Formal Difference Analysis and Unification on p-Norm Distribution Density Functions

LIU Zhengcai  ZHU Jianjun  WANG Huaiyu

ABSTRACT  The cause of the formal difference of p-norm distribution density functions is analyzed, two problems in the deduction of p-norm formulating are improved, and it is proved that two different forms of p-norm distribution density functions are equivalent. This work is useful for popularization and application of the p-norm theory to surveying and mapping.

KEYWORDS  p-norm distribution; density function; difference; equivalence

CLC NUMBER  P207

Introduction

Early in 1965, a researcher from Former Soviet Union put forward p-norm distribution density function of one variable [12]. In 1980, an American researcher also gave an analogous conclusion [22]. In 1993, a Chinese researcher, Sun Haiyan also deduced the p-norm distribution density function of one variable, imitating Gaussian deduction of normal distribution density function [32]. In 1991, 1996 and 1997, Sun Haiyan, Zhang Fangren, and Yu Zongchou deduced the p-norm distribution density function of multi-variable which was applied to the data processing in surveying and mapping [4-6]. Zhou Shijian put forward another form of one variable function [7], another form of multi-variable function [8], and a set of p-norm theory of measurement adjustment [9-12]. Sun put forward the theory of maximum likelihood estimation through the application of p-norm distribution to the field of measurement data processing. Zhou studied the p-norm robust estimation. However, there are formal differences of p-norm distribution density functions of one variable or multi-variables between Sun’s and Zhou’s, and the equivalence of them has not been proved for many years. It is inconvenient to popularize the p-norm theory in surveying and mapping.

1  Formal difference analysis on p-norm distribution density functions of one variable

Reference [3] deduced the p-norm distribution density function of one variable by imitating Gaussian deduction of normal distribution density function. Making \(n\) observations on some true value \(x\) with the same precision, we can get observation vector \(L\) with the following true-errors:

\[
\Delta_i = x - L_i \quad (i = 1, 2, \ldots, n)
\]

Let \(\hat{x}\) be the maximum likelihood estimation of \(x\), \(v_i\) be the correction of \(i\)th observation \(L_i\), then we can obtain

\[
v_i = x - L_i \quad (i = 1, 2, \ldots, n)
\]

In order to realize

\[
d \left( \sum_{i=1}^{n} |L_i - \hat{x}|^p \right)^{\frac{1}{p}} = \min
\]

the \(\hat{x}\) must satisfy such a demand that

\[
\frac{d}{dx} \left( \sum_{i=1}^{n} |L_i - x|^p \right) = 0
\]
because
\[ \frac{d}{dx} \left( |L_i - \hat{x}|^p \right) = -p |L_i - \hat{x}|^{p-2} (L_i - \hat{x}) \]

(5)

Substituting Eqs. (2) and (5) into Eq. (4), we obtain

\[ \sum_{i=1}^{n} p |v_i|^{p-2} v_i = 0 \]

(6)

Let
\[ y_i = p |v_i|^{p-2} v_i \]

(7)

Eqs. (6) and (7), with the form of carrying \( p \), are slightly different from those in Reference [3] in form, so list them here. Others of the detail deduction are the same as in Reference [3], so they are omitted here. The final form, given by Sun, of \( p \)-norm distribution density function of one variable is

\[ f(\Delta) = \frac{p^{(1-1/p)}}{2 \sigma^{3/2} (1/p) \Gamma(1/p)} \exp \left\{ -\frac{1}{p} \left( \frac{\Delta}{\sigma} \right)^p \right\} \]

(8)

Zhou's deduction eliminated coefficient \( p \) from Eq. (6), thus in Eqs. (6) and (7) there are no coefficient \( p \). In the same way, we can obtain the Zhou's \( p \)-norm distribution density function of one variable as follows,

\[ f(\Delta) = \frac{p^{(1-1/p)}}{2 \sigma^{3/2} (1/p) \Gamma(1/p)} \exp \left\{ -\frac{1}{p} \left( \frac{\Delta}{\sigma} \right)^p \right\} \]

(9)

where \( \sigma \) is a scale parameter.

From the above 2 deductions, we know that whether the coefficient \( p \) exists in Eqs. (6) and (7) or not will cause quite different forms of \( p \)-norm distribution density function of one variable. Using Eqs. (6) and (7) with coefficient \( p \), we obtain Sun's function Eq. (8). On the contrary, using Eqs. (6) and (7) without coefficient \( p \), we obtain Zhou's function Eq. (9). We call the difference as \( p \) difference between 2 forms of \( p \)-norm distribution density of one variable.

2 Equivalence of two forms of \( p \)-norm distribution density of one variable

By definition of variance \( \sigma^2 \) and Eq. (9), we obtain that

\[ \sigma^2 = \int_{-\infty}^{\infty} \Delta^2 f(\Delta) d\Delta = \frac{2}{\Gamma(1/p)} \int_{0}^{\infty} \delta^{\frac{1}{p}-1} \exp \left\{ -\frac{1}{p} \left( \frac{\Delta}{\sigma} \right)^p \right\} d\Delta \]

Let
\[ y = \frac{1}{p} \left( \frac{\Delta}{\sigma} \right)^p \]

then
\[ \Delta = (p y)^{1/p}, d\Delta = \sigma y^{1-1/p} dy \]

substituting above formulas into Eq. (9), we obtain

\[ \sigma^2 = \frac{2}{\Gamma(1/p)} \int_{0}^{\infty} \delta^{\frac{1}{p}-1} \exp \left\{ -\frac{1}{p} \left( \frac{\Delta}{\sigma} \right)^p \right\} d\Delta \]

Consequently, we obtain

\[ \sigma = \sigma y^{1/p} \int_{0}^{\infty} \frac{1}{\Gamma(1/p)} \exp \left\{ -\frac{1}{p} \left( \frac{\Delta}{\sigma} \right)^p \right\} d\Delta \]

(10)

Putting Eq. (10) into Eq. (9), we obtain

\[ f(\Delta) = \frac{p^{(1-1/p)}}{2 \sigma^{3/2} (1/p) \Gamma(1/p)} \exp \left\{ -\frac{1}{p} \left( \frac{\Delta}{\sigma} \right)^p \right\} \]

(11)

The right side of Eq. (11) is completely the same as Eq. (8). Therefore, using Eqs. (9) and (10), we can deduce Eq. (8). In the above deductive process, every step is reversible; similarly, using Eqs. (8) and (10), we can also deduce the Eq. (9). Therefore, Eq. (9) is equivalent to Eq. (8).

3 Equivalence of \( p \)-norm distribution density functions of multi-variable

3.1 Two forms of \( p \)-norm distribution density functions of multi-variable

In Reference [4], Sun put forward the definition of \( p \)-norm distribution density function of multi-variable as follows. Let \( B = (b_j) \) be a symmetric positive definite matrix of order \( n \), \( |B| \) denote the determinant of \( B \), \( \mu = [\mu_1, \mu_2, \ldots, \mu_n]^T \) be any real-valued columned vector, \( U \) be such an invertible square matrix of order \( n \) that \( UU^T = B \) (matrix \( B \) is of positive definiteness, so \( U \) exist certainly, then the \( p \)-norm dis-
distribution density function of multi-variable holds the form as follows:

\[
p(X) = \frac{1}{(2\pi)^{n/2} |R^{(1/p)}| |B|^{1/2}} \cdot \exp \left\{ - \left[ \sqrt{\frac{2}{p}} \left| U^{-1}(X - \mu) \right|_p \right]^{p} \right\}
\]

where \( |X|_p = \left( \sum_{i=1}^{n} |x_i|^{p} \right)^{1/p}, p > 0 \); \( X = [x_1, x_2, \ldots, x_n]^T \).

If \( D_x \) denotes covariance matrix, then \( D_x^{1/2} \cdot (D_x^{1/2})^T = D_x \), and so we can obtain another form of the \( p \)-norm distribution density function of multi-variable as:

\[
p(X) = \frac{|D_x|^{1/2} \cdot (X - \mu)^{\frac{1}{2}}}{2 \Gamma^{(1/p)} |B|^{1/2}} \cdot \exp \left\{ - \left[ \left( \frac{\Gamma(3/p)}{\Gamma(1/p)} \right)^{\frac{1}{p}} \left| D_x \right|^{\frac{1}{2}} (X - \mu) \right|_p \right\}
\]

Zhou Shijian’s multi-dimensional function of \( p \)-norm distribution density \( f_p(X) \) holds form as follows (changed original scale matrix to \( T \)):

\[
f_p(X) = \frac{p(1/p)^n}{2 \Gamma^{(1/p)} |T|^{1/2}} \cdot \exp \left\{ - \left[ \left( \frac{\Gamma(3/p)}{\Gamma(1/p)} \right)^{\frac{1}{p}} \left| T \right|^{\frac{1}{2}} (X - \mu) \right|_p \right\}
\]

where \( \mu \) indicates locating parameter, i.e. mathematical expectation vector; \( T \) denotes scale parameter matrix.

3.2 Relationship between covariance matrix and scaling parameter matrix

To prove the equivalence of Eqs. (13) and (14), we can find out the relation between covariance matrix and scaling parameter matrix, imitating the processing in Reference [4]. Let

\[
T = UU^T = (\tau_y)_{n \times n}
\]

\[
U = (u_y)_{n \times n} = T^{1/2}
\]

then

\[
\tau_y = \sum_{i=1}^{n} u_y^i
\]

\[
\tau_x = \sum_{i=1}^{n} u_x u_{\mu}
\]

and

\[
|U| = |T|^{1/2} = |T|^{1/2}
\]

By means of \( X = UY + \mu \), we obtain

\[
x_i = \mu_i + \sum_{j=1}^{n} u_{ij} y_j
\]

\[
(x_i - \mu_i)(x_j - \mu_j) = \sum_{k=1}^{n} u_{ik} u_{jk} y_k + \sum_{k \neq h} u_{ik} u_{jh} y_k y_h
\]

\[
dX = |U| dY = |T|^{1/2} dY
\]

Let \( t = \frac{1}{p} x^p \), then \( dt = x^{p-1} dx = (pt)^{\frac{p-1}{p}} dt \), we obtain

\[
\int_{-\infty}^{+\infty} e^{-\frac{1}{2} t^2} dt = \int_{-\infty}^{+\infty} dt = \int_{0}^{+\infty} e^{-\frac{1}{2} \frac{1}{t^2}} dt = p^{\frac{1}{2}} \Gamma\left( \frac{1}{p} \right)
\]

\[
\int_{-\infty}^{+\infty} y^p e^{-\frac{1}{2} t^2} dt = \int_{0}^{+\infty} y^p e^{-\frac{1}{2} \frac{1}{t^2}} dt = p^{\frac{1}{2}} \Gamma\left( \frac{3}{2} \right)
\]

\[
\int_{-\infty}^{+\infty} \exp \left\{ -\frac{1}{p} \left| Y \right|_p \right\} dY = \int_{-\infty}^{+\infty} \exp \left\{ -\frac{1}{p} \left| T \right|_p \right\} dY = \int_{-\infty}^{+\infty} \exp \left\{ -\frac{1}{p} \left| Y \right|_p \right\} dY
\]

Covariance between \( x_1, x_2 \) holds form as follows:

\[
cov(x, y) = \mathbb{E}[(x_i - \mu_j)(x_j - \mu_j)] = \int_{\mathbb{R}^n} (x_i - \mu_i)(x_j - \mu_j)f_p(X) dX
\]

\[
= \frac{p^{\frac{1}{2} - \frac{1}{p} n}}{2 \Gamma^{(1/p)} |T|^{1/2}} \int_{\mathbb{R}^n} \left\{ \sum_{i=1}^{n} u_{ii} y_i + \sum_{i \neq h} u_{ii} u_{hh} y_h y_h \right\} \cdot
\]

\[
\frac{e^{-\frac{1}{2} \frac{1}{t^2}}}{|T|^{1/2}} dt = \frac{e^{-\frac{1}{2} \frac{1}{t^2}}}{|T|^{1/2}} dt
\]
Putting Eqs. (20)-(24) into Eq. (25), we obtain

\[
\text{cov}(x_i, x_j) = \frac{\rho^{(1/2)_n}}{\Gamma_n^m} \left[ \sum_{k=h}^n u_k u_k \cdot \mathbf{y}_k \cdot \mathbf{y}_k \cdot \mathbf{y}_k \cdot \mathbf{y}_k \right] \exp \left\{ \frac{1}{\rho} \sum_{i=1}^n |x_i|^p \right\} dy_1 dy_2 \ldots dy_n.
\]

(25)

Thus the covariance matrix is obtained as follows:

\[
D_x = \rho^{1/2} \frac{\Gamma(3/p)}{\Gamma(1/p)} T
\]

i.e.,

\[
D_x^{1/2} = \left[ \frac{\rho^{1/2}}{\sqrt{\Gamma(1/p)}} \right] \left[ \frac{\Gamma(3/p)}{\Gamma(1/p)} T^{1/2} \right]^{T}
\]

\[
D_x^{1/2} = \left[ \frac{\rho^{1/2}}{\sqrt{\Gamma(1/p)}} \right]^{1/2}
\]

\[
|D_x|^{1/2} = \rho^{1/2} \left( \frac{\Gamma(3/p)}{\Gamma(1/p)} \right)^{1/2} |T|^{1/2}
\]

(26)

Furthermore,

\[
T = \rho^{-\frac{3}{2}} \frac{\Gamma(1/p)}{\Gamma(3/p)} D_x
\]

\[
T^{-1} = \rho^{1/2} \frac{\Gamma(3/p)}{\Gamma(1/p)} D_x^{-1}
\]

3.3 Equivalence of two different forms of \( p \) distribution density functions of multi-variable

Putting Eqs. (26) and (27) into Eq. (14), the equivalent proof can be given as follows:

\[
f_p(X) = \frac{\rho^{(1/2)_n}}{2\Gamma_n^m \Gamma(1/p)} |T|^{1/2} \cdot
\]

\[
\exp \left\{ -\frac{1}{\rho} \left[ \|T^{-1/2} (X - \mu) \|_p \right] \right\} = \rho^{(1/2)_n}
\]

\[
2\Gamma_n^m \Gamma(1/p) |D_x|^{1/2} \rho^{1/2} \frac{\Gamma(3/p)}{\Gamma(1/p)} |D_x|^{1/2} \cdot
\]

\[
\exp \left\{ -\frac{1}{\rho} \left[ \|X - \mu \|_p \right] \right\} = \rho(X)
\]

Consequently, from Eqs. (14), (26) and (27), we can deduce Eq. (13). Every step in the above deduction is reversible. Meanwhile the Eq. (14) can be also deduced from Eqs. (13), (26) and (27). Therefore, Eq. (13) is equivalent to Eq. (14). Sun’s \( p \)-norm distribution density function of multi-variables is equivalent to Zhou’s.

Although the forms are different, \( p \)-norm distribution density functions of one or many variables proposed by Sun Haiyan are equivalent to the corresponding one proposed by Zhou Shijian. Zhou’s has slightly more concise forms, but the scale parameter symbol \( \sigma \) and \( D \) do not accord with our professional custom (by means of our professional custom, \( \sigma \) denotes root mean square error, \( D \) denotes covariance matrix, but Zhou’s are not), so they are easily misunderstood. Therefore we suggest that a pair of Greek alphabets \( \tau \) and \( T \) denote scale parameter and scale parameter matrix separately.

ACKNOWLEDGEMENTS

Our thanks to Prof. Sun Haiyan and Zhou Shijian for giving support and encouragement.

(Continued on Page 186)