From Equations to Tri-quations and Multi-quations

Mourat Tchoshanov, Olga Kosheleva, and Vladik Kreinovich

University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA

Abstract

In general, an equation $A(x_1, \ldots, x_n) = B(x_1, \ldots, x_n)$ corresponds to the situation when we have two quantities $A(x_1, \ldots, x_n)$ and $B(x_1, \ldots, x_n)$ which are known to be equal, we know how each of these quantities depends on the unknown parameters $x_1, \ldots, x_n$, and we want to find the values of the unknowns $x_i$ from this equality – and from other similar equalities. In some practical situations, instead of two equal values, we have three (or more) quantities which are all equal to each other. In this paper, we explain how to find the unknowns $x_i$ in such situations, i.e., how to solve the corresponding “tri-quations” and “multi-quations”.

Mathematics Subject Classification: 93E24

Keywords: equations, data processing, Least Squares, tri-quations, multi-quations
1 From Equations to Tri-equations and Multi-equations: Idea

What is an equation. By definition, an equation is an expression of the type
\[ A(x_1, \ldots, x_n) = B(x_1, \ldots, x_n), \] (1)
where both the left-hand side \( A(x_1, \ldots, x_n) \) and the right-hand side \( B(x_1, \ldots, x_n) \) may depend on the unknowns \( x_1, \ldots, x_n \).

Simple equations. The simplest equations are equations of the type
\[ A(x_1, \ldots, x_n) = b \] (2)
for some constant \( b \) – and the simplest case of such a constant is \( b = 0 \).

Practical usefulness of simple equations. Simple equations, of the type (2), are typical in data processing, when we need to find the parameters of the system based on observations.

For example, we observe the motion of several planets and their satellites, and, based on this information, we want to estimate the weights of different planets.

In general, we know the dependence \( A_i(x_1, \ldots, x_n) \) the quantity \( A_i \) that we measure in the \( i \)-th measurement depends on the desired parameters \( x_1, \ldots, x_n \), we know the results \( b_i \) of the corresponding measurements. We can then use the corresponding equations
\[ A_i(x_1, \ldots, x_n) = b_i, \quad 1 \leq i \leq N, \] (3)
to find the desired values \( x_1, \ldots, x_n \).

Taking uncertainty into account. In practice, the measurements are never exact, there is always a measurement error \( \eta_i \neq 0 \); see, e.g., [2]. So, instead of the exact equations, we have approximate equations
\[ A_i(x_1, \ldots, x_n) = b_i + \eta_i, \quad 1 \leq i \leq N. \] (4)

In many cases, the measurement error is caused by the joint effect of many independent causes. In this case, the Central Limit Theorem implies that the measurement error is normally distributed [3]. It is usually assumed that the bias has been adjusted for, so the mean of this distribution is 0. Normal distribution with 0 mean has a probability density
\[ \rho(\eta_i) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot \exp \left( -\frac{\eta_i^2}{2\sigma^2} \right), \] (5)
for some value \( \sigma \).
Measurement errors corresponding to different measurements are usually independent. As a result, the probability density corresponding to the tuples \( \eta = (\eta_1, \ldots, \eta_N) \) is equal to the product of the probability densities corresponding to individual measurements:

\[
\rho(\eta) = \left( \frac{1}{\sqrt{2\pi \cdot \sigma}} \right)^N \prod_{i=1}^{N} \exp \left( -\frac{\eta_i^2}{2\sigma^2} \right).
\]

(6)

Since \( \exp(a) \cdot \exp(b) = \exp(a + b) \), this expression can be transformed into the following equivalent form

\[
\rho(\eta) = \left( \frac{1}{\sqrt{2\pi \cdot \sigma}} \right)^N \exp \left( -\frac{\sum_{i=1}^{N} \eta_i^2}{2\sigma^2} \right).
\]

(7)

It makes sense to select the parameters \( x_1, \ldots, x_n \) which are the most probable. Since \( \exp(x) \) is a strictly increasing function, maximizing the probability (7) is equivalent to minimizing the sum \( \sum_{i=1}^{N} \eta_i^2 \). Since \( \eta_i = b_i - A_i(x_1, \ldots, x_n) \), this means that we find the values \( x_1, \ldots, x_n \) from the condition

\[
\sum_{i=1}^{N} (b_i - A_i(x_1, \ldots, x_n))^2 \rightarrow \min_{x_1, \ldots, x_n}.
\]

(8)

This condition is known as the Least Squares approach [3].

**From simple equations to general equations.** In some practical situations, instead of (or in addition to) simple equations, we also have general (non-simple) equations. For example, we can use the fact that the energy is preserved, so if we have expressions for energy \( A_i(x_1, \ldots, x_n) \) and \( B_i(x_1, \ldots, x_n) \) at two different moments of time, these expressions should lead to the same value, i.e., we should have

\[
A_i(x_1, \ldots, x_n) = B_i(x_1, \ldots, x_n).
\]

(9)

A general equation can be reduced to a simple equation. In situations when the expressions \( A_i(x_1, \ldots, x_n) \) and \( B_i(x_1, \ldots, x_n) \) are exact, the equation (9) can be naturally transformed into an equivalent equation

\[
A_i(x_1, \ldots, x_n) - B_i(x_1, \ldots, x_n) = 0
\]

of type (2).

Let us show that a similar equivalent transformation is possible when we take uncertainty into account. Indeed, if \( A_i(x_1, \ldots, x_n) \) and \( B_i(x_1, \ldots, x_n) \)
are approximate expressions, this means that $A_i(x_1, \ldots, x_n) \approx E_i$ and $B_i(x_1, \ldots, x_n) \approx E_i$ for some unknown common value $E_i$, i.e., that we have

$$A_i(x_1, \ldots, x_n) = E_i + \eta_{ia} \quad (11a)$$

and

$$B_i(x_1, \ldots, x_n) = E_i + \eta_{ib} \quad (11b)$$

for some approximation errors $\eta_{ia}$ and $\eta_{ib}$. For these equations, the least squares approach leads to the formula

$$(A_i(x_1, \ldots, x_n) - E_i)^2 + (B_i(x_1, \ldots, x_n) - E_i)^2 \to \min_{x_1, \ldots, x_n, E_i}. \quad (12)$$

We can easily minimize this expression over $E_i$: namely, if we differentiate this expression over $E_i$ and equate the derivative to 0, we can conclude that

$$E_i = \frac{A_i(x_1, \ldots, x_n) + B_i(x_1, \ldots, x_n)}{2}. \quad (13)$$

Substituting this expression for $E_i$ into the left-hand side of the optimizing expression (12), we get the formula

$$\left( A_i(x_1, \ldots, x_n) - \frac{A_i(x_1, \ldots, x_n) + B_i(x_1, \ldots, x_n)}{2} \right)^2 + \left( A_i(x_1, \ldots, x_n) - \frac{A_i(x_1, \ldots, x_n) + B_i(x_1, \ldots, x_n)}{2} \right)^2 = \frac{1}{2} \cdot (A_i(x_1, \ldots, x_n) - B_i(x_1, \ldots, x_n))^2. \quad (14)$$

Minimizing this expression is equivalent to minimizing the expression

$$(A_i(x_1, \ldots, x_n) - B_i(x_1, \ldots, x_n))^2, \quad (15)$$

i.e., to the application of the Least Squares approach to the approximate equation (10).

**From equations to tri-quations and multi-quations.** What if we equate the values of energy at three or more moments of time? In this case, instead of two equal expressions $A_i(x_1, \ldots, x_n)$ and $B_i(x_1, \ldots, x_n)$, we have three (or more) equal expressions $A_{i,k}(x_1, \ldots, x_n), 1 \leq k \leq K$.

The situation when we have three equal quantities $K = 3$ is known as a tri-quation (see, e.g., [1]). By analogy, it is reasonable to call a general situation $K > 2$ a multi-quation.
Case of exact tri-quotations and multi-quotations. When all the expressions $A_{i,k}(x_1, \ldots, x_n)$ are exact, a system of general equations $A_{i,k}(x_1, \ldots, x_n) = A_{il}(x_1, \ldots, x_n)$ is equivalent to a system of simple equations

$$A_{i,k+1}(x_1, \ldots, x_n) - A_{i,k}(x_1, \ldots, x_n) = 0, \ 1 \leq k \leq K - 1.$$  \hfill (16)

**Remaining question.** What happens if we take uncertainty into account? If we simply apply the Least Squares approach to the system (16), we get the wrong answer, for two reasons.

First, the original multi-quotation is invariant under all possible permutations of the $K$ expressions, but the Least Squares formula for (16), i.e., the expression

$$\sum_{k=1}^{K-1} (A_{i,k+1}(x_1, \ldots, x_n) - A_{i,k}(x_1, \ldots, x_n))^2$$  \hfill (17)

is *not* invariant.

Second, the Least Squares formula (17) is only applicable when the approximation errors corresponding to equations (16) are independent. But if we consider the equations

$$A_{i,k}(x_1, \ldots, x_n) = E_i + \eta_{i,k}, \ k = 1, \ldots, K,$$  \hfill (18)

then we get

$$A_{i,k+1}(x_1, \ldots, x_n) - A_{i,k}(x_1, \ldots, x_n) = \eta_{i,k+1} - \eta_{i,k},$$  \hfill (18)

so for the approximation errors of two consequent equations we get

$$E[(\eta_{i,k+1} - \eta_{i,k}) \cdot (\eta_{i,k} - \eta_{i,k-1})] =$$

$$E[\eta_{i,k+1} \cdot \eta_{i,k}] - E[\eta_{i,k}^2] - E[\eta_{i,k+1} \cdot \eta_{i,k-1}] + E[\eta_{i,k} \cdot \eta_{i,k-1}].$$  \hfill (19)

Since approximation errors corresponding to different $k$ are independent, all the terms except one are equal to 0, so we get

$$E[(\eta_{i,k+1} - \eta_{i,k}) \cdot (\eta_{i,k} - \eta_{i,k-1})] = -\sigma^2,$$  \hfill (20)

meaning that the approximation errors $\eta_{i,k+1} - \eta_{i,k}$ and $\eta_{i,k} - \eta_{i,k-1}$ are *not* independent.

Since this natural reduction to a system of simple equations does not work if we take uncertainty into account, what should we do?
2 From Equations to Tri-quations and Multi-quations: Taking Uncertainty into Account

How to take uncertainty into account when solving tri-quations and multi-quations: analysis of the problem. For the system (18), the least squares approach leads to the formula

\[ \sum_{k=1}^{K} (A_{i,k}(x_1, \ldots, x_n) - E_i)^2 \rightarrow \max_{x_1, \ldots, x_n, E_i}. \]  

(21)

We can easily minimize this expression over \( E_i \): namely, if we differentiate this expression over \( E_i \) and equate the derivative to 0, we can conclude that

\[ E_i = \frac{1}{K} \cdot \sum_{k=1}^{K} A_{i,k}(x_1, \ldots, x_n). \]  

(22)

Substituting this expression into the left-hand side of the optimizing expression (21), we get the formula

\[ \sum_{k=1}^{K} (A_{i,k}(x_1, \ldots, x_n) - E_i)^2, \]  

(23)

which is nothing else but the sample variance of the values \( A_{i,1}(x_1, \ldots, x_n), \ldots, A_{i,K}(x_1, \ldots, x_n) \). Thus, we arrive at the following conclusion.

How to take uncertainty into account when solving tri-quations and multi-quations: recommendation. If instead of two approximately equal values, we have \( K > 2 \) approximately equal values \( A_{i,1}(x_1, \ldots, x_n), \ldots, A_{i,K}(x_1, \ldots, x_n) \), then, to find the most probable values of the parameters \( x_1, \ldots, x_n \), we should minimize the sample variance (22), (23) of the given \( K \) values.

Acknowledgments

This work was supported in part by the US National Science Foundation grants HRD-0734825, HRD-1242122, and DUE-0926721.

References

http://dx.doi.org/10.1007/0-387-29143-1


Received: November 14, 2015; Published: January 2, 2016