Application of Advanced Statistical Procedures for Adjustment of Measurement Results in Engineering Surveying

Abstract: Measurements in engineering surveying are aimed at determining the coordinates of the points of a geodetic control, spatially setting out a technical design of an engineering structure, determining the spatial coordinates of points (or their displacement) that represent an engineering structure, and identifying the displacement and deformation of a studied engineering structure. Provided that the aforementioned measurements are to represent the same engineering structure, such observation results should be settled (adjusted) in one calculation process. The application of the Gauss–Markov theorem for this adjustment using covariance matrix Cov(L) for observed values L is the classical approach for adjusting the results of surveying observations of various accuracy (taking into account accuracy weights).

Determining the displacements of points in the process of adjusting the results of periodic measurements, applying different methods of tying geodetic controls to national networks, and using various instruments and measurement methods result in the individual displacement components or coordinates of the observed points being determined with different accuracies. This circumstance forms the basis for the assumption that the estimated parameters (unknown values) should be random.

This paper will formulate the principles of estimation of Gauss–Markov models in which the estimated parameters (X) are random. For this purpose, methods for the prior definition of covariance matrix C_X for the estimated parameters will be provided, which will be used to determine the conditional covariance matrix of observation vector L and then to estimate the most probable values of the X parameters. Covariance matrix Cov(X) obtained as a result of this estimation will be used to define the limit values of the variances of these parameters.

Keywords: measurements in engineering surveying, Gauss–Markov model, diagonal covariance matrix

Received: 6 September 2017; accepted: 9 July 2018
1. Introduction

Engineering surveying uses the results of observations performed with a variety of instruments and measurement methods such as electronic total stations, precision levelers, satellite techniques and laser scanners, and precision instruments for short-distance measurements. The main purpose of these measurements is to determine the coordinates of the points representing the geodetic control or a studied engineering structure or to spatially set out the technical design of a structure.

Providing that the conducted measurements represent a selected engineering structure, the results of these observations should be adjusted in one calculation process. In the current surveying practice, the Gauss–Markov model is used to adjust the observation results, taking into account the diagonal covariance matrix for the observed values.

The Gauss–Markov model is also used to adjust the measurement results of multi-row controls, taking into account the apparent observational equations (pseudo-observations) for the coordinates of the reference points. For the pseudo-observations, the appropriate weights resulting from the accuracy of the coordinates of the analyzed points are assigned. Many authors of research studies in this field use a sequential adjustment of the measurement results; that is, adjustments carried out in many stages. Wiśniewski [1, 2] demonstrated that the formulated equations of the pseudo-observations for coordinates of points in sequential adjustment yield identical estimators as the adjustment of the whole geodetic network in one stage.

The issues related to the selection of the appropriate weights to adjust the observation results in geodetic control networks were the subject of the following studies: Baarda [3], Teunissen [4], Rao [5], Czaja [6], and Cross [7]. One of the methods for optimizing geodetic networks is a strategy for the balanced accuracy of observations developed by Kampmann [8] and Caspary [9]. Another proposal for the selection of observations and their weights was presented by Hekimoğlu [10] as well as by Kampmann and Krause [11]. The problem of selecting the appropriate weights for determining the coordinates of the points of geodetic networks is closely associated with the internal and external reliability of the network. The theory of network reliability is the subject of numerous scientific papers, and the precursors of these studies include Baarda [12, 13] and Pope [14]. Prószyński’s studies [15, 16] represent a significant accomplishment in this field as well. The issues related to the design of optimal horizontal and vertical control networks and their being tied to national geodetic networks are discussed in Dąbrowski [17, 18].

Engineering surveying uses horizontal and vertical control networks tied to national geodetic networks, periodic measurements of the networks of measurement points, various instruments and methods of observation, and various ways of stabilizing observation stands and measurement points. All of these observations
most frequently lead to the determination of the coordinates of individual points with different accuracies. This circumstance forms the basis for the assumption that the adjusted coordinates of the measurement points should be random, which means that a covariance matrix should be formulated a priori.

This paper will formulate the principles of estimation of the parameters of Gauss–Markov models and their variance applied to a network of geodetic points where the estimated parameters (coordinates of the points) will be random.

2. Theoretical Bases for Gauss–Markov model ($L, AX, H$) with Random Parameters

Horizontal and vertical angles, horizontal and spatial lengths, coordinates of 3-D points in a fixed reference system, height differences, and the displacement of selected points may be observed in the networks of measurement points used in engineering surveying. For each observed value of $\lambda$, the observational equation in the general form can be formulated as follows:

$$\delta_{\lambda} + d(\lambda) = \lambda_{\text{observ}} - \lambda_{\text{approx}}$$

(1)

where:

$\delta_{\lambda}$ – random deviation to observed value ($\lambda_{\text{observ}}$),

$d(\lambda)$ – the differential of a function describing the variability of analyzed component with respect to the coordinates of the points of the geodetic network defining that component,

$\lambda_{\text{approx}}$ – the approximate value of the analyzed component, determined subject to the approximate coordinates of the points of a geodetic network.

Let $L$ be a vector of random variables representing the differences between the observed values of the components in geodetic networks and their approximate values; i.e., $(\lambda_{\text{observ}} - \lambda_{\text{approx}})$. The average value of this vector can be described using fixed linear model $E(L) = AX$, where $X$ is the vector of the unknown parameters (adjustments to the approximated coordinates of the points). Matrix $A$ represents the matrix of coefficients defined by the values of the partial derivatives occurring in differential $d(\lambda)$. It is assumed that vector of unknown parameters $X$ also represents a random variable for which it is possible to determine a priori covariance matrix $C_X$.

Let matrix $H$ (whose inverse corresponds to weight matrix $P$) be the covariance matrix of observation $L$ with predetermined $X$; that is:

$$H = V(L/X)$$

(2)
Taking into account the above assumptions, the conditional covariance matrix of observation vector $L$ can be determined by the following dependence:

$$V(L) = E[V(L/X)] + V[E(L/X)] = H + V(AX) = H + ACXAT$$  \hspace{1cm} (3)

The estimation of the average value of vector $L$ will be performed using parameter estimators representing vector $X$ using covariance Matrix (3) and the method of least squares. For this purpose, square form $F$ will be formulated for random deviations (objective function) but by taking into account the conditional covariance matrix of observation vector $L$ for which the minimum relative to vector of unknowns $X$ will be sought; i.e.,

$$F = [(L - AX)^T(H + ACXAT)^{-1}(L - AX)] = \min$$  \hspace{1cm} (4)

The necessary condition for the minimum of Function (4) can be written in the following symbolic form:

$$\frac{\partial F}{\partial X} = 0$$  \hspace{1cm} (5)

After performing the differentiation of Function (4), a system of matrix equations is obtained, which satisfies Condition (5), which is the basis for the determination of the unbiased estimator of vector $\hat{X}$; i.e.,

$$\hat{X} = CXA^T(H + ACXAT)^{-1}L$$  \hspace{1cm} (6)

Further matrix transformations of Dependence (6) lead to an alternative formula for calculating the estimator of vector $X$ in the following form:

$$\hat{X} = (C^{-1}_X + A^T H^{-1} A)^{-1} A^T H^{-1} L$$  \hspace{1cm} (7)

In order to examine the effectiveness of this estimator, a full analysis of the variance and interval estimation is necessary.

The vector of random deviation $\delta$ to estimate linear model $AX$ is the difference between vector $L$ and its mean value $E(L) = A\hat{X}$; that is:

$$\delta = L - A\hat{X}$$  \hspace{1cm} (8)

The variance for the estimated model resulting from the mutual inconsistency of the survey results of the geodetic control network is defined by the following formula:

$$\hat{\sigma}_0^2 = \frac{\delta^T H^{-1} \delta}{n - u}$$  \hspace{1cm} (9)

where $n$ – the number of observed random variables (components in the geodetic control network), $u = R(A)$. 
The covariance matrix of the estimated vector of parameters $\hat{X}$ is defined by Variance (9) and the matrix contained in Formula (7); hence, it is expressed by the following formula:

$$\text{Cov}(\hat{X}) = \sigma_0^2(C^{-1}X + A^T H^{-1} A)^{-1}$$  \hspace{1cm} (10)

Components on the diagonal of Matrix (10) determine the variances of the individual estimated parameters, and their square root is the standard deviation $\sigma(X_i)$ of these parameters.

In order to determine the significance level of the values of the estimated parameters, it is necessary to estimate the limit value of their variance or standard deviation at a predetermined confidence level of $(1 - \alpha)$. The functional relationship that represents the estimated variance $\sigma^2(\hat{X}_i)$ of the analyzed parameter $(\hat{X}_i)$ and the tested variance $\sigma^2(X_i)$ of this parameter (taking into account $k = n - u$ degrees of freedom) is determined by the chi-square ($\chi^2$) and takes the following form:

$$\chi^2 = \frac{k \cdot \sigma^2(\hat{X}_i)}{\sigma^2(X_i)}$$  \hspace{1cm} (11)

Pearson demonstrated that the variability of the above relationship for a static test can be presented by a gamma function that, under the appropriate boundary conditions, describes the density distribution of the probability of random variable $\chi^2$ (known as the chi-square distribution in short). The chi-square distribution for different degrees of freedom $k = n - u$ takes on a different form (as illustrated in Figure 1).
The chi-square distribution can be defined by its quantiles \( \chi^2(\alpha; k) \), determined by the probability density function and significance index \( \alpha \). Quantile \( \chi^2(\alpha; k) \) is the length of the abscissa on axis 0\( \chi^2 \) that (from the entire field under the density function diagram with a surface of 1) cuts off the area of \( \alpha \) (as illustrated in Figure 2).

![Fig. 2. Quartiles of chi-square distribution](image)

The area marked in Figure 2 satisfies inequality \( \chi^2 > \chi^2(\alpha; k) \) that represents the probabilities of the values of \( (1 - \alpha) \), which is called the confidence level. This relationship can be expressed in the following analytical form:

\[
P[\chi^2 > \chi^2(\alpha; k)] = 1 - \alpha
\]

(12)

After substituting \( \chi^2 \) with Expression (11), the relationship of the estimated variance and tested variance was obtained in conjunction with distribution quantile \( \chi^2 \); i.e.:

\[
P = \left[ \frac{k \cdot \sigma^2(\hat{X})}{\sigma^2(X)} > \chi^2(\alpha; k) \right] = 1 - \alpha
\]

(13)

The above relationship will always occur if the expression in the square brackets is satisfied; i.e.:

\[
\frac{k \cdot \sigma^2(\hat{X})}{\sigma^2(X)} > \chi^2(\alpha; k)
\]

(14)

The transformation of the above inequality leads to the following condition for the tested variance, which is a limiting variance for a confidence level of \( (1 - \alpha) \):

\[
\sigma^2(X) \leq \frac{k \cdot \sigma^2(\hat{X})}{\chi^2(\alpha; k)}
\]

(15)
An interpretation of the above inequality is as follows: at a confidence level of 
\((1 – \alpha)\), the maximum value of the tested variance will always be less than or equal to 
the estimated variance multiplied by coefficient \(k/\chi^2(\alpha; k)\). For practical applications, 
the values of these coefficients were determined for the selected degrees of freedom 
(from 2 to 10) and the selected confidence levels (from 0.99 to 0.60) as illustrated in 
Table 1.

**Table 1.** Coefficients for determining limit value of variance or [limit standard deviation]  
for estimated parameters (coordinates of geodetic network points)

| \(k/(1 – \alpha)\) | 0.99    | 0.95    | 0.90    | 0.80    | 0.60    |
|---------------------|---------|---------|---------|---------|---------|
| \(k = 2\)           | 100.00  | 19.42   | 9.48    | 4.48    | 1.96    |
| \(k = 3\)           | 26.09   | 8.52    | 5.14    | 2.98    | 1.60    |
| \(k = 4\)           | 13.47   | 5.62    | 3.76    | 2.43    | 1.45    |
| \(k = 5\)           | 9.02    | 4.37    | 3.10    | 2.13    | 1.37    |
| \(k = 6\)           | 6.88    | 3.67    | 2.72    | 1.95    | 1.31    |
| \(k = 7\)           | 5.65    | 3.23    | 2.47    | 1.83    | 1.27    |
| \(k = 8\)           | 4.86    | 2.93    | 2.29    | 1.74    | 1.24    |
| \(k = 9\)           | 4.31    | 2.71    | 2.16    | 1.67    | 1.22    |
| \(k = 10\)          | 3.90    | 2.53    | 2.06    | 1.62    | 1.20    |

In a practical example, the coefficient values demonstrated in Table 1 can be 
interpreted as follows: if the horizontal geodetic network consisting of 7 points 
was measured with angles and side lengths with reference to a geodetic network 
of a higher level in the total number of \(n = 18\), then the value of coefficient \(k/\chi^2(\alpha; k)\) 
for a confidence level of \(\alpha\), and degrees of freedom \(k = 18 – 4 = 4\) is 3.76. This also 
means that the limit value of the standard deviation for the estimated coordinates 
of the points of the analyzed control is 1.9 times the value of the estimated standard 
deviation.

The determination and selection of the components of covariance matrix \(C_X\) 
for vector of unknown parameters \(X\) and conditional covariance matrices \(H\) for the 
observation vector directly influence the efficiency of the estimated vector of param-
eters \(\hat{X}\) and on its reliability as well.

3. **Applications of Gauss–Markov Model with Random Parameters**

The Gauss–Markov model (G-M) with random parameters can be used in the 
several issues related to engineering surveying.
These are:
- determination of point displacements based on periodic geodetic measurements,
- determination of geometric parameters of building structures and technical facilities for control and rectification purposes,
- adjustment of measurement results when establishing geodetic control networks that are tied to national networks,
- adjustment of measurement results performed with electronic total stations and GPS technology, taking into account a different stabilization of the geodetic network points.

This paper will formulate the principles for the prior defining of the components of covariance matrix $C_X$ for the first application area; i.e., the periodic surveys of geodetic control networks that are to form the basis for determining the displacements of the points representing the engineering structures.

If the geodetic control network is observed periodically, the purpose of each survey will be to determine the most likely coordinates of its points in a fixed reference system. The differences in the adjusted coordinates of the periodically observed points are determined by the components of their displacements. The determination of a local reference system occurs in the process of adjusting the periodic measurements. The adoption of a reference system for the adjustment of the observation results from the first periodic measurement is of particular importance. The approximate coordinates of the geodetic network points needed to compare the observational equations should always be determined in the local system based on the accurate measurements of the components in the analyzed networks.

The use of the G-M model with random parameters to adjust the results of the first measurement requires a suitable design of covariance matrix $C_X$ for the estimated adjustments to the approximated coordinates of the points. If all of the observed points of the geodetic network have an equally accurate stabilization (both for the stands and targets) and will be measured with the same instruments, then covariance matrix $C_{X1}$ should be diagonal, and its components should take the predetermined values for the variances of the individual coordinates of the points. After applying a variance analysis to the differential of the section length as the linear form of the differentials of the coordinates of the endpoints, the following formula for the variance of the section length is obtained:

$$\sigma^2(d) = 2(\sigma_w^2)$$ (16)

assuming that the variances for the individual coordinates of the points ($\sigma_w^2$) have the same value. If an electronic total station with nominal accuracy of 5 mm (for example) is used for the measurement of the geodetic control network, then this value should be used in Formula (16) for $\sigma^2(d) = 25 \text{ mm}^2$. 
Thus, the variance for the individual coordinates of the points will be at a level of

$$\sigma_w^2 = \frac{1}{2} \sigma^2(d) = 12.5 \text{ mm}^2$$  \hspace{1cm} (17)

Such a variance should become the components on the main diagonal of matrix $C_{XI}$ which will be used to adjust the horizontal geodetic control network in the first periodic measurement. In the case of the leveling network of points, the value of the variance of the benchmark height is set at the accuracy level of reading from the level staff.

Based on the adjustment of the results of the first periodic measurement, the most probable correction vector to the approximate coordinates of the points can be determined (as illustrated by Formula [7]) recorded in the following form:

$$\hat{X}_{I} = (C_{XI}^{-1} + A^T H^{-1} A)^{-1} A^T H^{-1} L_I$$  \hspace{1cm} (18)

and the covariance matrix for this correction vector (as illustrated in Formula [10]) is expressed as follows:

$$\text{Cov}(\hat{X}_I) = \sigma_0^2 (C_{XI}^{-1} + A^T H^{-1} A)^{-1}$$  \hspace{1cm} (19)

For the formulation of the observational equations for the second periodic measurement, the adjusted coordinates of the points from the first measurement should be used, while matrix $\text{Cov}(\hat{X}_I)$ obtained in the process of adjusting the first measurement will be covariance matrix $C_X$ for the adjustment of the second measurement. The adjustment of the observation results of the second periodic measurement in the form of equation

$$\hat{X}_{II} = [\text{Cov}(\hat{X}_I)]^{-1} + A^T H^{-1} A]^{-1} A^T H^{-1} L_{II}$$  \hspace{1cm} (20)

provides the most likely components of the displacements of the analyzed points; i.e.: 

$$\hat{X}_{II} = [u_{x1} \ u_{y1} \ ... \ u_{x_i} \ u_{y_i}]^T = \hat{U}_{II}$$  \hspace{1cm} (21)

The covariance matrix for such components of the point displacements is expressed by the following formula:

$$\text{Cov}(\hat{X}_{II}) = \sigma_{II}^2 [\text{Cov}(\hat{X}_I)]^{-1} + A^T H^{-1} A]^{-1}$$  \hspace{1cm} (22)

After using the coefficients contained in Table 1, the limit values of the variance will be defined for the determined displacement components. Those points whose displacement components exceed the limit values of their variances should be regarded as shifted points at the adopted confidence level.
4. Conclusions

The proposed algorithm for the estimation of the Gauss–Markov model with random parameters to adjust the results of the periodic measurements of geodetic control networks allows us to determine the most likely coordinates of the points and their displacement components as well as the standard deviations for these parameters.

The values of the coefficients contained in Table 1 (defined by chi-square distribution quantiles) make it possible to determine the limit values of the standard deviations for the determined components of the point displacements that form the basis for identifying the shifted points.

The analyses contained in this paper can be successfully used at the design stage of geodetic control networks that are periodically observed in terms of selecting the appropriate observation components and the accuracy of their measurement.

References

[1] Wiśniewski Z.: Algebra macierzy i statystyka matematyczna w rachunku wyrównawczym: teoria i zadania. Wydawnictwo Uniwersytetu Warmińsko-Mazurskiego, Olsztyn 2000.
[2] Wiśniewski Z.: Rachunek wyrównawczy w geodezji: (z przykładami). Wydawnictwo Uniwersytetu Warmińsko-Mazurskiego, Olsztyn 2005.
[3] Baarda W.: Statistical concepts in geodesy. Netherlands Geodetic Commission, Publications on Geodesy, New Series 2, No. 4, Delft 1967.
[4] Teunissen P.J.G.: Adjustment theory – an introduction. VSSD, Delft 2000.
[5] Rao C.R.: Modele liniowe statystyki matematycznej. PWN, Warszawa 1982.
[6] Czaja J.: Modele statystyczne w informacji o terenie. Skrypty Uczelniane – Akademia Górniczo-Hutnicza im. Stanisława Staszica, nr 1465, Wydawnictwa AGH, Kraków 1996.
[7] Cross P.A.: Numerical methods in network design. [in:] Grafarend E.W., Sanso F. (eds.), Optimization and design of geodetic networks, Springer, Berlin – Heidelberg – New York 1985, pp. 132–168.
[8] Kampmann G.: Robuste Deformationsanalyse mittels balancierter Ausgleichung. Allgemeine Vermessungs-Nachrichten: AVN. Zeitschrift für alle Bereiche der Geodäsie und Geoinformation, vol. 101, 1994, 1, pp. 8–18.
[9] Caspary W.: Anmerkungen zur balancierten Ausgleichung. Zeitschrift für Vermessungswesen, vol. 123, no. 8, 1998, pp. 271–273.
[10] Hekimoglu S.: Change of the diagonal elements of the hat matrix under changing weight and changing position of an observation. Zeitschrift für Vermessungswesen, vol. 123, no. 8, 1998, pp. 266–271.
Zastosowanie zaawansowanych procedur statystycznych do wyrównywania wyników pomiarów w geodezji inżynieryjnej

Streszczenie: Celem pomiarów w geodezji inżynieryjnej może być: wyznaczanie współrzędnych punktów osnowy realizacyjnej, wytyczenie w przestrzeni projektu technicznego obiektu inżynierskiego, wyznaczenie przestrzennych współrzędnych punktów lub ich przemieszczeń reprezentujących obiekt inżynierski oraz określenie przemieszczeń i odkształceń badanego obiektu inżynierskiego. Jeżeli wyżej wymienione pomiary odnoszą się do obiektu inżynierskiego, to takie wyniki obserwacji powinny być uzgadniane (wyrównywane) w jednym procesie obliczeniowym. Zastosowanie do tego wyrównania modeli Gaussa–Markowa z wykorzystaniem macierzy kowariancji Cov(L) dla wielkości obserwowanych L stanowi kluczowe postępowanie wyrównywania różnodokładnych wyników obserwacji geodezyjnych, z uwzględnieniem wag dokładności. Wyznaczanie przemieszczeń punktów w procesie wyrównywania wyników okresowych pomiarów, stosowanie różnych sposobów nawiązywania osnow realizacyjnych do sieci państwowych oraz wykorzystywanie różnych przyrządów i metody pomiaru – wszystko to powoduje, że poszczególne składowe przemieszczeń lub współrzędne obserwowanych punktów będą określane z różną dokładnością. Ta okoliczność jest podstawą założenia, że szacowane parametry (niewiadome) powinny mieć charakter losowy.
W artykule sformułowano zasady estymacji modeli Gaussa–Markowa, w których szacowane parametry $X$ mają charakter losowy. W tym celu podano sposoby określania a priori macierzy kowariancji $C_X$ dla estymowanych parametrów, która została wykorzystana do wyznaczenia macierzy kowariancji warunkowych wektora obserwacji $L$, a następnie do estymacji najbardziej prawdopodobnych wartości parametrów $\hat{X}$. Uzyskana w wyniku tej estymacji macierz kowariancji $\text{Cov}(X)$ została wykorzystana do ustalenia granicznych wartości wariancji tych parametrów.

Praktyczne zastosowanie proponowanego sposobu estymacji modelu G-M do wyznaczania pionowych przemieszczeń powierzchni osuwiska, dla parametrów losowych, zostało zilustrowane na przykładzie fragmentu niwelacyjnej sieci punktów.

**Słowa kluczowe:** pomiary w geodezji inżynieryjnej, model Gaussa–Markova, diagonalna macierz kowariancyjna