(\frac{\partial^2\varphi}{\partial x_n\partial x_1}\right)_E & \left(\frac{\partial^2\varphi}{\partial x_n\partial x_2}\right)_E & \cdots\cdots\left(\frac{\partial^2\varphi}{\partial x_n^2}\right)_E\end{pmatrix}\\&R_1=\frac{1}{3!}\left\{\left(\frac{\partial^3\varphi}{\partial x_1^3}\right)_E\Delta_1^3+\left(\frac{\partial^3\varphi}{\partial x_1^2\partial x_2}\right)_E\Delta_1^2\Delta_2+\cdots\cdots\right\}+\cdots\end{aligned}\right\} \tag{3}$$

其中偏导数$\left(\frac{\partial\varphi}{\partial x_i}\right)_E$，$\left(\frac{\partial^2\varphi}{\partial x_i\partial x_j}\right)_E$的下标E表示用数学期望代入计算。则（2）可写为

$$Y=\varphi(E(X))+\alpha^T\Delta+\frac{1}{2}\Delta^T\beta\Delta+R_1 \tag{4}$$

对上式取数学期望可得

$$E(Y)=\varphi(E(X))+\frac{1}{2}E(\Delta^T\beta\Delta)+E(R_1)$$

因

$$E(\Delta^T\beta\Delta)=E(tr(\Delta^T\beta\Delta))=E(tr(\beta\Delta\Delta^T))=tr(\beta D_X) \tag{5}$$

所以

$$E(Y)=\varphi(E(X))+\frac{1}{2}tr(\beta D_X)+E(R_1) \tag{6}$$

而

$$Y-E(Y)=\alpha^T\Delta+\frac{1}{2}(\Delta^T\beta\Delta-tr(\beta D_X))+R_1-E(R_1)$$

$$(Y-E(Y))^2=\alpha^T\Delta\Delta^T\alpha+\frac{1}{4}(\Delta^T\beta\Delta-tr(\beta D_X))^2+\alpha^T\Delta\Delta^T\beta\Delta-\alpha^T\Delta tr(\beta D_X)+\cdots\cdots$$

所以

$$D_Y=E\{(Y-E(Y))^2\}=\alpha^TD_X\alpha+\frac{1}{4}E\{(\Delta^T\beta\Delta)^2-2\Delta^T\beta\Delta tr(\beta D_X)+(tr(\beta D_X))^2\}$$
$$+E\{\alpha^T\Delta\Delta^T\beta\Delta\}-E\{\alpha^T\Delta tr(\beta D_X)\}+R$$

因式中

$$\left.\begin{aligned}&E\{\alpha^T\Delta tr(\beta D_X)\}=\alpha^TE(\Delta)tr(\beta D_X)=0\\&E\{\Delta^T\beta\Delta tr(\beta D_X)\}=(tr(\beta D_X))^2\end{aligned}\right\}$$

故

$$D_Y=\alpha^TD_X\alpha-\frac{1}{4}(tr(\beta D_X))^2+\frac{1}{4}E\{(\Delta^T\beta\Delta)^2\}+E\{\alpha^T\Delta\Delta^T\beta\Delta\}+R \tag{7}$$

因

$$\Delta^T\beta\Delta=\overline{\beta}^T\overline{A}$$

其中，$A=\Delta\Delta^T$，“—”表示矩阵按列拉直成向量。

所以

$$(\Delta^T\beta\Delta)^2=\overline{\beta}^T\overline{A}\,\overline{A}^T\overline{\beta}=\overline{\beta}^T(A\otimes A)\overline{\beta}$$

其中“⊗”表示矩阵的克罗内克积。

则 $(\Delta^T\beta\Delta)^2$ 的数学期望为

$$E(\Delta^T\beta\Delta)^2=E\{\overline{\beta}^T(A\otimes A)\overline{\beta}\}$$

$$=\overline{\beta}^T\begin{pmatrix}\mu_1^4 & \rho_{1112}\cdots\cdots\rho_{111n}\cdots\cdots\rho_{1n11} & \rho_{1n12}\cdots\cdots\rho_{1n1n}\\ \vdots & \vdots\quad\vdots\quad\vdots\quad\cdots\cdots\quad\vdots & \vdots\quad\cdots\cdots\quad\vdots\\ \rho_{11n1} & \rho_{11n2}\cdots\cdots\rho_{11nn}\cdots\cdots\rho_{1nn1} & \rho_{1nn2}\cdots\cdots\rho_{1nnn}\\ \vdots & \vdots\quad\vdots\quad\vdots\quad\vdots\quad\vdots & \vdots\quad\vdots\quad\vdots\\ \rho_{n111} & \rho_{n112}\cdots\cdots\rho_{n11n}\cdots\cdots\rho_{nn11} & \rho_{nn12}\cdots\cdots\rho_{nn1n}\\ \vdots & \vdots\quad\vdots\quad\vdots\quad\cdots\cdots\quad\vdots & \vdots\quad\vdots\quad\vdots\\ \rho_{n1n1} & \rho_{n1n2}\cdots\cdots\rho_{n1nn}\cdots\cdots\rho_{nnn1} & \rho_{nnn2}\cdots\cdots\mu_n^4\end{pmatrix}_{n^2\times n^2}\overline{\beta} \tag{8}$$

其中

$$\rho_{ijkl}=E\{(x_i-E(x_i))(x_j-E(x_j))(x_k-E(x_k))(x_l-E(x_l))\}$$

i，j，k，l 是自然数段 1 到 n 中的任一个数，但不能同时相等。

$$\mu_i^4 = E(\Delta_i^4), \quad i = 1, 2, \cdots n$$

它是随机变量 x_i 的四阶中心距。

因

$$\alpha^T\Delta\Delta^T\beta\Delta = \alpha^T\Delta\overline{A}^T\overline{\beta}$$

所以它的数学期望为：

$$E(\alpha^T\Delta\Delta^T\beta\Delta) = E(\alpha^T\Delta\overline{A}^T\overline{\beta}) = \alpha^T E(\Delta\overline{A}^T)\overline{\beta}$$

$$= \alpha^T\begin{pmatrix} \mu_1^3 & \rho_{112}\cdots\cdots\rho_{11n}\cdots\cdots\rho_{1n1}\cdots\cdots\rho_{1nn} \\ \rho_{211} & \rho_{212}\cdots\cdots\rho_{21n}\cdots\cdots\rho_{2n1}\cdots\cdots\rho_{2nn} \\ \cdots\cdots & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \rho_{n11} & \rho_{n12}\cdots\cdots\rho_{n1n}\cdots\cdots\rho_{nn1}\cdots\cdots\mu_n^3 \end{pmatrix}\overline{\beta} \tag{9}$$

其中

$$\rho_{ijk} = E(\Delta_i\Delta_j\Delta_k) = E\{(x_i - E(x_i))(x_j - E(x_j))(x_k - E(x_k))\}$$

i，j，k 是自然数 1 到 n 中任一数，但不全相等。

$$\mu_i^3 = E(\Delta_i^3) = E\{(x_i - E(x_i))^3\} \qquad i = 1, 2, \cdots n$$

它是随机变量 x_i 的三阶中心距。

方差计算式（7）中的R为

$$R = E\{2\alpha^T\Delta(R_1 - E(R_1)) + (\Delta^T\beta\Delta - tr(\beta D_x))(R_1 - E(R_1))\} + E\{(R_1 - E(R_1))^2\} \tag{10}$$

顾及（8），（9），（10）三式就可以计算任一观测值函数的精度。公式（7）是精度计算的一般公式，可以用在一般情况下的精度计算，例如，x_1，x_2，$\cdots x_n$彼此相关，或是其中的一部分相关，或是彼此独立，它们的分布可以是正态的，也可以是其它类型的。

二、独立观测值含有二次项的协方差传播公式

当x_1，x_2，…，x_n是彼此独立的观测值时，

$$d_{x_ix_j} = E(\Delta_i\Delta_j) = \begin{cases} d_{ii} & i = j \\ 0 & i \neq j \end{cases}$$

即 D_x 为对角阵。

$$\rho_{ijkl} = \begin{cases} d_{ii}d_{kk} & i = j \quad k = l \quad i \neq k \\ 0 & \text{其它} \end{cases}$$

$$\mu_i^4 = E(\Delta_i^4)$$

且当x_1，x_2，…，x_n服从正态分布时，

$$E\{\alpha^T\Delta\Delta^T\beta\Delta\} = 0$$

$$\mu_i^4=\int_{-\infty}^{\infty}\Delta_i^4\frac{1}{\sqrt{2\pi d_{ii}^{1/2}}}e^{-\frac{\Delta_i^2}{2d_{ii}}}d\Delta_i=3d_{ii}^2$$

如果我们只考虑到二次展开项，略去高阶项R，便得到一个含有二次项的中误差传播公式，用$\tilde{m}_y^2$记之，得

$$E\{(\Delta^T\beta\Delta)^2\}=\sum_{i=1}^n\sum_{j=1}^n\left(\frac{\partial^2\varphi}{\partial x_i^2}\right)_E\left(\frac{\partial^2\varphi}{\partial x_j^2}\right)_E d_{ii}d_{jj}+2\sum_{i=1}^n\sum_{j=1}^n\left(\frac{\partial^2\varphi}{\partial x_i\partial x_j}\right)_E^2 d_{ii}d_{jj}$$

$$=3\sum_{i=1}^n\left\{\left(\frac{\partial^2\varphi}{\partial x_i^2}\right)_E d_{ii}\right\}^2+2\sum_{i<j}d_{ii}d_{jj}\left\{\left(\frac{\partial^2\varphi}{\partial x_i^2}\right)_E\left(\frac{\partial^2\varphi}{\partial x_j^2}\right)_E+2\left(\frac{\partial^2\varphi}{\partial x_i\partial x_j}\right)_E^2\right\}$$

所以

$$\tilde{m}_y^2=\sum_{i=1}^n\left(\frac{\partial\varphi}{\partial x_i}\right)_E^2 d_{ii}+\frac{1}{2}D^TB_ED=\alpha^TD_x\alpha+\frac{1}{2}D^TB_ED \tag{11}$$

其中

$$D^T=(d_{11},\ d_{22}\cdots d_{nn})$$

$$B_E=\begin{pmatrix}\left(\frac{\partial^2\varphi}{\partial x_1^2}\right)_E^2 & \left(\frac{\partial^2\varphi}{\partial x_2\partial x_1}\right)_E^2 & \cdots\cdots\left(\frac{\partial^2\varphi}{\partial x_1\partial x_n}\right)_E^2\\ \left(\frac{\partial^2\varphi}{\partial x_2\partial x_1}\right)_E^2 & \left(\frac{\partial^2\varphi}{\partial x_2^2}\right)_E^2 & \cdots\cdots\left(\frac{\partial^2\varphi}{\partial x_2\partial x_n}\right)_E^2\\ \cdots\cdots & \cdots\cdots & \cdots\cdots\\ \left(\frac{\partial^2\varphi}{\partial x_n\partial x_1}\right)_E^2 & \left(\frac{\partial^2\varphi}{\partial x_n\partial x_2}\right)_E^2 & \cdots\cdots\left(\frac{\partial^2\varphi}{\partial x_n^2}\right)_E^2\end{pmatrix}$$

（11）式就是当随机变量x_1、x_2、…、x_n相互独立且服从正态分布时，一般函数$Y=\varphi(x_1,\ x_2,\ \cdots x_n)$的顾及二次项的精度计算公式。

通常，E(X)是未知的，但已知一组观测值x_1，x_2，…x_n，在公式（11）中用x_1，…，x_n代替其相应的期望值$E(x_1)$，$E(x_2)$，…$E(x_n)$，得到Y的精度估计公式，用$\hat{m}_y^2$记之。

$$\hat{m}_y^2=\alpha^TD_x\alpha+\frac{1}{2}D^TBD \tag{12}$$

其中B矩阵是B_E中把X代替E(X)后得到的。

公式（12）就是含有二次项的中误差传播公式。

特别是，如果只取公式（7）中右边的第一项，并用观测值x_1，x_2，…，x_n代替相应的期望值，即得到仅含有一次项的中误差传播公式：

$$\hat{m}_y^2=\alpha^TD_x\alpha$$

因此，含有一次项的协方差传播公式是本文公式的近似。

改写公式（12）成为：

$$\widehat{m}_y^2=\sum_{i=1}^{n}\left(\frac{\partial\varphi}{\partial x_i}\right)^2 d_{ii}+\frac{3}{4}\sum_{i=1}^{n}\left(\frac{\partial^2\varphi}{\partial x_i^2}\right)^2 d_{ii}^2+\frac{1}{2}\sum_{i<j}\left[\left(\frac{\partial^2\varphi}{\partial x_i^2}\right)\left(\frac{\partial^2\varphi}{\partial x_j^2}\right)\right.$$

$$\left.+2\left(\frac{\partial^2\varphi}{\partial x_i\partial x_j}\right)^2\right]d_{ii}d_{jj}-\frac{1}{4}\left(\sum_{i=1}^{n}\left(\frac{\partial^2\varphi}{\partial x_i^2}\right)d_{ii}\right)^2$$

如果我们略去上式右边的最后一项，即得到H.Wolf的含有二次项的中误差传播公式。因此H.Wolf的含有二次项的中误差传播公式是本文公式的近似，且它们有关系式：

$$\widehat{m}_y^2=\widehat{m}_{yw}^2-\frac{1}{4}\left(\sum_{i=1}^{n}\left(\frac{\partial^2\varphi}{\partial x_i^2}\right)d_{ii}\right)^2 \tag{13}$$

其中，$\widehat{m}_{yw}^2$ 代表H.Wolf的计算公式。

这一结论的证明是容易的，只要我们在公式（6）中略去右边第二项，并按本文的推导过程即可得到H.Wolf的计算公式。

三、观测值相关时含有二次项的协方差传播公式

在公式（7）中略去高阶项R即得到观测值相关时的含有二次项的协方差传播公式。

根据矩阵分解得到另一个计算公式。

设随机变量 x_1，$x_2\cdots$，x_n 是具有方差阵为 D_x 的正态分布随机向量，则由于 D_x 的正定性可以分解为

$$D_x=T\lambda T^T$$

T是 $n\times n$ 阶正交阵，λ 是 D_x 的 n 个特征值构成的对角阵。

作变量替换

$$Z=T^TX$$

则

$$E(Z)=T^TE(X),\qquad D_Z=\lambda$$

即向量Z中的各分量 z_1，z_2，…，z_n 相互独立，具有方差阵为 λ。

根据本文的推导过程可得

$$\widetilde{m}_y^2=\alpha^TD_x\alpha+\frac{1}{2}g^TCg \tag{14}$$

其中，α 的表示式如式（3），

$$g^T=(\lambda_1,\ \lambda_2,\ \cdots,\ \lambda_n),$$
$$C=(c_{ij}^2)=([T_i^T\beta T_j]^2)$$

T_i 是正交阵T的第 i 列向量。

公式（14）就是观测值相关时的含有二次项的协方差传播公式。

四、含有二次项的协方差传播公式的应用举例

例 1：图 1 是一个简单的变形监测控制网，A、B 为固定点，C点布设在可能的变形区。设C_I、C_{II}分别表示C点的第一、二期观测。为了检验C点是否产生了变动，有时利用 C_I和C_{II}间的距离及其精度构成统计量进行检验[4]。记C_I 和C_{II}的坐标差分别为 Δx、Δy，则

$$f = s^2 = \Delta x^2 + \Delta y^2$$

由于任何相关的二个正态随机变量都可以独立化。因此，为了便于计算、比较。不妨设 Δx，Δy 是两个相互独立的随机变量，且 $E(\Delta x) = 1\text{mm}$，$E(\Delta y) = 1\text{mm}$，$m^2_{\Delta x} = 1\text{mm}^2$，$m^2_{\Delta y} = 1\text{mm}^2$。

按惯用的协方差传播公式可得到 f 的方差为

$$\tilde{m}_1^2 = 8\text{mm}^4$$

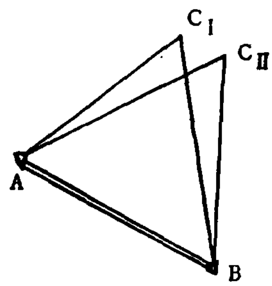

图 1

而

$$a_1 = \frac{\partial f}{\partial \Delta x}\Big|_E = 2\cdot\Delta x\Big|_E = 2$$

$$a_2 = \frac{\partial f}{\partial \Delta y}\Big|_E = 2\cdot\Delta y\Big|_E = 2$$

$$a_{11} = \frac{\partial^2 f}{\partial \Delta x^2}\Big|_E = 2, \quad a_{22} = \frac{\partial^2 f}{\partial \Delta y^2}\Big|_E = 2, \quad a_{12} = \frac{\partial^2 f}{\partial \Delta x \partial \Delta y}\Big|_E = 0$$

所以，按H.Wolf 的公式得到 f 的方差为

$$\tilde{m}_2^2 = \sum_{i=1}^{2} a_i^2 m_i^2 + \frac{3}{4}\sum_{i=1}^{2}(a_{ii} m_i^2)^2 + \frac{1}{2}(a_{11}a_{22} + 2a_{12}^2)m_1^2 m_2^2$$

$$= 8 + \frac{3}{4}(2^2 + 2^2) + \frac{1}{2}(2 \times 2 + 0) \times 1 \times 1 = 16\text{mm}^4$$

按公式（11）计算得到 f 的方差为

$$\tilde{m}_3^2 = \alpha^T D_x \alpha + \frac{1}{2} D^T B_E D = \begin{pmatrix} 2 \\ 2 \end{pmatrix}^T \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \end{pmatrix} + \frac{1}{2}(1,\ 1)\begin{pmatrix} 2^2 & 0 \\ 0 & 2^2 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 12\text{mm}^4$$

而根据正态随机向量的概率密度计算得到 f 的方差为

$$\tilde{m}_4^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [\Delta x^2 + \Delta y^2 - E(\Delta x^2 + \Delta y^2)]^2 \frac{1}{2\pi} e^{-\frac{1}{2}[(\Delta x - 1)^2 + (\Delta y - 1)^2]} d_{\Delta x} d_{\Delta y}$$

$$= 12\text{mm}^4$$

例 2：设 $y = x_1^2 + x_2^2 + \cdots + x_n^2$，$x_1$，$x_2$，…，$x_n$ 相互独立，且都服从标准正态分布，则

由公式（11）计算得到

$$\tilde{m}_f^2 = 2n$$

由H. Wolf 的公式计算得到

$$\tilde{m}_{yw}^2 = 2n + n^2$$

且由统计学易知

$$y \sim \chi^2(n)$$

$$\therefore \quad \tilde{m}_y^2 = 2n$$

由此可见，对于二次多项式，本文得到的含有二次项的协方差传播公式与真值一样，而H.Wolf 的是它的近似。且在变形分析中，在某些情况下，仅利用惯用的协方差传播公式是不够的，顾及二次项是必要的。

从文中的例子可以看到，当观测值的数学期望很小时，可能要考虑使用含有二次项的公式，但视函数形式而定，也就是说，精度计算时是否使用含有二次项的公式，决定于函数的形式、观测值数学期望及其方差的大小。

本文在写作过程中得到了刘大杰付教授、陈永奇博士的热情帮助，表示衷心感谢。

参 考 文 献

[1] H.Wolf, Das Fehlerfortpflanzungsgesetz mit Gliedern II Ordnung, ZFV. März. 1961.

[2] H.Wolf 著，方佩竹译，平差计算（实用公式），测绘出版社，1983.

[3] 陶本藻，含有二次项的误差传播定律的推导，测绘通报，3,1983.

[4] J.Martusewicz, New Conception of Determination of Displacements, FIG XVII International Congress, Sofia, Bulgaria, 1983.

Variance—Covariance Propagation for a Nonlinear Function

Xu Peiliang

Abstract

The general expression of variance—covariance propagation for a nonlinear function is derived. By simplifying it, the corresponding formulas with Order I and Order II are obtained. The formula with Order II presented by H.Wolf is proved to be an approximation of the above formula.

[Key words] nonlinear; variance-covariance propagation; application