Rigorous Method of Weights Calculation in Adjustment of Measurement Data with Different Qualities

Ye Xiaoming1*  Ding shijun1

1. School of Geodesy and Geomatics, Wuhan University, Wuhan, Hubei, China, 430079.

Abstract: In the traditional measurement theory, precision is defined as the dispersion of measured value, and is used as the basis of weights calculation in the adjustment of measurement data with different qualities, which leads to the trouble that trueness is completely ignored in the weight allocation. In this paper, following the pure concepts of probability theory, the measured value (observed value) is regarded as a constant, the error as a random variable, and the variance is the dispersion of all possible values of an unknown error. Thus, a rigorous formula for weights calculation and variance propagation is derived, which solves the theoretical problem of determining the weight values in the adjustment of multi-channel observation data with different qualities. The results show that the optimal weights are not only determined by the covariance array of observation errors, but also related to the model of adjustment.

Keywords: measurement theory, random variable, uncertainty, survey adjustment, weight.

1. Introduction

The traditional measurement error theory is interpreted according to the logic of error classification, in which the measured value (observed value) and random error are regarded as random variables, and the variance is the dispersion of the measured value or random error; while the systematic error is regarded as constant without variance[1,2,3,4,5,6]. In this way, when multiple observations with different qualities are obtained from a single measurand, according to the conceptual logic of existing measurement theory, the weight of each observed value is calculated by its precision, and finally the weighted average value of each observed value is taken as the final measured value.

However, the biggest trouble faced by this method is that the systematic error of each observation is different from each other. Some observations with large systematic error but small random error get larger weights, which can make the final measured value have better precision, but its trueness will be worse. Although we can explain the existing theory based on the assumption that the systematic error does not exist or can be ignored, this assumption is usually not tenable in actual measurement.

For example, in the field of geodesy, the multiplicative constant error of a geodimeter is considered as systematic error, its measured value given by metrology field is usually used to correct the observed value, and people think that the residual systematic error after correction can be ignored. However, the multiplicative constant error of geodimeter is originally the residual error after correction in the field of instrument manufacturing, which comes from the residual error of quartz crystal frequency after temperature correction, and still changes with temperature. In the field of instrument, the limited range of this error is also given. Therefore, it is actually meaningless to use a test sample given by metrology field to correct the observed value.

Therefore, this theory with obvious defects often makes the measurement practice in trouble.

For example, in 2020, the National Bureau of surveying, mapping and geographic information of China announced that the elevation of Mount Everest is 8848.86m, which is actually an adjustment value obtained from the trigonometric leveling data of geodimeter and GNSS satellite survey data according to a certain weight proportion. However, at present, we have not seen the

*xmye@sgg.whu.edu.cn
public information about weight calculation, nor the official data about error evaluation.

In recent years, literature \[7,8,9,10,11,12,13\] proposed an interpretation method of measurement theory based new concepts, which clearly pointed out that the traditional measurement theory's understanding of constant and random variable concepts is inconsistent with probability theory. Because the constant in probability theory refers to a numerical value, and the random variable refers to a quantity whose numerical value is unknown, this theory regards the observed value (measured value) as a constant, the error and the true value as random variables, and the variance of error is the dispersion of all possible values of the error. In this conceptual system, errors cannot be classified as systematic error and random error.

Although these documents have been published formally, they do not deal with the problem of weight calculation in multi-channel data adjustment, so this paper will complete the mathematical deduction of this calculation principle. This deduction will completely overcome the confusion of error categories in traditional measurement theory, and completely solve the probability evaluation of error of adjustment value.

2. Basic concepts

2.1. Constant and random variable

In probability theory, the constant is a numerical value, such as 100, 150, \(x=100, x=8848.86\), and so on. Unlike constants, random variable is an unknown quantity whose actual value cannot be given. Because the random variable is unknown or uncertain, we can only describe the probability range of its value. In order to study its probability range, it is necessary to study the distribution range of all its possible values (sample space), while all possible values refer to the set of test values of random variables under all permitted possible test conditions. Mathematical expectation and variance are the numerical expression of probability range of random variable.

2.2. Mathematical expectation and variance

For a random variable \(X\) with all possible values \(\{x_i\}\), there is \(X \in \{x_i\}\), \(p_i\) is the probability of occurrence of event \(X = x_i\), that is, \(p_i = P(X = x_i)\) (Continuous random variables correspond to the probability density function \(p(x)\)), and its mathematical expectation and variance are respectively defined as:

\[
\begin{align*}
E(X) &= \sum_{i=1}^{n} p_i x_i \quad \text{or} \quad E(X) = \int_{-\infty}^{+\infty} xp(x)dx \\
\sigma^2(X) &= E[X - E(X)]^2
\end{align*}
\] (2-1)

Its meaning is that although the random variable \(X\) is unknown, it exists within a probability interval with \(E(X)\) as the center and \(\sigma^2(X)\) as the width evaluation. In other words, the mathematical expectation \(E(X)\) and the variance \(\sigma^2(X)\) are the evaluation of the probability interval of \(X\).

For a constant \(C\), because all its possible values are itself, there are:

\[
\begin{align*}
E(C) &= C \\
\sigma^2(C) &= E[C - E(C)]^2 = 0
\end{align*}
\] (2-2)

For example, when there is \(x = 8848.86\), there must be \(E(x) = E(8848.86) = 8848.86\) and \(\sigma^2(x) = \sigma^2(8848.86) = 0\). It can be seen that in traditional measurement theory, using the form of \(\sigma^2(x)\) to express the precision of measured value is actually a conceptual confusion, and even according to the so-called frequency interpretation, the equation \(\sigma^2(x) = \sigma^2(8848.86) = 0\) also cannot be refuted.

2.3. The probability expressions of measured value, error and true value

Different from the traditional measurement theory, in references\[7,8,9,10,11,12,13\], Fig 1 is used to
interpret its basic measurement concepts. Among them, an observed value \( x_0 \) is a sample within all possible observed values \( \{ x_i \} \) and is used as the final measured value, and \( \{ x_i \} \) is the sample space of random variable \( X \).

Because the measured value \( x_0 \) is a numerical value and belongs to a constant, according to the formula (2-2), there are:

\[
\begin{align*}
E(x_0) &= x_0 \\
\sigma^2(x_0) &= 0
\end{align*}
\]  

(2-3)

For the unknown deviation \( \Delta_A = x_0 - E(X) \), because it has the same sample space \( \{ x_i - E(X) \} \) as the random variable \( \Delta X = X - E(X) \), we can use \( \Delta X = X - E(X) \) to express the deviation \( \Delta_A \), that is, \( \Delta_A = X - E(X) \). Therefore, there are

\[
\begin{align*}
E(\Delta_A) &= E[X - E(X)] = 0 \\
\sigma^2(\Delta_A) &= E(\Delta_A^2) = E[(X - E(X))^2]
\end{align*}
\]  

(2-4)

Its concept meaning is that variance \( \sigma^2(\Delta_A) \) is the dispersion of all possible values \( \{ x_i - E(X) \} \) of deviation \( \Delta_A \), and is the evaluation value of the probability interval of deviation \( \Delta_A \).

For the unknown deviation \( \Delta_B = E(X) - X_T \), because it comes from the previous measurement, its formation principle is similar to that of deviation \( \Delta_A \), and it also has all its possible values, so there are

\[
\begin{align*}
E(\Delta_B) &= 0 \\
\sigma^2(\Delta_B) &= E(\Delta_B^2)
\end{align*}
\]  

(2-5)

For example, the multiplicative constant error of a geodimeter, which is generally considered as systematic error without variance in the field of geodesy, is actually the output error in the field of instrument manufacturing, and the instrument manufacturer also uses its error samples under all possible measurement conditions to count its probability interval (usually expressed by the maximum permissible error MPE). That is, the multiplicative constant error of a geodimeter also has its variance like the output deviation in the field of geodesy.

That is to say, both \( \Delta_A \) and \( \Delta_B \) are unknown deviations, have their own variances, and cannot be classified as random error and systematic error respectively. That is, the concepts of precision and trueness were abolished.

In this way, there are

\[
\Delta = \Delta_A + \Delta_B
\]  

(2-6)
\[ E(\Delta) = E(\Delta_A) + E(\Delta_B) \]
\[ = 0 \]
\[ \sigma^2(\Delta) = E[\Delta - E(\Delta)]^2 \]
\[ = E(\Delta^2) \]
\[ = E(\Delta_A + \Delta_B)^2 \]
\[ = \sigma^2(\Delta_A) + \sigma^2(\Delta_B) \]  

The variance \( \sigma^2(\Delta) \) is the evaluation value of the probability range of the total error \( \Delta \), is also the dispersion of all possible values of the total error \( \Delta \), and is called as the uncertainty of the error \( \Delta \). Please note that this uncertainty is not the dispersion of the measured value, which is different from the traditional measurement theory.

Therefore, for the true value \( X_T \), there are:
\[ X_T = x_0 - \Delta \]  
\[ E(X_T) = E(x_0) - E(\Delta) \]
\[ = x_0 \]
\[ \sigma^2(X_T) = E[X_T - E(X_T)]^2 \]
\[ = E(x_0 - \Delta - x_0)^2 \]
\[ = \sigma^2(\Delta) \]  

In conclusion, the probability expressions of measured value, error and true value are shown in Table 1\(^{[10][11][12]}\).

**Table 1. The probability expressions of measured value, error and true value**

| Measured value \( x_0 \) | Error \( \Delta \) | True value \( X_T \) |
|--------------------------|------------------|---------------------|
| \( \begin{cases} E(x_0) = x_0 \\ \sigma^2(x_0) = 0 \end{cases} \) | \( \begin{cases} E(\Delta) = 0 \\ \sigma^2(\Delta) = E(\Delta^2) \end{cases} \) | \( \begin{cases} E(X_T) = x_0 \\ \sigma^2(X_T) = \sigma^2(\Delta) \end{cases} \) |

\( E(\Delta) = 0 \) means that the mean value of all possible values of error \( \Delta \) after measurement is zero, because once \( E(\Delta) = C \neq 0 \), this \( C \) must be corrected to the measured value \( x_0 \).

The above conclusion is derived by using an observed value \( x_0 \) as the final measured value. When multiple observations \( x_i \) are used to obtain the best measured value, the analysis process is shown in Section 3.

### 2.4. Regularity and randomness of error

Any error has variance to evaluate its probability range, and the regular error is of course the same, because the variance of error is the dispersion of error values under all possible measurement conditions. Fig 2 is a conceptual diagram describing the regularity and randomness of the periodic error of the geodimeter. When observing the corresponding relationship between the error value and the range, we can see the regularity. When observing the density distribution of all possible values of the error, we can see the randomness.

![Fig 2. Regularity and randomness of sine error](image-url)
2.5. The contribution form of error to repeated measurement

Because errors vary with various measurement conditions, the variation of measurement conditions in repeated measurement determines the contribution form of errors to repeated observations (deviation, dispersion, outlier or both). It is precisely because errors can lead to the dispersion of repeated observations, so the variance of any error can be obtained through experimental statistics, and these data exist in the instrument specifications or instructions.

2.6. The law of covariance propagation

Because any error has variance, the concept of variance can be extended to any two errors, which is called covariance. It comes from the common component of the two errors.

Thus, for an error array $\Delta X = (\Delta x_1, \Delta x_2, \cdots, \Delta x_n)^T$, the definition of variance is as follows:

$$D(\Delta X) = E(\Delta XX^T)$$

This is the law of covariance propagation, which is the propagation relation of probability interval between errors, or the propagation relation of the dispersion of all possible values of errors, rather than the propagation relation of the dispersion of measured value explained by traditional theory.

After the data processing, the difference between the measured value and the true value is an unknown deviation. Therefore, the task of measurement is to adjust the measured value to minimize the variance of the deviation and submit its value.

3. The solution of weight values

The smaller the error is, the better the quality of the measured value is. However, because the error value is unknown, and we can only use the variance of the error to describe its probability interval, so we can only think that the quality of the measured value is the best when the variance of the error is the smallest. Therefore, the mathematical problem we are faced with is to minimize the variance of the error of the final measured value by assigning the weights of each observed value.

3.1. Direct measurement model for single measurand

As shown in Fig1, when $n$ different observed values $x_i$ are obtained by repeated observation of a measurand, assuming that the best measured value is $y$, the error equation is as follows:

$$\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \begin{bmatrix} y \end{bmatrix}$$

Its measurement model is $V = X - Y$, the equation set (3-1) is regarded as one of many equation sets, and each $v_i$, $x_i$ and $y$ is a sample in the sample spaces of random variables $V$, $X$ and $Y$. 

5
respectively. Therefore, according to the formula in Table 1, there are:

\[
E(V) = 0
\]

\[
\sigma^2(V) = E(V^2)
\]

\[
\approx \frac{\sum v_i^2}{n}
\]

\[
= \frac{1}{n} (x_1^2 + x_2^2 + \cdots + x_n^2 - 2x_1y - 2x_2y - \cdots - 2x_ny + ny^2)
\]

The best measured value is the value when \( \sigma^2(V) \) or \( \sum v_i^2 \) takes the minimum value, and this is a minimum problem of quadratic function, so making \( \frac{d(\sigma^2(V))}{dy} = 0 \), the best measured value is obtained as follows:

\[
y = \frac{1}{n} \sum_{i=1}^{n} x_i
\]

(3-3)

This is the least square method. As you can see, in the above deduction process of least square method, each \( x_i \) and \( y \) represents a numerical value respectively, instead of the random variable considered by traditional measurement theory, which is consistent with the conceptual principle in Section 2.

Of course, formula (3-3) is actually an equal weight processing method, that is, the weight of each observation \( x_i \) is \( \frac{1}{n} \). However, if \( n \) observation errors \( \Delta x_i \) have different variances and covariances, that is, the quality of each observation \( x_i \) is different from each other, then the equal weight processing is not scientific. At this time, different weights must be given to each observed value to ensure that the quality of the final measured value is the best. Therefore, in order to study each weight \( p_i \) when \( \sum p_i v_i^2 \) takes the minimum value, the equation (3-1) is modified as:

\[
\begin{bmatrix}
 p_1^{1/2} v_1 \\
p_2^{1/2} v_2 \\
\vdots \\
p_n^{1/2} v_n 
\end{bmatrix} = \begin{bmatrix}
 p_1^{1/2} x_1 \\
p_2^{1/2} x_2 \\
\vdots \\
p_n^{1/2} x_n 
\end{bmatrix} - \begin{bmatrix}
 p_1^{1/2} 0 & \cdots & 0 \\
 0 & p_2^{1/2} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & p_n^{1/2}
\end{bmatrix} \begin{bmatrix}
 1 \\
 1 \\
\vdots \\
1
\end{bmatrix}
\]

(3-4)

Similarly, by using the least square method, there is

\[
y = \left( p_1 x_1 + p_2 x_2 + \cdots + p_n x_n \right) / \sum_{i=1}^{n} p_i
\]

(3-5)

Please note that in the derivation process of the least square method, it is impossible to obtain the best weights \( p_i \) by making \( \frac{d(\sigma^2(V))}{dp_i} = 0 \). Actually, the best weight \( p_i \) originally depends on the evaluation of \( \Delta x_i = x_i - x_T \) rather than the value of \( v_i = x_i - y \).

For the convenience of calculation, making \( \sum_{i=1}^{n} p_i = 1 \), the formula (3-5) becomes:

\[
y = p_1 x_1 + p_2 x_2 + \cdots + p_n x_n
\]

(3-6)

Now, using the quality of each \( x_i \), we begin to solve the general calculation method of each weight \( p_i \).
Taking the total differential of equation (3-6), the error propagation equation is obtained as follows:

\[ \Delta y = p_1 \Delta x_1 + p_2 \Delta x_2 + \cdots + p_n \Delta x_n \]  

(3-7)

Applying the law of covariance propagation (2-14) to equation (3-7), there is:

\[ \sigma^2(\Delta y) = \begin{bmatrix} p_1 & p_2 & \cdots & p_n \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \\ \vdots \\ \Delta x_n \end{bmatrix} \]  

(3-8)

\[ = p_1^2 \sigma_1^2 + p_2^2 \sigma_2^2 + \cdots + p_n^2 \sigma_n^2 + 2p_1p_2\sigma_{12} + 2p_1p_3\sigma_{13} + \cdots \]

Please note that the variance \( \sigma_i^2 \) is the dispersion \( \sigma^2(\Delta x_i) \) of all possible values of the error \( \Delta x_i = x_i - x_i' \), which expresses the evaluation of the probability interval of the error \( \Delta x_i \). This concept of variance is consistent with probability theory, but it is different from the concept of dispersion \( \sigma^2(x_i) \) of observation \( x_i \) in traditional measurement theory.

When the value of the function \( \sigma^2(\Delta y) \) is the minimum, the quality of the measured value \( y \) is the best, which is a problem of finding the minimum value of the function \( \sigma^2(\Delta y) \) under the restriction condition \( \sum_{i=1}^{n} p_i = 1 \). Therefore, according to the Lagrange multiplier method, the following functions are formed:

\[ Y = \sigma^2(\Delta y) + 2\lambda(p_1 + p_2 + \cdots + p_n - 1) \]

\[ = p_1^2 \sigma_1^2 + p_2^2 \sigma_2^2 + \cdots + p_n^2 \sigma_n^2 + 2p_1p_2\sigma_{12} + 2p_1p_3\sigma_{13} + \cdots \]

(3-9)

\[ + 2\lambda(p_1 + p_2 + \cdots + p_n - 1) \]

By making the partial differential equation of each unknown \( p_i \) equal to 0, the Lagrange equation is obtained as follows:

\[ p_1 \sigma_1^2 + p_2 \sigma_{12} + p_3 \sigma_{13} + \cdots + p_n \sigma_{1n} + \lambda = 0 \]

\[ p_1 \sigma_{21} + p_2 \sigma_2^2 + p_3 \sigma_{23} + \cdots + p_n \sigma_{2n} + \lambda = 0 \]

\[ \vdots \]

\[ p_1 \sigma_{n1} + p_2 \sigma_{n2} + p_3 \sigma_{n3} + \cdots + p_n \sigma_n^2 + \lambda = 0 \]

\[ p_1 + p_2 + \cdots + p_n - 1 = 0 \]

(3-10)

That is:

\[ \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} & 1 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \]

(3-11)

Because the covariance array of error array \( \Delta X = [\Delta x_1, \Delta x_2, \cdots, \Delta x_n]^T \) is:
\[
\begin{align*}
\textbf{D}(\Delta \textbf{X}) &= E \begin{bmatrix}
\Delta x_1 \\
\Delta x_2 \\
\vdots \\
\Delta x_n
\end{bmatrix} \\
&= \begin{bmatrix}
\Delta x_1 \\
\Delta x_2 \\
\vdots \\
\Delta x_n
\end{bmatrix} \\
&= \begin{bmatrix}
\sigma_{11}^2 & \cdots & \sigma_{1n} \\
\sigma_{21} & \cdots & \sigma_{2n} \\
\vdots & \cdots & \vdots \\
\sigma_{n1} & \cdots & \sigma_{nn}^2
\end{bmatrix}
\end{align*}
\]

(3-12)

So the equation (3-11) can be written as follows:

\[
\begin{bmatrix}
1 \\
1 \\
\vdots \\
1 \\
1 \\
1 \\
0 \\
\lambda
\end{bmatrix}
\begin{bmatrix}
p_1 \\
p_2 \\
\vdots \\
p_n \\
\lambda
\end{bmatrix}
= \begin{bmatrix}
0 \\
0 \\
\vdots \\
0 \\
1
\end{bmatrix}
\]

(3-13)

Therefore

\[
\begin{bmatrix}
p_1 \\
p_2 \\
\vdots \\
p_n \\
\lambda
\end{bmatrix}
= \textbf{D}(\Delta \textbf{X})
\begin{bmatrix}
1 \\
1 \\
\vdots \\
1 \\
1
\end{bmatrix}
\]

(3-14)

And according to formula (3-13), there is

\[
\begin{align*}
\textbf{D}(\Delta \textbf{X}) &+ \begin{bmatrix}
1 \\
1 \\
\vdots \\
1 \\
1
\end{bmatrix} \lambda = \begin{bmatrix}
0 \\
0 \\
\vdots \\
0 \\
0
\end{bmatrix}
\end{align*}
\]

(3-15)

That is

\[
\begin{align*}
\textbf{D}(\Delta \textbf{X}) &+ \begin{bmatrix}
1 \\
1 \\
\vdots \\
1 \\
1
\end{bmatrix} \lambda = \begin{bmatrix}
-\lambda \\
-\lambda \\
\vdots \\
-\lambda
\end{bmatrix}
\end{align*}
\]

(3-16)

By substituting equation (3-16) into equation (3-8), we can get:

\[
\sigma^2(\Delta y) = \begin{bmatrix}
p_1 \\
p_2 \\
\vdots \\
p_n
\end{bmatrix} \textbf{D}(\Delta \textbf{X}) \\
\begin{bmatrix}
p_1 \\
p_2 \\
\vdots \\
p_n
\end{bmatrix}
\]

\[
= \begin{bmatrix}
p_1 \\
p_2 \\
\vdots \\
p_n
\end{bmatrix} \begin{bmatrix}
-\lambda \\
-\lambda \\
\vdots \\
-\lambda
\end{bmatrix}
\]

(3-17)

\[
= -\lambda
\]

8
Finally, by substituting formula (3-14) into formulas (3-6) and (3-17), the optimal measured value \( y \) and its error evaluation \( \sigma_y^2(\Delta y) \) can be obtained.

In formula (3-11), if each \( \sigma_y = 0 \), we can get the following results:

\[
p_1 : p_2 : \cdots : p_n = \frac{1}{\sigma_i^2} : \frac{1}{\sigma_j^2} : \cdots : \frac{1}{\sigma_n^2} \quad (3-18)
\]

This is the reason why we used to use a single \( \frac{1}{\sigma_i^2} \) to determine the weight ratio. Obviously, the formula (3-18) does not hold when the hypothesis \( \sigma_y = 0 \) does not exist.

Next, let's see the actual cases.

Example 1, as shown in Figure 1, when \( n \) observations \( x_i \) are obtained by using the same sensor to directly repeat the measurement, please calculate the final measured value according to formula (3-6).

Because \( \Delta x_i = \Delta_{B_i} + \Delta_{A_i} \), so according to the formula (2-8), there is:

\[
\sigma_i^2 = \sigma_{A_i}^2 + \sigma_{B_i}^2
\]

For any two observation errors \( \Delta x_i \) and \( \Delta x_j \), according to the formula (2-12), there is:

\[
\sigma_{jk} = E(\Delta x_i \Delta x_j) = E[(\Delta_{B_i} + \Delta_{A_i})(\Delta_{B_j} + \Delta_{A_j})] = E(\Delta_{B_i}^2) = \sigma_{B_i}^2
\]

Therefore, according to the formula (3-11), there is:

\[
\begin{bmatrix}
\sigma_{A_i}^2 + \sigma_{A_j}^2 & \sigma_{A_i}^2 & \cdots & \sigma_{A_j}^2 & 1 \\
\sigma_{A_i}^2 & \sigma_{A_i}^2 + \sigma_{A_j}^2 & \cdots & \sigma_{A_j}^2 & 1 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
\sigma_{A_i}^2 & \sigma_{A_i}^2 & \cdots & \sigma_{A_j}^2 + \sigma_{A_j}^2 & 1 \\
1 & 1 & \cdots & 1 & 0
\end{bmatrix}
\begin{bmatrix}
p_1 \\
p_2 \\
\vdots \\
p_n \\
\lambda
\end{bmatrix}
= 
\begin{bmatrix}
0 \\
0 \\
\vdots \\
0
\end{bmatrix}
\]

By solving the equation, we get the following results:

\[
p_1 = p_2 = \cdots = p_n = \frac{1}{n} \quad \lambda = -\sigma_{A_i}^2 \frac{\sigma_{A_j}^2}{n}
\]

Finally, according to (3-6) and (3-17), there are:

\[
y = \frac{x_1 + x_2 + \cdots + x_n}{n}, \quad \sigma^2(\Delta y) = \sigma_{A_i}^2 + \sigma_{B_i}^2 \frac{\sigma_{A_j}^2}{n}
\]

It can be seen that in this case, the root cause of \( p_1 = p_2 = \cdots = p_n = \frac{1}{n} \) is that each error \( \Delta x_i \) has the same variance \( \sigma_i^2 = \sigma_{A_i}^2 + \sigma_{B_i}^2 \) and any two errors \( \Delta x_i \) and \( \Delta x_j \) have the same covariance \( \sigma_{jk} = \sigma_{A_i}^2 \). Moreover, it can be seen that the traditional measurement theory classifies
\( \Delta_B \) as systematic error, which is actually to negate the correlation between \( \Delta x_i \) and \( \Delta x_j \).

Example 2, as shown in Fig 3, in order to obtain the accurate diameter value of the copper wire, the observed values of 5 times, 10 times, 15 times and 20 times of the diameter are measured with a caliper, and 4 observed values are shown in Table 2. Please give the best measured value of diameter and the evaluation of uncertainty.

| Table 2. Observed values |
|--------------------------|
| \( i \) | 1 | 2 | 3 | 4 |
| \( a_i \) | 5 | 10 | 15 | 20 |
| \( l_i \) (mm) | 2.87 | 5.72 | 8.60 | 11.37 |

The four measured values are as follows:
\[ x_1 = l_1/a_1 = 0.574 \]
\[ x_2 = l_2/a_2 = 0.572 \]
\[ x_3 = l_3/a_3 = 0.5733 \]
\[ x_4 = l_4/a_4 = 0.5685 \]

The error \( \Delta l_i \) of the observed value \( l_i \) consists of three parts: fixed error \( K \), proportional error \( R \) and non-uniform dividing error \( c_i \). That is:
\[ \Delta l_i = K + l_i R + c_i \]

According to the definition (2-13) of variance, there is
\[
\text{D(AX)} = \text{AD(}\Delta l\text{A)}\text{I}
\]
\[
= \left[ \begin{array}{cccc}
\sigma_k^2 & l_1^2 \sigma_R^2 + \sigma_c^2 & \sigma_k^2 & \sigma_k^2 \\
\sigma_k^2 & l_2^2 \sigma_R^2 + \sigma_c^2 & \sigma_k^2 & \sigma_k^2 \\
\sigma_k^2 & l_3^2 \sigma_R^2 + \sigma_c^2 & \sigma_k^2 & \sigma_k^2 \\
\sigma_k^2 & l_4^2 \sigma_R^2 + \sigma_c^2 & \sigma_k^2 & \sigma_k^2
\end{array} \right]
\]

Because \( x_i = l_i/a_i \) and \( \Delta x_i = \Delta l_i/a_i \), there is
\[
\left[ \begin{array}{c}
\Delta x_1 \\
\Delta x_2 \\
\Delta x_3 \\
\Delta x_4
\end{array} \right] = \left[ \begin{array}{cccc}
a_{i1}^{-1} & 0 & 0 & 0 \\
0 & a_{i2}^{-1} & 0 & 0 \\
0 & 0 & a_{i3}^{-1} & 0 \\
0 & 0 & 0 & a_{i4}^{-1}
\end{array} \right] \left[ \begin{array}{c}
\Delta l_1 \\
\Delta l_2 \\
\Delta l_3 \\
\Delta l_4
\end{array} \right]
\]

That is
\[ \Delta \text{X} = \Delta \text{AL} \]

According to the law of covariance propagation (2-14), there is:
\[
\text{D(AX)} = \text{AD(}\Delta l\text{A)}\text{I}
\]

Assuming \( \sigma_K = \pm 0.01 \text{mm} \), \( \sigma_R = \pm 1 \times 10^{-5} \) and \( \sigma_c = \pm 0.02 \text{mm} \), there is:

10
According to the formula (3-14), there is:
\[
\begin{bmatrix}
    p_1 \\
p_2 \\
p_3 \\
p_4 \\
\end{bmatrix} = \begin{bmatrix}
    -0.014285714 \\
0.085714286 \\
0.299999996 \\
0.628571432 \\
\end{bmatrix}
\]

According to the formula (3-6), there is:
\[
y = \begin{bmatrix}
p_1 & p_2 & p_3 & p_4 \\
x_1 \\
x_2 \\
x_3 \\
x_4 \\
\end{bmatrix} = 0.5702 \text{(mm)}
\]

According to the formula (3-17), there is:
\[
\sigma^2(\Delta y) = -\lambda \\
= 9.14286 \times 10^{-7} \text{(mm}^2)\]

Therefore, the uncertainty is:
\[
\sigma(\Delta y) \approx \pm 0.00096 \text{(mm)}
\]

As you can see, \( p_1 \) is a negative value, which is beyond our inherent concept. In the past, we used to follow the conceptual logic of traditional theory to determine weights according to precision, ignoring the correlation between errors, so that \( \mathbf{D}(\Delta \mathbf{x}) \) is a diagonal matrix, and \( p_i \) is always positive. Now, variance is the dispersion of all possible values of error, and any error has variance and covariance, so \( \mathbf{D}(\Delta \mathbf{x}) \) is no longer a diagonal matrix, and \( p_i \) can be negative.

However, this negative weight is also a positive contribution to the final measured value. In this case, \( a_4 \) is the smallest, and the fixed error \( K \) is too harmful to the observed value \( x_4 \). Therefore, Lagrange algorithm can only choose negative weights to correct this harm, which is exactly the scientific point of this algorithm.

### 3.2. Indirect measurement model for single measurand

The observation error equation is as follows.
\[
\begin{bmatrix}
p_1^{1/2} v_1 \\
p_2^{1/2} v_2 \\
\vdots \\
p_n^{1/2} v_n \\
\end{bmatrix} = \begin{bmatrix}
p_1^{1/2} x_1 \\
p_2^{1/2} x_2 \\
\vdots \\
p_n^{1/2} x_n \\
\end{bmatrix} - \begin{bmatrix}
p_1^{1/2} 0 & \cdots & 0 \\
0 & p_2^{1/2} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & p_n^{1/2} \\
\end{bmatrix} \begin{bmatrix}
a_1 \\
a_2 \\
\vdots \\
a_n \\
\end{bmatrix} \quad (3-19)
\]

According to the least square method, there is:
\[
y = \sum \frac{p_i a_i x_i}{\sum p_i a_i^2} \quad (3-20)
\]

Similar to the above direct measurement model, in order to realize the minimum value of \( \sigma^2(\Delta y) \), it is assumed that \( \sum p_i a_i^2 = 1 \), and the formula (3-20) becomes:
\[
y = p_1 a_1 x_1 + p_2 a_2 x_2 + \cdots + p_n a_n x_n
\]
\[
= \begin{bmatrix}
    a_1 & a_2 & \cdots & a_n
\end{bmatrix}
\begin{bmatrix}
    p_1 & 0 & \cdots & 0 \\
    0 & p_2 & \cdots & 0 \\
    \vdots & \vdots & \ddots & \vdots \\
    0 & 0 & \cdots & p_n
\end{bmatrix}
\begin{bmatrix}
    x_1 \\
    x_2 \\
    \vdots \\
    x_n
\end{bmatrix}
\] (3-21)

That is
\[
y = A^T P X
\] (3-22)

Among them, \[ A = \begin{bmatrix}
    a_1 \\
    a_2 \\
    \vdots \\
    a_n
\end{bmatrix}, \quad P = \begin{bmatrix}
    p_1 & 0 & \cdots & 0 \\
    0 & p_2 & \cdots & 0 \\
    \vdots & \vdots & \ddots & \vdots \\
    0 & 0 & \cdots & p_n
\end{bmatrix}, \quad X = \begin{bmatrix}
    x_1 \\
    x_2 \\
    \vdots \\
    x_n
\end{bmatrix}. \]

Taking the total differential of formula (3-22), the error propagation equation is obtained as follows:
\[
\Delta y = A^T P \Delta X
\] (3-23)

According to the law of covariance propagation (2-14), there is:
\[
\sigma^2(\Delta y) = A^T PD(\Delta X)PA
\]
\[
= p_1^2 a_1^2 \sigma_1^2 + p_2^2 a_2^2 \sigma_2^2 + \cdots + p_n^2 a_n^2 \sigma_n^2 + 2 p_1 p_2 a_1 a_2 \sigma_1 \sigma_2 + 2 p_1 p_3 a_1 a_3 \sigma_1 \sigma_3 + \cdots
\] (3-24)

Similarly, according to Lagrange multiplier method and omitting the derivation process, it is concluded that:
\[
\begin{bmatrix}
    a_1^2 \sigma_1^2 & a_1 a_2 \sigma_{12} & \cdots & a_1 a_n \sigma_{1n} & a_1^2 \\
    a_2 a_1 \sigma_{21} & a_2^2 \sigma_2^2 & \cdots & a_2 a_n \sigma_{2n} & a_2^2 \\
    \vdots & \vdots & \ddots & \vdots & \vdots \\
    a_n a_1 \sigma_{n1} & a_n a_2 \sigma_{n2} & \cdots & a_n^2 \sigma_n^2 & a_n^2
\end{bmatrix}
\begin{bmatrix}
    p_1 \\
    p_2 \\
    \vdots \\
    p_n
\end{bmatrix}
= \begin{bmatrix}
    0 \\
    0 \\
    \vdots \\
    0
\end{bmatrix}
\] (3-25)

Making \[ A' = \begin{bmatrix}
    a_1 & 0 & \cdots & 0 \\
    0 & a_2 & \cdots & 0 \\
    \vdots & \vdots & \ddots & \vdots \\
    0 & 0 & \cdots & a_n
\end{bmatrix} \], the formula (3-25) becomes:
\[
\begin{bmatrix}
    p_1 \\
    p_2 \\
    \vdots \\
    p_n
\end{bmatrix}
= A'D(\Delta X)A'
\] (3-26)

It can be seen that the best weights are actually determined by \( D(\Delta X) \) and \( A' \) together.

And because according to (3-26), there is:
\[ A'D(\Delta X)A' \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix} = \begin{bmatrix} a_1^2 \\ a_2^2 \\ \vdots \\ a_n^2 \end{bmatrix} [-\lambda] \]  
(3-27)

So there is

\[ \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix}A'D(\Delta X)A' \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix} = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix}A'D(\Delta X)A' \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix} [-\lambda] \]  
(3-28)

And because there is:

\[ A^TPD(\Delta X)PA = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix}A'D(\Delta X)A' \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix} \]  
(3-29)

So formula (3-24) becomes:

\[ \sigma^2(\Delta y) = A^TPD(\Delta X)PA = -\lambda \]  
(3-30)

Example 3, the data of example 2 is still used, and the solution is carried out according to the (3-19) model.

The observation error equation is as follows.

\[
\begin{bmatrix} l_1^{1/2}v_1 \\ l_2^{1/2}v_2 \\ l_3^{1/2}v_3 \\ l_4^{1/2}v_4 \end{bmatrix} = \begin{bmatrix} l_1^{1/2}l_1 \\ l_2^{1/2}l_2 \\ l_3^{1/2}l_3 \\ l_4^{1/2}l_4 \end{bmatrix} - \begin{bmatrix} p_1^{1/2} \\ p_2^{1/2} \\ p_3^{1/2} \\ p_4^{1/2} \end{bmatrix} a_1 y = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix}
\]

There are \( L = \begin{bmatrix} l_1 \\ l_2 \\ l_3 \\ l_4 \end{bmatrix} = \begin{bmatrix} 2.87 \\ 5.72 \\ 8.6 \\ 11.37 \end{bmatrix}, A = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = \begin{bmatrix} 5 \\ 10 \\ 15 \\ 20 \end{bmatrix} \) and \( A' = \begin{bmatrix} 5 \\ 10 \\ 15 \\ 20 \end{bmatrix} \).

Because \( \Delta l_i = K + l_iR + c_i \),

According to the definition (2-13) of variance, there is

\[
D(\Delta L) = \begin{bmatrix} \Delta l_1 \\ \Delta l_2 \\ \Delta l_3 \\ \Delta l_4 \end{bmatrix}
\]
\[
\begin{pmatrix}
\sigma_k^2 + l_1^2 \sigma_n^2 + \sigma_i^2 & \sigma_k^2 + l_1^2 \sigma_n^2 & \sigma_k^2 + l_1^2 \sigma_n^2 \\
\sigma_k^2 + l_1^2 \sigma_n^2 & \sigma_k^2 + l_1^2 \sigma_n^2 + \sigma_i^2 & \sigma_k^2 + l_1^2 \sigma_n^2 \\
\sigma_k^2 + l_1^2 \sigma_n^2 & \sigma_k^2 + l_1^2 \sigma_n^2 + \sigma_i^2 & \sigma_k^2 + l_1^2 \sigma_n^2 + \sigma_i^2 \\
\end{pmatrix}
\]

Substituting \(\sigma_k = \pm 0.01 \text{mm}, \ \sigma_n = \pm 1 \times 10^{-5}\) and \(\sigma_i = \pm 0.02 \text{mm},\) and according to the formula (3-26), there is:

\[
\begin{bmatrix}
p_1 \\ p_2 \\ p_3 \\ \vdots \\ p_n \\
\end{bmatrix} = \begin{bmatrix} a_1^2 & a_2^2 & a_3^2 & a_4^2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\
\end{bmatrix}^{-1} \begin{bmatrix} 0 \\ -0.000571429 \\ 0.000857143 \\ 0.00133333 \\ 0.001571429 \\ \end{bmatrix}
\]

According to formula (3-22), the best measured value is as follows:

\[
y = A^T PX = 0.5702(\text{mm})
\]

According to the formula (3-30), there is:

\[
\sigma^2(\Delta y) = -\lambda = 9.14286 \times 10^{-7}(\text{mm}^2)
\]

Therefore, the uncertainty is \(\sigma(\Delta y) = \pm 0.00096(\text{mm})\)

It can be seen that the final results of example 3 and example 2 are the same.

3.3. **Indirect measurement model for multiple measurands**

The observation error equation is as follows.

\[
\begin{bmatrix}
p_{1/2}^1 v_1 \\ p_{1/2}^2 v_2 \\ \vdots \\ p_{1/2}^n v_n \\
\end{bmatrix} = \begin{bmatrix}
p_{1/2}^1 x_1 \\ p_{1/2}^2 x_2 \\ \vdots \\ p_{1/2}^n x_n \\
\end{bmatrix} - \begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1t} \\ a_{21} & a_{22} & \cdots & a_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ a_{nt} & a_{nt} & \cdots & a_{nt} \\
\end{bmatrix} \begin{bmatrix}
y_1 \\ y_2 \\ \vdots \\ y_t \\
\end{bmatrix} \tag{3-31}
\]

According to the least square method, there is:

\[
Y = [A^T PA]^{-1} A^T PX \tag{3-32}
\]

Taking the total differential of formula (3-32), the error propagation equation is obtained as follows:

\[
\Delta Y = [A^T PA]^{-1} A^T P \Delta X \tag{3-33}
\]

However, in the error propagation equation (3-33), the propagation law of each error \(\Delta y_j\) is different from each other, so it is difficult to ensure that the variance of all errors can be minimized at the same time. Therefore, it is suggested to use formula (3-14) to obtain a relative fair weight ratio.

4. **The main mistakes of traditional measurement theory in weight calculation method**
The traditional theory explains its measurement concept according to Fig 4.

1. Any observed value \( x_i \) (or measured value) is a numerical value and belongs to a constant, and there must be \( \mathbb{E}(x_i) = x_i \) and \( \sigma^2(x_i) = 0 \). The traditional theory regards \( x_i \) as a random variable, which violates the basic concept of probability theory. For example, in case 2, according to the pure concept of probability theory, there must be:

\[
\begin{align*}
\mathbb{E}(x_i) &= \mathbb{E}(0.574) = 0.574 & \sigma^2(x_i) &= \sigma^2(0.574) = 0 \\
\mathbb{E}(x_j) &= \mathbb{E}(0.572) = 0.572 & \sigma^2(x_j) &= \sigma^2(0.572) = 0 \\
\mathbb{E}(x_k) &= \mathbb{E}(0.5733) = 0.5733 & \sigma^2(x_k) &= \sigma^2(0.5733) = 0 \\
\mathbb{E}(x_l) &= \mathbb{E}(0.5685) = 0.5685 & \sigma^2(x_l) &= \sigma^2(0.5685) = 0 
\end{align*}
\]

It can be seen that using \( \sigma^2(x_i) \) as the basis to determine the weight \( p_i \) is a kind of conceptual logic confusion.

2. Because \( x_i \) is a constant, the error \( x_i - \mathbb{E}(X) \) is a deviation, which is the same as the deviation \( \mathbb{E}(x) - x_T \). Therefore, the theory of error classification is not correct.

3. The precision \( \sigma^2(x_i) \) in the traditional theory is actually the evaluation of the probability range of the deviation \( x_i - \mathbb{E}(x) \), rather than the dispersion of the observed value \( x_i \). Therefore, the variance of deviation \( \mathbb{E}(x) - x_T \) is ignored when the weight is determined only by precision.

5. Conclusion

According to the pure concept of probability theory, both the observed value and the measured value are numerical values and belong to constant, while the error and true value are unknown and belong to random variable; the variance of error is the dispersion of all possible values of the error, and is the evaluation value of the probability interval where the error exists; any error is a deviation and has its variance, so it can't be classified as systematic error or random error. Based on these concepts, and analysing the covariance propagation relationship between the observation error and the final error, it is a pure mathematical problem to give the weight assignment with the minimum
variance, and the dilemma of traditional measurement theory has been solved.

In an adjustment model, the weight allocation of observations is not only determined by the covariance array of observation error, but also related to the parameters in the adjustment model, which is particularly important for a single measurand model. For the adjustment model with multiple measurands, because the errors of multiple measured values come from different propagation laws, the variances of all errors cannot be minimized at the same time, and only the covariance array of observation errors can be used as the only basis to determine the weight.

References
1 JCGM 200:2012, International vocabulary of metrology — Basic and general concepts and associated terms (VIM).
2 JCGM 100:2008, Guide to the Expression of Uncertainty in Measurement (GUM).
3 JIF1001-2011, General Terms in Metrology and Their Definitions.
4 JIF1059-2012, Evaluation and Expression of Uncertainty in Measurement.
5 GB/T14911-2008, Basic Terms of Surveying and Mapping.
6 School of Geodesy and Geomatics, Wuhan University. Error Theory and Foundation of Surveying Adjustment. Wuhan University Press, 2016 6.
7 Xiao-ming Ye, Mo Ling, Qiang Zhou, Wei-nong Wang, Xue-bin Xiao. “The New Philosophical View about Measurement Error Theory”. Acta Metrologica Sinica, volume 36, 2015, 6, Pages: 666-670.
8 Ye Xiao-ming, Xiao Xue-bin, Shi Jun-bo, Ling Mo. The new concepts of measurement error theory [J]. Measurement, Volume 83, April 2016, Pages: 96-105.
9 Ye Xiao-ming, Liu Hai-bo, Ling Mo, Xiao Xue-bin. The new concepts of measurement error's regularities and effect characteristics [J]. Measurement, Volume 126, October 2018, Pages: 65-71.
10 Ye Xiao-ming, Ding Shi-jun. Comparison of variance concepts interpreted by two measurement theories [J]. Journal of Nonlinear and Convex Analysis, Volume 20, Number 7, 2019, Pages: 1307-1316.
11 Shi Huisheng, Ye Xiaoming, Xing Cheng, Ding Shijun. Origin and Evolution of Conceptual Differences between Two Measurement Theories [J]. Fuzzy Systems and Data MiningVI, 2020.11.
12 Shi Huisheng, Ye Xiaoming, Xing Cheng, Ding Shijun. A New Theoretical Interpretation of Measurement Error and Its Uncertainty [J]. Discrete Dynamics in Nature and Society, 2020. DOI: https://doi.org/10.1155/2020/3864578.
13 Ye Xiaoming. Measurement Error Theory Based on New Concepts. Hubei Science and Technology Press, 2017.