HELMERT TRANSFORMATION FEATURES WITHIN A TWO-DIMENSIONAL SPACE

Ph.D, Docent Victor Ziborov,
ziborov@ukr.net, ORCID: 0000-0003-4397-1782,
Ph.D, Docent Igor Patrakeyev,
ipatr@ukr.net, ORCID: 0000-0002-0448-8790

Kyiv national university of construction and architecture

KEY WORDS: Helmert transformation, four-parameter transformation of plane coordinates, least squares method.

Relevance. The Helmert coordinate transformation problem occurs in many areas of engineering activity. In the field of geoinformatics, all geographic information systems (ArcGIS, QGIS, MapInfo and others) have the ability to link raster images using the Helmert method. All local coordinate systems must be referenced to a national geodetic reference. This is a requirement of regulatory documents in almost any country. All construction networks must be linked to the state geodetic network. In photogrammetry, the Helmert transform is also used along with the affine transform. In astrometry, digital images are converted to a star's coordinate system based on the Helmert transform. We have listed a far from complete list of engineering areas where Helmert transformations are used. The reason for such demand for this particular method of coordinate transformation is its comparative simplicity. In addition, the Helmert transformation ensures a conformal map, that is, preservation of angles, and therefore the shape of infinitesimal figures when transforming coordinates from one system to another system.

Very often, by Helmert transformation they mean a seven-parameter transformation in three-dimensional space (seven-parameter Helmert transformation) (see, for example, [3, 4]). This is an important task, and the features that are solved in the transformation in two-dimensional space must (if possible) be transferred to three-dimensional space. However, it's not that simple. There is a so-called curse of dimensionality, when as the dimension of a problem increases, the complexity of its solution increases significantly.

In our case, we have two coordinate systems in two-dimensional space. In the simplest case, the origins of coordinates of these systems coincide. In this case, one system is rotated relative to the other system at an angle θ. It is very easy to obtain formulas relating these two coordinate systems using polar coordinates R and φ, see fig. 1.

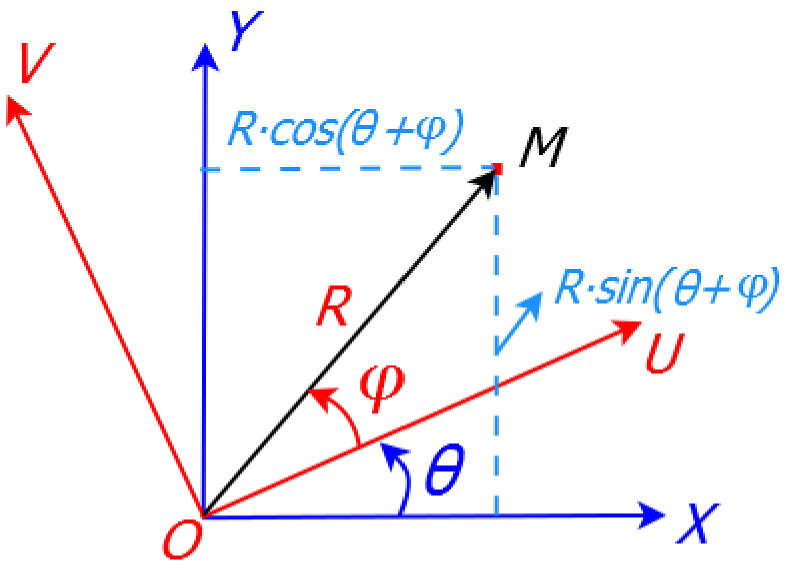

Figure 1. Communication between two rectangular systems

From the figure it is clear that:

$$\begin{cases} X = R \cdot \cos(\theta + \varphi) = R \cdot \cos\theta \cdot \cos\varphi - R \cdot \sin\theta \cdot \sin\varphi \\ Y = R \cdot \sin(\theta + \varphi) = R \cdot \sin\theta \cdot \cos\varphi + R \cdot \cos\theta \cdot \sin\varphi \end{cases}$$

And since polar coordinates are related to the coordinate system $U0V$ like this:

$$\begin{cases} U = R \cdot \cos\varphi; \\ V = R \cdot \sin\varphi, \end{cases}$$

Then we get the following expression:

$$\begin{cases} X = U \cdot \cos\theta - V \cdot \sin\theta, \\ Y = U \cdot \sin\theta + V \cdot \cos\theta. \end{cases}$$

The last expression connects two rectangular systems when one system is rotated relative to the other system by an angle of rotation θ. In this case, the origins of coordinates of both systems coincide.

However, in the general case, the origins of coordinates of both systems do not coincide. Moreover, there is not only a shift X_0 and Y_0, but also scaling s:

$$\begin{cases} X = X_0 + s \cdot U \cdot \cos\theta - s \cdot V \cdot \sin\theta \\ Y = Y_0 + s \cdot U \cdot \sin\theta + s \cdot V \cdot \cos\theta \end{cases} \qquad (1)$$

Note that the Helmert transform takes into account only one scale. In photogrammetry, scaling along two directions (for example, along both coordinate axes) is taken into account. Therefore, in photogrammetry, the affine transformation is often used. The same applies to scanning paper plans and maps in geoinformatics to take into account paper deformation. Meanwhile, it should be remembered that an affine transformation will not provide a conformal mapping. That is, during an affine transformation, the angles between the curves will be distorted. This means that the shape of infinitesimal figures will not be preserved.

Let's write down the transformation (1) in matrix-vector form:

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} X_0 \\ Y_0 \end{bmatrix} + s \cdot \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} U \\ V \end{bmatrix}$$

In this case, it is easy to obtain the inverse transformation formulas. First, we move the shift vector to the left side, and then we divide both sides of the equation by the scale:

$$\begin{bmatrix} X - X_0 \\ Y - Y_0 \end{bmatrix} \cdot s^{-1} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} U \\ V \end{bmatrix}$$

Now we multiply both sides of the matrix equation on the left by the inverse matrix of the rotation matrix. This is easy to do since all rotation matrices are orthogonal. That is, to get the inverse matrix, you should simply transpose it:

$$\begin{bmatrix} U \\ V \end{bmatrix} = s^{-1} \cdot \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} X - X_0 \\ Y - Y_0 \end{bmatrix}. \qquad (2)$$

Thus, we got the formulas (2) for inverse coordinate transformation.

To uniquely determine the four parameters of the Helmert coordinate transformation X_0, Y_0, θ and s two points are needed whose coordinates are known in both coordinate systems. In real practice, geodetic, photogrammetric measurements, geoinformatics, and astrometry produce redundant measurements, so the problem is solved using the least squares method.

A specific program, for example, implemented in geodetic measurement processing systems, solves the conversion problem in two stages. The first step is to determine the parameters of the least squares transformation using common points. And the second stage is the calculation of the coordinates of points from one system to another that do not have common points, based on the obtained transformation parameters.

Let's look at a specific example of a two-dimensional Helmert transform using a numerical example. We want to clarify many details. In our problem, for simplification, but without loss of generality, only three points are given, the coordinates of which are known in two coordinate systems $X0Y$ and $U0V$:

Table 1. Coordinate catalog

N	X	Y	U	V

1	93168.687	43687.203	30173.173	30438.572
2	88685.536	39866.932	25769.805	26524.720
3	88652.913	42237.428	25687.734	28894.456

The coordinate values in this table are presented in meters; they are typical for solving problems solved by surveyors using them. Here, each coordinate has eight significant digits. Note that coordinates in the Gauss–Krüger projection have twelve significant digits accurate to the nearest millimeter. For comparison, to store a single-precision real number in any programming system, seven significant digits are provided. Therefore, it is necessary to use double variables of type double. In this case, we will have 15 significant digits to store a real number. We will consider this sufficient for operations with coordinates in the Gauss-Kruger projection, but with some reservations.

Transformation equations are given:

$$\begin{cases} X = X_0 + s \cdot U \cdot \cos\theta - s \cdot V \cdot \sin\theta \\ Y = Y_0 + s \cdot U \cdot \sin\theta + s \cdot V \cdot \cos\theta \end{cases}$$

It is necessary to find the law of transition from the coordinate system $U0V$ into the system $X0Y$. That is, find unknown parameters $X_0,\ X_0,\ s$ and θ using the least squares method.

In the given transformation equations, we have a nonlinear relationship between the unknown parameters. In the general case, one can expand the transformation equations into a Taylor series with respect to unknown parameters and limit oneself to the linear terms of the expansion. In this case, it is necessary to calculate approximate parameter values. In solving the Helmert transformation problem, some developers follow exactly this path.

However, we will take a different route. To linearize the transformation equations, we make the replacement: $\alpha = s \cdot \cos\theta$; $\beta = s \cdot \sin\theta$. Then the transformation equations will have the form:

$$\begin{cases} X = X_0 + U \cdot \alpha - V \cdot \beta \\ Y = Y_0 + U \cdot \beta + V \cdot \alpha \end{cases}$$

Now we have the same four unknown parameters: $X_0,\ Y_0,\ \alpha$ and β. But there is a linear connection between them. In the case of such linearization, there is no need to calculate approximate values of the unknowns. After finding the parameters $X_0,\ Y_0,\ \alpha$ and β, we can easily calculate the parameters s and θ:

$$\alpha^2 + \beta^2 = (s \cdot \cos\theta)^2 + (s \cdot \sin\theta)^2$$
$$s^2 = \alpha^2 + \beta^2$$

$$\theta = arctan\frac{\beta}{\alpha}$$

Following the scheme for solving the problem using the least squares method, we write the equations of residuals in general form:

$$\begin{cases} r_i^x = X_0 + U_i \cdot \alpha - V_i \cdot \beta - X_i; \\ r_i^y = Y_0 + V_i \cdot \alpha + U_i \cdot \beta - Y_i. \end{cases} \qquad (3)$$

It is clear that only the coordinates will receive residuals X and Y. That is, we assume the fact that the coordinates U and V derived from more precise measurements. If this is not the case, then the inverse problem should be solved (2).

Looking at the equations of residuals (3), let's make normal equations:

$$\begin{cases} [aa] \cdot X_0 + [ab] \cdot Y_0 + [ac] \cdot \alpha + [ad] \cdot \beta = [al], \\ [ab] \cdot X_0 + [bb] \cdot Y_0 + [bc] \cdot \alpha + [bd] \cdot \beta = [bl], \\ [ac] \cdot X_0 + [cb] \cdot Y_0 + [cc] \cdot \alpha + [cd] \cdot \beta = [cl], \\ [ad] \cdot X_0 + [db] \cdot Y_0 + [dc] \cdot \alpha + [dd] \cdot \beta = [dl], \end{cases} \qquad (4)$$

We will not solve these normal equations. We will need them in our further discussions.

Next, we compose the equations of residuals in numerical form:

$$\begin{cases} r_1^x = X_0 + 30173.173 \cdot \alpha - 30438.572 \cdot \beta - 93168.687 \\ r_2^x = X_0 + 25769.805 \cdot \alpha - 26524.720 \cdot \beta - 88685.536 \\ r_3^x = X_0 + 25687.734 \cdot \alpha - 28894.456 \cdot \beta - 88652.913 \\ r_1^y = Y_0 + 30438.572 \cdot \alpha + 30173.173 \cdot \beta - 43687.203 \\ r_2^y = Y_0 + 26524.720 \cdot \alpha + 25769.805 \cdot \beta - 39866.932 \\ r_3^y = Y_0 + 28894.456 \cdot \alpha + 25687.734 \cdot \beta - 42237.428 \end{cases}$$

Let's write the last system in matrix-vector form:

$$\begin{bmatrix} r_1^x \\ r_2^x \\ r_3^x \\ r_1^y \\ r_2^y \\ r_3^y \end{bmatrix} = \begin{bmatrix} 1 & 0 & 30173.173 & -30438.572 \\ 1 & 0 & 25769.805 & -26524.720 \\ 1 & 0 & 25687.734 & -28894.456 \\ 0 & 1 & 30438.572 & 30173.173 \\ 0 & 1 & 26524.720 & 25769.805 \\ 0 & 1 & 28894.456 & 25687.734 \end{bmatrix} \cdot \begin{bmatrix} X_0 \\ Y_0 \\ \alpha \\ \beta \end{bmatrix} - \begin{bmatrix} 93168.687 \\ 88685.536 \\ 88652.913 \\ 43687.203 \\ 39866.932 \\ 42237.428 \end{bmatrix}$$

or:

$$r = A \cdot z - l.$$

Here r – this is the vector of residuals,
A – matrix of residuals equations,
$z = \begin{bmatrix} X_0 \\ Y_0 \\ \alpha \\ \beta \end{bmatrix}$ – vector of unknowns,
l – free members vector.

Next, it would be possible to solve the system of residuals equations using, for example, the pseudoinverse matrix:

$$z = A^{+} \cdot l, \qquad (5)$$

where $A^{+} = (A^T \cdot A)^{-1} \cdot A^T$ – pseudoinverse matrix of matrix A.

However, first we calculate the condition number of the matrix A. This number turns out to be equal

$$\kappa(A) = \|A^{+}\| \cdot \|A\| = 5.9 \cdot 10^5.$$

Here $\kappa(A)$ – matrix condition number A,
$\|A^{+}\|$ – norm of pseudoinverse matrix A,
$\|A\|$ – matrix of A norm.

Condition number $\kappa(A)$ characterizes the task of (5). It shows how many times the output value can change $\frac{\|\Delta z\|}{\|z\|}$ with a small change in input data $\frac{\|\Delta l\|}{\|l\|}$. Norm $\|\cdot\|$ means the Euclidean norm of the corresponding vector.

A large value of the condition number ($5.9 \cdot 10^5$) indicates poor stability of the matrix A^{+}. It is clear that we cannot consider the result of solving the problem reliable $z = A^{+} \cdot l$, if the condition number $\kappa(A)$ equals more than half a million. Therefore, we will try to reduce the condition number of the matrix of residuals equations.

Let's solve this problem differently. To do this, we introduce the so-called centered coordinates. Let's find the average between all coordinates X_i, average between coordinates Y_i. Let's calculate the same for U_i and for V_i.
We will get four numbers:

$$\bar{X} = \frac{1}{3}\sum X_i, \ \bar{Y} = \frac{1}{3}\sum Y_i, \ \bar{U} = \frac{1}{3}\sum U_i \ \text{и} \ \bar{V} = \frac{1}{3}\sum V_i.$$

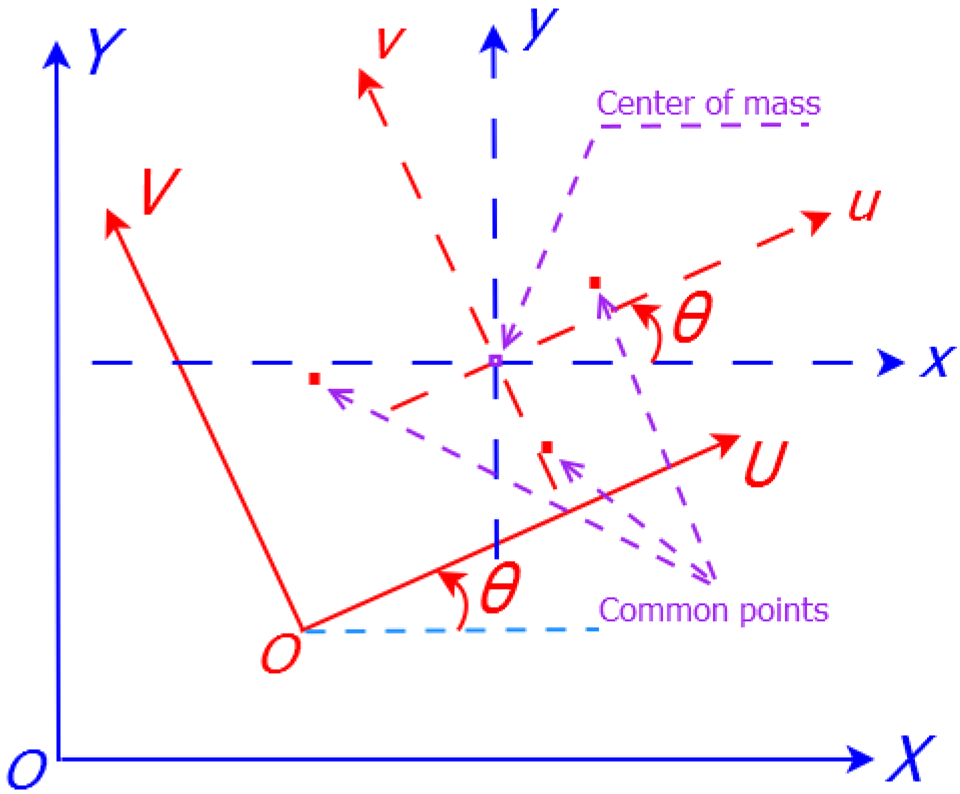

Figure 2. Transition to centered coordinates

These four coordinates point to one point (see Fig. 2), which is called the Center of mass. Let us move the origins of coordinates of both systems in parallel to this point. Since parallel translation has occurred, the rotation angle θ has not changed between the new coordinate systems.

For the center of gravity point we write the transformation equations:

$$\begin{cases} \bar{X} = X_0 + \bar{U} \cdot \alpha - \bar{V} \cdot \beta \\ \bar{Y} = Y_0 + \bar{U} \cdot \beta + \bar{V} \cdot \alpha \end{cases} \qquad (6)$$

As we see, if we knew the values α and β, then we could easily calculate the shift parameters X_0 and Y_0 using these formulas.

Now let's calculate the centered coordinates x_i, y_i, u_i and v_i. To do this, from each measured coordinate we subtract the corresponding coordinate of the center of gravity point:

$$x_i = X_i - \bar{X}; \quad y_i = Y_i - \bar{Y}; \quad u_i = U_i - \bar{U}; \quad v_i = V_i - \bar{V};$$

Now let's write the residuals equations for centered coordinates in matrix-vector form:

$$\begin{bmatrix} r_1^x \\ r_2^x \\ r_3^x \\ r_1^y \\ r_2^y \\ r_3^y \end{bmatrix} = \begin{bmatrix} +2962.9357 & -1819.3227 \\ -1440.4323 & +2094.5293 \\ -1522.5033 & -275.2067 \\ +1819.3227 & +2962.9357 \\ -2094.5293 & -1440.4323 \\ +275.2067 & -1522.5033 \end{bmatrix} \cdot \begin{bmatrix} \alpha \\ \beta \end{bmatrix} - \begin{bmatrix} +2999.6417 \\ -1483.5093 \\ -1516.1323 \\ +1756.6820 \\ -2063.5890 \\ +306.9070 \end{bmatrix}$$

Note that in these equations there is no longer X_0, Y_0, since, as can be seen in Figure 2, the origins of the centered coordinate systems coincide. In this case, the condition number of the matrix A will be equal to the minimum possible value for the condition number:

$$\kappa(A) = 1$$

Note that the less $\kappa(A)$, the better". That is, the errors in the solution will be smaller relative to the errors in the condition. The best condition number is 1.

Recall that in general, for centered coordinates, the equations of residuals have the form:

$$\begin{cases} r_i^x = u_i \cdot \alpha - v_i \cdot \beta - x_i \\ r_i^y = v_i \cdot \alpha + u_i \cdot \beta - y_i \end{cases} \qquad (7)$$

That is, now normal equations already look like this:

$$\begin{cases} \alpha \cdot \sum (u_i^2 + v_i^2) + \beta \cdot \sum (-u_i \cdot v_i + v_i \cdot u_i) = \sum (u_i \cdot x_i + v_i \cdot y_i) \\ \alpha \cdot \sum (-u_i \cdot v_i + v_i \cdot u_i) + \beta \cdot \sum (-v_i^2 + u_i^2) = \sum (-v_i \cdot x_i + u_i \cdot y_i) \end{cases}$$

As you can see, the elements in the side diagonal are equal to zero. That is, the last system of normal equations split into two independent equations. From these equations it follows:

$$\alpha = \frac{\sum (u_i \cdot x_i + v_i \cdot y_i)}{\sum (u_i^2 + v_i^2)}; \qquad (8)$$

$$\beta = \frac{\sum (-v_i \cdot x_i + u_i \cdot y_i)}{\sum (u_i^2 + v_i^2)}; \qquad (9)$$

Here it is necessary to draw some intermediate conclusions. In the beginning we had four normal equations (4) with four unknown parameters. However, thanks to the introduction of centered coordinates, we did not have to solve these normal equations. Source system (4) split into four independent equations: (8), (9) to calculate $\alpha\ and\ \beta$ and (6) to calculate shift parameters X_0 and Y_0.

It should be noted that to calculate α and β it is absolutely not necessary to move both systems to the center of gravity point. It is enough to place the origin of the system coordinates at the center of gravity $U0V$. In this case the parameters α and β, on which the rotation angle and scale factor depend can be calculated using other formulas:

$$\alpha = \frac{\sum(u_i \cdot X_i + v_i \cdot Y_i)}{\sum(u_i^2 + v_i^2)} = 0.99957326776067;$$
$$\beta = \frac{\sum(-v_i \cdot X_i + u_i \cdot Y_i)}{\sum(u_i^2 + v_i^2)} = -0.0208737106442;$$

That is, to calculate the parameters α and β it is not necessary to center the coordinates X and Y. Usually this fact is not taken into account.

Now found α and β can be used to calculate the scale factor s and the angle of rotation of one coordinate system relative to another θ:

$$s = \sqrt{\alpha^2 + \beta^2} = 0.99979119290870,$$
$$\theta = arctan\frac{\beta}{\alpha} = -1°11'46.724"$$

Now according to the formulas (6) we calculate the shift parameters:

$$\begin{cases} X_0 = \bar{X} - \bar{U} \cdot \alpha + \bar{V} \cdot \beta = 62373.0296, \\ Y_0 = \bar{Y} - \bar{U} \cdot D - \bar{V} \cdot \alpha = 13891.4630. \end{cases}$$

Let us estimate the values of the residuals that the coordinates receive X и Y. To do this you can use the formula (3) and (7):

$$\begin{bmatrix} r_1^x \\ r_2^x \\ r_3^x \\ r_1^y \\ r_2^y \\ r_3^y \end{bmatrix} = \begin{bmatrix} 0.0056 \\ -0.0289 \\ 0.0233 \\ 0.0168 \\ 0.0206 \\ -0.0375 \end{bmatrix}$$

Now let's check the correctness of calculations of the transformation parameters. This control follows from our following reasoning. When solving the least squares problem, we are looking for the minimum of the function:

$$S = \sum r^2 = r^T \cdot r = r^T \cdot (A \cdot z - l) = r^T \cdot A \cdot z - r^T \cdot l$$

But

$$r^T \cdot A \cdot z = z^T \cdot A^T r = z^T \cdot A^T (A \cdot z - l) = z^T \cdot (A^T A \cdot z - A^T \cdot l) = z^T \cdot \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = 0.$$

From here

$$S = r^T \cdot r = -r^T \cdot l.$$

Last equality $r^T \cdot r = -r^T \cdot l$ can serve as a control for the correctness of solving the least squares problem. Note that most modern authors do not know and do not use this equality to control the solution.

For our problem, we find both of these sums:

$$\sum r^2 = 0.0035245568$$

$$\sum r \cdot l = -\,0.0035245574$$

As you can see, both sums coincide in absolute value and are opposite in sign. This indicates the correctness of the decision.

Let's evaluate the accuracy:

$$m = \sqrt{\frac{\sum r^2}{2 \cdot n}} = \sqrt{\frac{0.003524}{2 \cdot 3}} = 0.0242$$

Now we can write the system transition key U and V to X and Y:

$$\begin{cases} X = 62373.0296 + 0.99957326776067 \cdot U - (-0.0208737106442) \cdot V \\ Y = 13891.4630 + (-0.0208737106442) \cdot U + 0.99957326776067 \cdot V \end{cases}$$

Conclusions

Thus, in this article we examined the features of the 2-D Helmert transform problem. To do this, we looked at a specific example of a two-dimensional Helmert transform using a numerical example.

We solved the residuals equations using a pseudoinverse matrix. However, the condition number of the residuals equation matrix turned out to be more than half a million. Such a large value of the condition number indicates poor stability of the matrix A^+. Therefore, we reduced the condition number of the residuals equation matrix by introducing centered coordinates. In this case, the condition number of the matrix A turned out to be equal to one, which is the best for assessing the stability of the matrix.

Control of the solution is given through a comparison of the sum of squares of the residuals and the sum of the products of residuals by free terms in the residuals equation.

REFERENCES

1.Victor V. Ziborov (2014). Gross errors diagnostic in the task of Helmert coordinates transformation with the use of recurrent formula // Engineering Geodesy, Kiev: KNUCA - ISSUE 61, 2014 - P. 15-21. http://nbuv.gov.ua/UJRN/Ig_2014_61_5
2. https://wiki.gis.com/wiki/index.php/Helmert_transformation
3. Geodesy, www.absoluteastronomy.com/encyclopedia/G/Ge/Geodesy.htm
4. G.A.Watson (2006). Computing Helmert transformations//Journal of Computational and Applied Mathematics 197 (2006) 387 – 394 https://doi.org/10.1016/j.cam.2005.06.047
5. https://mathworld.wolfram.com/search/?query=helmert&x=0&y=0
6. https://www.mdpi.com/2220-9964/9/9/494