Solution of overdetermined systems of equations using the conforming subsystem selection

V A Fursov\textsuperscript{1,2}, A P Kotov\textsuperscript{1,2} and Ye V Goshin\textsuperscript{1,2}

\textsuperscript{1}Samara National Research University, Moskovskoye shosse, 34, Samara, Russia, 443086
\textsuperscript{2}Image Processing Systems Institute of RAS - Branch of the FSRC "Crystallography and Photonics" RAS, Molodogvardejskaya street 151, Samara, Russia, 443001

e-mail: fursov@ssau.ru, kotov@ssau.ru, goshine@yandex.ru

Abstract. Within this research, we consider an overdetermined system of equations generated on a small number of observations. The degrees of freedom of the overdetermined system and the number of observations are roughly equal. We propose a modification of the Conformed Estimates Method (CEM). The method improvement consists of the usage of an auxiliary system, which is formed from the initial system by reducing the columns. The estimates can be obtained on the higher-level auxiliary subsystem, which is the freest from noise. We consider our method in the problem of computing Rational Polynomial Coefficients (RPCs) from well-established Ground Control Points (GCPs).

1. Problem statement

We consider a problem in determining the estimates $\hat{c}$ of the $M$-dimensional parameter vector $c = [c_1, c_2, \ldots, c_M]^T$ of the overdetermined system

$$Y = Xc + \xi,$$

where $Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}$, $X = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,M} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,M} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N,1} & x_{N,2} & \cdots & x_{N,M} \end{bmatrix}$, and $\xi = [\xi_1, \xi_2, \ldots, \xi_N]^T$ is a vector of error term. It is assumed that the number of observations, according to the system (1), is small. Accordingly, the number of degrees of freedom $S = N - M$ is comparable with the number $M$. The formulated problem is found in a wide range of science applications. Examples can be found in [1, 2].

The least squares estimate is obtained from equation (1), as shown in equation (2):

$$\hat{c} = (X^T X)^{-1} X^T Y.$$
free. To find such a subsystem, a set of estimates is computed for each applicant-subsystem. The measure of closeness is derived for each set of the estimates based on lower-layer subsystems (low level) belonging to a higher-layer subsystem (high level). Thus, the best solution is obtained on this higher-layer subsystem.

The feature of CEM is to generate all possible subsystems of high level. For each of these subsystems, all possible subsystems of the lower level are formed. Obviously, even with a relatively small dimension of the original system (1), it is computationally difficult. In [5] we have proposed an algorithm for the generation of conformed subsystems which ensures the preservation of “good” properties of estimates at lower computational costs.

The algorithm is designed as an iterative procedure, where at each iteration one gross error is rejected. If there are more errors, the procedure is repeated the required number of times. All \( N \) higher-layer subsystems consecutively formed with not included one row of the initial system (1):

\[
\mathbf{Y}_l = \mathbf{X}_l \mathbf{c}_l + \xi_l, \quad l = 1, N,
\]

where \( \mathbf{X}_l \) is a \((N - 1) \times M\)-dimensional matrix of the higher-layer subsystem \( l \).

Further, we can generate \( (N - 1) \) lower-layer subsystems from each subsystem (3):

\[
\mathbf{Y}_k = \mathbf{X}_k \mathbf{c}_k + \xi_k, \quad k = 1, N - 1,
\]

where \( \mathbf{X}_k \) is an \( M \times M \)-dimensional matrix formed by shift instruction. Every row of the higher-layer subsystem \( l \) uses the \( M \) number in the lower-layer subsystem. Next, the conformity function is calculated at each \( l \) higher-layer subsystem (3):

\[
W(l) = \frac{2}{(N - 1)(N - 2)} \sum_{i,j=1 \atop i \neq j}^{N-1} \| \xi_{i,j} - \hat{\xi}_{i,j} \|, \quad (5)
\]

where \( \| \xi_{i,j} - \hat{\xi}_{i,j} \| \) is the Euclidean distance of one pair of estimates from all possible pairwise combinations of \( (N - 1) \) estimates, which can be obtained in a \( k \) lower-layer subsystem. The most conformed subsystem \( \hat{l} \) on a high level can be defined as:

\[
W(\hat{l}) = \min_{l \in L} W(l), \quad (6)
\]

in which the final estimate \( \hat{\mathbf{c}} \) is then calculated. The choice of a \( \hat{l} \) subsystem performed by criterion (6), which ensures a comparability of the values of conformity function \( W(l) \) because all the higher-level and lower-level subsystems have the same dimension, respectively. Each subsystem of the high level \( \hat{l} \) have the same number \( (N - 1) \) of lower-level subsystems and the same scheme of their enumeration. These requirements are needed to generate a conformed subsystem \( \hat{l} \). Note that the final parameter estimation of the conformed subsystem \( \hat{l} \) can be carried out with any known method.

In this paper, we develop the idea of conformed estimation. In particular, we consider a new algorithm for forming a conformed subsystem using an auxiliary system. We can get it from the original overdetermined system (1) by reducing the columns.

2. Method

The goal of the method is to form an auxiliary overdetermined system,

\[
\mathbf{Y} = \mathbf{X} \hat{\mathbf{c}} + \tilde{\xi}, \quad (7)
\]

where \( \mathbf{X} \) is an \( N \times P \)-dimensional matrix with \( P < M \). The parameter estimation of the auxiliary system (7) is carried out by using the algorithm for the generation of conformed subsystems described in the previous section. The algorithm is applied to the system (7) in the same way as if it were applied to the original system (1). Thus, identical sets of row indices are obtained. For this, the requirement to save all variables with errors of the auxiliary system in the same rows of the original system (1) must
be fulfilled. Thus, the auxiliary system is formed by reducing columns that contain measurement
errors or contain errors present in other columns.

We assume that to meet the above requirements in the matrix $X$ of the system (1) it suffices to leave
the $P$ columns. The columns can always be rearranged. Accordingly, we will assume that the
last $(M - P)$ columns can be reduced. The reducing of these columns is achieved by multiplying the
matrix $X$ on the block $M \times P$-dimensional matrix:

$$
\tilde{X} = X \cdot P,
$$

where $P = \begin{bmatrix} E_p \\ 0 \end{bmatrix}$, $E_p$ is an all-ones $P \times P$-matrix.

The resulting $N \times P$ matrix $\tilde{X}$, as well as the original matrix $X$, forms an $N$-dimensional space.
However, if the original matrix $X$ had a $M$-dimensional basis, then the matrix $\tilde{X}$ will have $P$
linearly independent rows ($P$-dimensional basis). Consequently, the overdetermined auxiliary system
(7) will have the only least-squares estimate. It’s the $(P \times 1)$ vector:

$$
\hat{c} = (\tilde{X}^T \tilde{X})^{-1} \tilde{X}^T Y.
$$

The vectors of estimates $\hat{c}$ and $\hat{c}$ differ in the dimension and in the numerical values of the
components, since they are the result of the least-squares method (LSM)-projection of the vector $Y$
onto different subspaces. However, in this case we are not interested in the accuracy of parameter
estimation of the original model (1). Our goal is to identify the most noise-free subsystem. If, as a
result of the transformation $\tilde{X} = X \cdot P$, all variables that may contain errors occur at least once (only
their repeated occurrences or combinations are excluded), then the row numbers of the most
conformed higher-level subsystem extracted from systems (1) and (7) are most likely to match.

The requirement of preserving all variables containing errors is crucial in transformation $\tilde{X} = X \cdot P$
to determining the row numbers of the most conformed subsystem. With that said, the algorithm can
be presented as shown in Figure 1.

![Figure 1. Algorithm for the generation of conformed subsystems using reducing the columns.](image)

The described method is considered on the example of determining the parameters of the Rational
Polynomial Coefficients (RPC) model of image formation by a satellite camera imaging. This model
establishes the relationship between the geodetic coordinates object and its coordinates in the image
using the RPCs.

## 3. Problem of determining the RPC-model of image formation

Let’s consider the application of the proposed method to solving the problem of determining the
parameters of the RPC-model of image formation in the Earth remote sensing systems. The spacecraft
is in motion when taking images. Therefore, there are image transformations and distortions associated
with external parameters (motion parameters of the spacecraft), along with distortions associated with
internal camera parameters (focal length, scaling parameters and distortion). Since the external
parameters of the camera change in flight, they are usually refined using an RPC model.

To determine the RPCs (parameters of the model), a test image is used that contains a number of
the geo-reference points, whose coordinates and relative heights are known. The georeferenced points
are generated using the Global Positioning System (GPS) and landmarkers on the surface of the earth.
They can be randomly placed over the image. The sets of unknown RPCs \( a, b, c, d \) are the parameters of the satellite imaging model of the camera:

\[
Y = \left( a^T u \right) / \left( b^T u \right), \quad X = \left( c^T u \right) / \left( d^T u \right). \tag{10}
\]

Here \( Y, X \) – the normalised image coordinates:

\[
Y = \left( y - Y_0 \right) / Y_s, \quad X = \left( x - X_0 \right) / X_s, \tag{11}
\]

where \( y, x \) – the pixel coordinates of the row and column, \( Y_0, X_0 \) – the offset factors for the row and column, \( Y_s, X_s \) – the scale factors for the row and column.

The vector \( u \) is composed of various degrees (usually up to the third degree inclusive) of the coordinates of a three-dimensional point – \( P, L, H \) obtained by normalising the geodetic coordinates \( \varphi, \lambda, h \):

\[
P = \left( \varphi - \varphi_0 \right) / \varphi_s, \quad L = \left( \lambda - \lambda_0 \right) / \lambda_s, \quad H = \left( h - h_0 \right) / h_s, \tag{12}
\]

where \( \varphi, \lambda, h \) – latitude, longitude and height respectively, \( \varphi_0, \lambda_0, h_0 \) – the offset factors and \( \varphi_s, \lambda_s, h_s \) – the scale factors. Further, without loss of generality, we consider the simplest case of a first-order model with a vector

\[
u = [1, L, P, H]^T.
\]

The problem lies in estimating the parameters \( \varphi, \lambda, h \), \( P, L, H, \) of a model from a set of known three-dimensional coordinates, \( Y, X, \) and their corresponding coordinates, \( Y, X, \) observed on the multi-view images [6-9]. For the \( N \) control points of the set of coordinates, according to (13), we can write the matrix equation (hereinafter, unlike formula (1), we use the following notation)

\[
Y = MJ + \xi, \tag{13}
\]

where

\[
M = \begin{bmatrix}
1 & L_1 & P_1 & H_1 & -Y_1L_1 & -Y_1P_1 & -Y_1H_1 \\
1 & L_2 & P_2 & H_2 & -Y_2L_2 & -Y_2P_2 & -Y_2H_2 \\
& \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1 & L_N & P_N & H_N & -Y NL_N & -Y NP_N & -Y NH_N
\end{bmatrix},
\]

\[
J = [a_0, a_1, a_2, a_3, b_1, b_2, b_3]^T, \quad Y = [Y_1, \ldots, Y_N]^T,
\]

and \( \xi = [\xi_1, \xi_2, \ldots, \xi_N]^T \) – the random error vector, associated with model order errors, inaccurate definition of three-dimensional geodesic coordinates and errors in determining the coordinates of control points on an image.

Traditionally [1,2], the Least-Squares Method (LSM) is used to solve the formulated problem. In accordance with (13), the least-squares estimate of the parameter vector composed of the unknown RPCs, we get:

\[
\hat{J} = (M^TM)^{-1} M^T \bar{Y}. \tag{14}
\]

It is known that the LSM is very sensitive to gross errors such as failures [10-11]. The best result is achieved using the Least Absolute Deviations (LAD) method. We compare these methods with the proposed modification of the Conformed Estimates Method (CEM) [3-5]. Modification consists of generating a conformed subsystem using an auxiliary system obtained from the original system (13) by reducing the columns.

It is easy to see that, in the last three columns of the matrix \( M \) in (13), the variables \( L, P, H \) present as factors of the vector components \( Y, \), which are also present in the previous columns. Since the vector \( Y \) is also present on the left side of system (13), it is clear that after reducing the last three columns from the matrix \( M \) all variables that may contain errors will remain in the transformed system. Thus, an auxiliary system (7) can be used to determine the most conformed subsystem. In the notation of this section it can be written as

\[
Y = \bar{M}J + \bar{\xi}, \tag{15}
\]
where \( \mathbf{M} = \begin{bmatrix} 1 & L_1 & P_1 & H_1 \\ 1 & L_2 & P_2 & H_2 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & L_N & P_N & H_N \end{bmatrix} \),

the vector \( \mathbf{Y} \) is the same as in the original system (13). For the auxiliary system (15), the above formulated requirement for the preservation of all variables that may contain errors (in this case \( Y, P, L, H \)) is satisfied. Therefore, we can apply the algorithm for the generation of conformed subsystems shown in Figure 1. Note, according to (10), two sets of RPCs \( \mathbf{J}_y \) and \( \mathbf{J}_x \) must be determined for one image. Let us calculate the conformity function (5) as the mean of the functions \( W_r(l) \) and \( W_X(l) \).

4. Experimental results

We used \( N = 10 \) points uniformly distributed over the “Mountains” image as the initial data (Figure 2). To test how the accuracy of parameter estimations depend on the values of the conformity function (5), we introduced an additional random error into the coordinates \( (Y_i, X_i), i = 1,N = 10 \) of each of the 10 points. The additional errors were formed by generating a sequence of uniformly distributed random values in the interval \([-0.1, 0.1]\).

Table 1 shows the values of the conformity functions (5) for the 10 higher-level subsystems obtained in 10 implementations of the algorithm (Figure 1), with one row excluded. The values of the conformity functions (5) for all 10 \((N-1)\) -dimension subsystems obtained in each \( k \) implementation are grouped in columns. In each \( k \) column, \( l \) there is the number of the higher-level subsystem. The number \( l \) coincides with the row number of the source system (13). An additional random error was introduced into this row.

It is easy to see that in 9 out of 10 cases the minimum value of the conformity function is achieved if the row with an additional error is excluded. Only in the third implementation \((k = 3)\) is the minimum value of the conformity function (5) reached by dropping the first row in which no additional error was entered.

| \( l \) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.70 | 1.61 | 0.77 | 4.25 | 1.29 | 2.80 | 1.11 | 2.35 | 4.15 | 2.97 |
| 2 | 2.75 | 0.70 | 1.11 | 6.72 | 1.80 | 5.09 | 1.56 | 2.45 | 3.78 | 5.75 |
| 3 | 2.56 | 1.76 | 0.89 | 4.91 | 1.29 | 2.72 | 1.24 | 2.44 | 4.80 | 3.20 |
| 4 | 3.20 | 2.03 | 1.00 | 0.80 | 2.21 | 2.12 | 1.02 | 2.21 | 4.59 | 2.39 |
| 5 | 2.43 | 1.54 | 0.88 | 4.26 | 0.76 | 2.42 | 0.97 | 2.09 | 3.96 | 2.72 |
| 6 | 3.08 | 1.66 | 1.27 | 3.81 | 1.22 | 0.94 | 1.28 | 2.23 | 3.97 | 3.31 |
| 7 | 3.05 | 1.28 | 0.91 | 4.90 | 1.03 | 3.03 | 0.71 | 1.84 | 4.39 | 2.76 |
| 8 | 3.41 | 2.04 | 1.22 | 4.65 | 1.70 | 2.94 | 1.36 | 1.06 | 5.45 | 4.29 |
| 9 | 3.59 | 2.48 | 1.52 | 4.60 | 1.58 | 3.00 | 2.07 | 2.68 | 0.88 | 4.24 |
| 10 | 3.38 | 1.79 | 1.21 | 4.84 | 1.80 | 2.95 | 1.20 | 2.43 | 4.48 | 0.78 |

Figure 2. The “Mountain” image with the sets of training points - □ and test points - ▲.
Apparently, the additional error (which was formed as a uniformly distributed random value from the interval \([-0.1, 0.1]\)) turned out to be comparable with the model errors. This is also confirmed by the data in Table 2, in which the root-mean-square error (RMSE) values were obtained on the higher-level subsystems. The results show that the minimum values of RMSE are achieved on higher-level subsystems which are obtained by eliminating rows using the criterion of the conformed function (6). As expected, noticeable differences are obtained only for implementation with a number \(k = 3\). For this implementation the values of RMSE on all subsystems are significantly less than for other implementations. It confirms the explanation given above.

### Table 2. The RMSE values for 10 implementation (in pixels).

| \(k\) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|------|---|---|---|---|---|---|---|---|---|----|
| 1    | **2.15** | 3.35 | 1.72 | 14.03 | 3.05 | 5.42 | 1.07 | 3.61 | 8.29 | 10.20 |
| 2    | 9.14 | **1.97** | 1.96 | 10.74 | 7.04 | 2.85 | 2.40 | 3.61 | 7.30 | 9.32 |
| 3    | 2.90 | 3.68 | **2.18** | 14.34 | 5.13 | 6.06 | 1.37 | 2.94 | 9.83 | 9.39 |
| 4    | 9.14 | 2.21 | 1.95 | **2.19** | 7.40 | 6.32 | 2.24 | 4.15 | 7.42 | 10.70 |
| 5    | 9.92 | 3.66 | 1.96 | 11.82 | **2.36** | 6.17 | 2.50 | 4.36 | 10.30 | 9.12 |
| 6    | 6.89 | 3.02 | 1.87 | 13.37 | 6.29 | **1.88** | 2.21 | 1.50 | 9.79 | 8.12 |
| 7    | 2.96 | 2.91 | 1.35 | 12.02 | 6.54 | 1.98 | **1.22** | 3.57 | 5.74 | 10.30 |
| 8    | 9.21 | 3.46 | 1.82 | 9.29 | 7.58 | 6.09 | 2.53 | **2.33** | 8.34 | 8.20 |
| 9    | 7.94 | 3.28 | 0.98 | 9.06 | 7.04 | 5.19 | 2.30 | 4.09 | **1.74** | 9.08 |
| 10   | 9.52 | 3.45 | 2.21 | 13.43 | 7.20 | 4.79 | 2.52 | 2.81 | 9.82 | **2.09** |

### Figure 3. The errors for 20 samples: LSM (\(^\ast\)), LAD (\(+\)) and CEM (\(*\)).

Figure 3 shows a graph of changes in estimation errors for 20 small data sets by three methods: LSM (noted by the sign \(^\ast\)), Least Absolute Deviations method (LAD – \(+\)) and Conformed Estimates Method (CEM – \(*\)).

Table 3 shows the RMSE values and mean absolute error (MAE) values obtained by the same three methods on a test example for 1000 implementations.

### Table 3. Comparison of methods on 1000 repeated samples.

|         | LSM     | LAD     | CEM     |
|---------|---------|---------|---------|
| **RMSE** | 1.2410  | 0.5118  | 0.2298  |
| **MAE**  | 0.9843  | 0.2897  | 0.0581  |

The use of the described algorithm provides a reduction in the RMSE values of the parameter estimates by more than 2 times compared with the LAD and more than 5 times compared with the LSM.

### 5. Conclusion
The results show that the method of forming a conformed subsystem using an auxiliary system obtained from the original system by reducing the columns provides a more accurate estimate of the...
RPCs compared with the LSM and LAD methods. A dependence is established between the accuracy of solutions on the subsystems of the higher-level subsystem and the conformed function corresponding to this subsystem.

6. References

[1] Mudrov V I, Kushko V L 1976 *Methods of processing of measurements* (Moscow: Sovetskoe radio) p 192

[2] Lawson C L and Hanson R J 1995 *Solving least squares problems* (Siam) 15

[3] Fursov V A 2003 Estimates Conformity Principle in the Problems of Identification *International Conference on Computational Science* 463-470

[4] Fursov V A, Gavrilov A V and Goshin Ye V 2016 Parametric Identification by Means of Consequent Formation of a Conformed Estimations *IFAC-PapersOnLine* 49(13) 123-128

[5] Fursov V A, Gavrilov A V and Kotov A P 2018 Prediction of Estimates' Accuracy for Linear Regression with a Small Sample Size *41st International Conference on Telecommunications and Signal Processing (TSP)* 679-685

[6] Zheng E, Wang K, Dunn E and Frahm J M 2015 Minimal solvers for 3D geometry from satellite imagery *Proceedings of the IEEE International Conference on Computer Vision* 738-746

[7] Grodecki J, Dial G 2003 Block adjustment of high-resolution satellite images described by rational polynomials *Photogrammetric Engineering & Remote Sensing* 69(1) 59-68

[8] Hu Y, Tao C V and Croitoru A 2004 Understanding the rational function model: methods and applications *International archives of photogrammetry and remote sensing* 20(6)

[9] Tao C V, Hu Y 2001 A comprehensive study of the rational function model for photogrammetric processing *Photogrammetric Engineering & Remote Sensing* 67(12) 1347-1358

[10] Afanasyev A A, Zamyatin A V 2017 Hybrid methods for automatic landscape change detection in noisy data environment *Computer Optics* 41(3) 431-440 DOI: 10.18287/2412-6179-2017-41-3-431-440

[11] Kholopov I S 2017 Projective distortion correction algorithm at low altitude photographing *Computer Optics* 41(2) 284-290 DOI: 10.18287/0134-2452-2017-41-2-284-290

Acknowledgments

Methods and algorithms were developed with the support of the Ministry of Science and Higher Education of the Russian Federation (project # 2.891.2017) and RFBR (project # 17-29-03112), experimental studies - in the framework of the state assignment of the IPSI RAS - branch of the Federal Scientific-Research Centre "Crystallography and Photonics" of the RAS (agreement № 007-Г3/13363/26).