Method of indirect measurements on the basis of vector-observation

S A Popov

Yaroslav-the-Wise Novgorod State University, ul. B. St. Peterburgskaya, 41, 173004, Veliky Novgorod, Russia
E-mail: Stanislav.Popov@novsu.ru

Abstract. The method of indirect measurements is considered, which allows to increase the accuracy of estimation of the measured sought value by building a multiresponse model of measurements by means of vector-observation. The expressions for calculation of the model coefficient estimates and covariance matrices of measurement errors are given. It is shown, how by means of indirect measurements and the multiresponse model the refined estimation of the sought value and its confidence interval are calculated.

1. Indirect measurements using vector-observation

Indirect measurements refer to such measurements, in which the numerical value of the sought measured value is determined by the known functional dependence through other values that can be directly measured.

\[ x = \sum_{i} a_i y_i, \]

where \( x \) is the sought measured value, \( y_i \) is the value of the indirect measurement, \( a_i \) the correcting coefficient, \( m \) is the number of indirect measurements of the sought value \( x \).

The order of calculations of the sought value by the measured values in the indirect measurements is set by the Recommendations [1] which set the guidelines for the estimation of the indirect measured sought value and the errors of the measurement result at linear dependence and lack of correlation between errors of measurements of arguments; in case of non-linear dependence and lack of correlation between measurements errors of arguments; for correlated error measurements of arguments when there are rows of individual values of the measured arguments.

To improve the accuracy of the measured sought value in this work it is proposed to use indirect measurements, in which to assess the desired value nonlinear dependence of the sought value and its indirect measurements (calibrating function) is used in the form of vector-observation and covariance matrix of the errors of observations of indirect measurements.

2. Calculation of non-linear calibration function coefficient estimates

Indirect multiple (group) measurements can be used to increase the accuracy of the sought value by increasing the number of measured indirect values and increasing the accuracy of their estimations.

For example, a group of sensors can be used to increase accuracy by using an indirect method to measure the same sought value [2]. This article offers the method of group indirect measurement consisting of two stages.
At the stage of the function construction, the calibration function for each sensor is constructed based on the model of dependence of the sought value on sensors responses, adequately connecting the values of indirect measurements with the measured sought value.

In the second stage by means of the calibration function according to the results of measurement of sensors responses, the sought value and its confidence interval are calculated. Here it is shown how on the basis of indirect measurements the estimate of the sought value and its variance are calculated by means of indirect measurements.

The first stage defines the type of multiresponse calibration function and evaluates its coefficients; the multiresponse model of the dependence of the sought value on the sensor responses can be represented as a multiresponse function, known accurate within number coefficients:

$$ Y = F(B, x) + E $$

(1)

where \( Y = (y_1, \ldots, y_m)^T \) is a vector-observation of the responses of each sensor \( y_i \) at the given sought value \( x \), \( B = (b_1, \ldots, b_l)^T \) is a vector of the coefficients of the model, \( E = (e_1, \ldots, e_m)^T \) is the a vector of additive errors of observation of sensor readings, \( F(B, x) = (f_1(B, x), \ldots, f_m(B, x))^T \) is a vector of functions simulating the dependence of sensor responses from the sought value for each sensor.

The vector of estimates of the model coefficients (1) in the general view is calculated at the stage of construction of the calibration function on the results of the experiment in the form of \( n \) indirect measurements of the sought value \( x \) by means of \( m \) sensors on the basis of the following iteration procedure [2]:

$$ B_{k+1} = B_k + \left[ \sum_{j=1}^{n} P(B', x_j)P(B', x_j)^T \right]^{-1} \sum_{j=1}^{n} P(B', x_j) \left[ Y_j - F(B', x_j) \right], $$

(2)

where \( x \) is the magnitude of the sought value, \( n \) is the number of sought values in the experiment,

The matrix of derivative functions \( F(B, x) \) by coefficients is defined as

$$ P(B, x) = \frac{\partial F(B, x)}{\partial B} = \left\{ \frac{\partial f_1(B, x)}{\partial B}, \ldots, \frac{\partial f_m(B, x)}{\partial B} \right\}. $$

Covariance matrix of coefficient estimates (2) is calculated by the formula

$$ V_n = \left\{ \sum_{j=1}^{n} P(B_j, x_j) V_e P(B_j, x_j)^T \right\}^{-1}, $$

(3)

where \( V_e \) is the covariance matrix of observational errors.

3. Calculation of non-linear calibration function coefficient estimates

The type of the model in the expression (1), nonlinear within coefficients \( B \), is selected on the basis of the study of physical processes that connect indirect and direct measurements. In most cases, it is possible to use polynomial models with a polynomial degree not higher than four. This model can be represented as the following multiresponse, linear within coefficients [3]:

$$ Y = P^T(x)B + E, $$

(4)

where \( x \) is the sought value, \( Y = (y_1, \ldots, y_m)^T \) are the observed values (responses) of each of the \( m \) sensors, \( B = (b_1, \ldots, b_l)^T \) are the estimated coefficients of the model (1), \( E = (e_1, \ldots, e_m)^T \) are the errors of observations of the sought values, \( P(X) \) is function describing the type of dependence of the sought value \( x \) from the indirect measurements \( Y \) which has a dimension equals to the number of model coefficients.

By comparing multiple models as a working model (1), a third-degree polynomial is accepted, for which the number of coefficients is equal \( k = 4m \) upon condition that each response is described by
the vector of independent coefficients. Then the matrix $P(x)$ having the dimension $(m \times km)$ takes the following view:[4]:

$$
P(x) = \begin{bmatrix}
1 & x_1 & \ldots & x^3 & 0 & 0 & 0 & \ldots & 0 \\
0 & 0 & 0 & 1 & \ldots & x^3 & 0 & \ldots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & \ldots & x^3
\end{bmatrix}.
$$

(5)

For linear within coefficients model (4) the estimations of the coefficients are calculated by the results of the experiment in the form of indirect measurements of $n$ values of the sought value $x$ for $m$ sensors according to the following formula [2]:

$$
B = \left( \sum_{j=1}^{n} P(x_j) P(x_j)^T \right)^{-1} \sum_{j=1}^{n} P(x_j) Y_j,
$$

(6)

where $x_j, j=1,n$, are the sought values, set in the experiment, $V_k$ is the covariance matrix of observational errors, $Y_j$ is the vector of indirect measurements, $n$ is the number of indirect measurements.

As an estimation of covariance matrix $V_k$ the matrix $S_k$ which is calculated by the following formula is used:

$$
S_k = \frac{1}{v-1} \sum_{j=1}^{n} (Y_j - \bar{Y})(Y_j - \bar{Y})^T,
$$

(7)

where $v$ is the number of response observations for the constant value of the sought $x$ value.

4. Check the adequacy of the model

As the method of checking the adequacy of the multiresponse model residual analysis is used. To test the hypothesis of equality to zero the mathematical expectation of the residual was used, $T^2$ is the Hotelling statistics [3] in the form

$$
T^2 = \frac{1}{nm-k} \sum_{j=1}^{n} R_j^T S_k^{-1} R_j,
$$

(8)

where $R_j = Y_j - F(B, x_j)$ is the vector of the residues of the mode in $j$-th observation, and the matrix $S_k$ was calculated by the formula (7).

If the residues have zero mathematical expectation, then the statistics (8) approximately has the Fischer distribution with degrees of freedom $f_1 = n - k / m$ and $f_2 = v - 1$.

If inequality holds

$$
T^2 < F_{f_1,f_2,1-\alpha},
$$

(6)

where $F_{f_1,f_2,1-\alpha}$ is the quantile of the Fisher distribution with degrees of freedom $f_1$ and $f_2$ at the level of confidence $\alpha = 0.9$.

Then the hypothesis for equality to zero of mathematical expectation of residues is accepted with confidence probability $1 - \alpha$.

5. Design of experiment

The values $x_j, j=1,n, \text{(design of experiment)}$ set in the experiment are used during calculating the calibration coefficient estimates (6). The accuracy of the coefficient evaluation depends on these values (3). The optimal design of experiment is calculated by a numerical method using this iterative procedure.
The design of experiment in this case is a set of control points, \( \mathbf{O} = \{x_1, x_2, \ldots, x_n\} \) where \( n \) is the number of points in experiment. The \( D \)-optimality criterion requires such arrangement of points in the design area \( \Sigma \), in which the determinant of the matrix \( \mathbf{V}_n \) is minimal. Thus, \( D \)-optimal design is defined as follows:

\[
\text{det} \, \mathbf{V}_n(\mathbf{O}^*) = \min_{\mathbf{x} \in \Sigma} \text{det} \, \mathbf{V}_n(\mathbf{x}) = \min_{\mathbf{x} \in \Sigma} \left[ \sum_{j=1}^{m} \mathbf{P}(\mathbf{x}_j) \mathbf{V}_j \mathbf{P}^T(\mathbf{x}_j) \right]^{-1},
\]

where \( \mathbf{O}^* = \{x_1^*, x_2^*, \ldots, x_n^*\} \) is the optimal design of experiment in terms of the criterion of \( D \)-optimality.

The covariance matrix of the target function \( \mathbf{Y} \) calculated by the model for the vector of independent variables \( \mathbf{x} \) equals

\[
\text{Var}[\mathbf{Y}(\mathbf{O}^*)] = \mathbf{V}_x(\mathbf{O}^*) = \mathbf{P}^T(\mathbf{x}) \mathbf{V}_x \mathbf{P}(\mathbf{x}).
\]

Extreme task (7) has a dimension \( ml \), where \( l \) is the number of model coefficients, \( m \) is the number of observed parameters, is complicated in the sense of convergence computational procedure and the cost of machine time when using standard procedures for the extremum search. In this case the problem is complicated by the condition that the design \( \mathbf{O} \) should be discrete, i.e. the control points should belong to some discrete set of permissible points area, and precise, since the number of control points should be precisely defined before starting the experiment. In view of the above limitations the following modified algorithm of the optimal design of experiment search is proposed [4].

1. Specify the number of control points \( n \) and determine the arbitrary initial design of the experiment \( \mathbf{O}^1 = \{x_1^1, x_2^1, \ldots, x_n^1\} \).
2. The first point \( x_i \) is selected to search the local extreme.
3. Local extreme task is solved for one changeable control point \( x_i \) and fixed remaining points \( x_1, x_2, \ldots, x_i, \ldots, x_n \) and the design of experiment \( \mathbf{O}^j \) is determined in the form (7)

\[
\mathbf{O}^j = (x_1^j, x_2^j, \ldots, x_i^j, \ldots, x_n^j) = \max_{x \in \Sigma} \sum_{j=1}^{m} \mathbf{P}(\mathbf{x}_j) \mathbf{V}_j \mathbf{P}^T(\mathbf{x}_j)
\]

For the model linear within coefficients type of three response polynomial of the third degree.  
4. The procedure in item 3 is repeated for all points in the design.
5. Items 3-4 are repeated until the specified accuracy of calculations, specified number of iterations or before the expiration of the specified time of the account has been reached.

6. Evaluation of the sought value and its variance

After the first stage of indirect measurements based on the results of the experiment determined the kind of multi-response calibration function (4) and calculated estimates of its coefficients, one can proceed to the measurement of the sought value \( \mathbf{x} \) by indirect method. In the second stage by means of the calibration function according to the results of measurement of sensors responses, \( \mathbf{Y} = (y_1, \ldots, y_m)^T \) the sought value \( \mathbf{x} \) and its confidence interval are calculated.

The iterative procedure of calculating the vector of estimates of the sought value for each sensor can be presented as follows [1]:

\[
\mathbf{X}^{(i+1)} = \mathbf{X}^{(i)} - \left[ \frac{\partial \Phi(\mathbf{B}, \mathbf{X}', \mathbf{Y})}{\partial \mathbf{X}} \right]^{-1} \Phi(\mathbf{B}, \mathbf{X}', \mathbf{Y}),
\]

where \( \Phi(\mathbf{B}, \mathbf{Y}, \mathbf{X}) = \mathbf{Y} - \mathbf{F}(\mathbf{B}, \mathbf{X}), \mathbf{X} = (x_1, \ldots, x_m)^T \) is the vector of estimates of the sought value for each sensor.

The covariance matrix (8) for all sensors depends on the matrices \( \mathbf{V}_n \) (3) and \( \mathbf{V}_e \) (defined in the first stage):
\[
V_x = AV_xA^T + GV_xG^T,
\]
where matrices \(A\) and \(G\) are calculated as follows:
\[
A = \left[ \frac{\partial \Phi(B, X, Y)}{\partial B} \right]^T \left[ \frac{\partial \Phi(B, X, Y)}{\partial X} \right]^{-1} - \Phi(B, X, Y) \left[ \frac{\partial \Phi(B, X, Y)}{\partial X} \right]^2 \left[ \frac{\partial \Phi(B, X, Y)}{\partial B} \right],
\]
\[
G = \left[ \frac{\partial \Phi(B, X, Y)}{\partial Y} \right]^T \left[ \frac{\partial \Phi(B, X, Y)}{\partial X} \right]^{-1} - \Phi(B, X, Y) \left[ \frac{\partial \Phi(B, X, Y)}{\partial X} \right]^2 \left[ \frac{\partial \Phi(B, X, Y)}{\partial Y} \right].
\]

The best linear estimate of the sought value \(\bar{x}\) is calculated according as follows:
\[
\bar{x} = (J^TV_x^{-1}J)^{-1}J^TV_x^{-1}\hat{x}
\]
(10)

where \(J\) is a vector with dimension \(m\) whose elements are equal to one, \(\hat{x}\) is a vector of local estimates obtained from formula (8).

Variance \(s_x^2\) of estimate (10) is
\[
s_x^2 = (J^TV_x^{-1}J)^{-1},
\]
(11)

that is less than the least variance of local estimates.

With independent estimates \(\hat{x}\), i.e. when the covariance matrix \(V_x\) is a diagonal matrix, expressions (10) and (11) are converted into a known formula for of varying accuracy observations.

7. Results of group indirect measurements on the basis of multiresponse functions

Large values of correlation coefficients of local estimates significantly increase the variance of the overall estimate, so if one does not take into account the strong correlation, the variance of the estimate of the sought value will be significantly underestimated.

With large positive correlation coefficients, the estimate of average \(\bar{x}\) may be less than the least of the local estimates \(x_i\). Let us, for example, for \(m = 2\) (two sensors) known local estimates are \(x_1 = 50\) and \(x_2 = 60\). Let, also, the covariance matrix \(V_x\) equals
\[
V_x = \begin{bmatrix} 1 & 1.2 \\ 1.2 & 2 \end{bmatrix}.
\]

Then from (10) it is possible to get the estimate of the sought value \(\bar{x} = 46.67\) at variance of estimation of average (11) \(s_x^2 = 0.94\). Without taking into account the correlation, i.e. considering the matrix \(V_x\) as diagonal, we get \(\bar{x} = 53.33\) and \(s_x^2 = 0.67\). Thus, the error of estimating of the sought value and its covariance without correlation can be 14.3% and 28.5%, respectively. This is a fairly significant error, so in calculating the average on varying accuracy observations, the verification of the independence of local estimates is necessary. If these values are not independent, use expressions (10) and (11) to estimate the average.

The method of indirect measurements based on multiresponse function was used to estimate the magnitude of the homogeneous color (sought value) of computer image by the pixel color values of the photoreceiver matrix [5]. As the object of research the CCD matrix of a photoreceiver ISX455AQF of type with image size 1712 \(\times\) 1200 pixels (2 MP) was used. Images of homogeneous colors of the RAL D2 system containing 1625 color samples were carried out. Images were executed in macromode several times for each color with the rotation of 45°; then the color coordinates were averaged. As a result, for standard color coordinates in the CIE LAB system with known color coordinates were calculated colors for each pixel. As a result of calculations it was established that the coefficients in the model (4) in this case can be accepted as the same for all pixels of the photoreceiver matrix. Thus,
the number of coefficients of the overall image model (4) is assumed to be four, so the matrix $\mathbf{P}^T(x)$ had a dimension $(4 \times m)$ and took the following form.

$$
\mathbf{P}^T(x) = \begin{bmatrix}
1 & x_1 & \ldots & x_4 \\
\vdots & \vdots & \ddots & \vdots \\
1 & x_m & \ldots & x_m
\end{bmatrix}.
$$

Calculation of average values of color coordinates on homogeneous colors of RAL system showed that the error of color reproduction by the photoreceiver matrix is (6-8)%. Calculation of color coordinates according to the formula (8) on the basis of the built model (4) and obtained covariance matrix of these estimates $\mathbf{V}_x$ (8) allowed to adjust this error to the values of (0.8-2)%.

**Conclusion**

Indirect measurements in the engineering allow to estimate the values for which direct measurements are difficult or impossible. For this purpose the set of sensors is used, according to which indications it is possible to calculate the sought value with minimal error. The more such sensors are, the more accurate the sought value is measured. To ensure the most accurate evaluation of the measured value, it is necessary to take into account the type of dependence of sensor readings from the measured sought value and their statistical relations. It is possible to increase the accuracy of evaluation by using multiresponse model that describes the dependence of sensor readings on the measured sought value. This paper describes the method of estimating indirect measurements based on the building of a multiresponse model. This method involves design and carrying out an experiments to select a type of model, calculate its coefficients, and covariance matrices. Then, on the basis of the received information calculation of point and interval estimation of the measured value is carried out. The carried out experiment on estimation of homogeneous color of digital image has shown, that this method allows in this case to increase accuracy of estimation up to (5-7)%.

**References**

[1] *Recommendation. State system for ensuring the uniformity of measurements. Indirect measurements. Evaluation of measurement results and estimation of their errors. MI 2083-90 1991.* Standards Publishing House 1-4
[2] Rencher, Alvin C. 2002 *Methods of multivariate Analysis* A John Wiley & Sons, Inc. Publication. Brigham Young University pp339-345
[3] Popov S. A. and Klykov N. N. 2013 *Vestn. Novgorod. Gos. Univ.* 2 (75) pp 96-98
[4] Popov S. A. 2016 *Vestn.estn. Novgorod. Gos. Univ.* 4 (95) pp 53-57
[5] Popov S.A., Emelyanov G.M. 2005 *Pattern Recognition and Image Analysis* 15(2) pp 63-65