ESTYMACJA PARAMETRÓW LINIOWYCH MODELI

Czaja, J., Preweda, E.

Modelowanie danych przestrzennych

(pp. 1-12)

Warszawa, 2000

Materiały II Ogólnopolskiego Seminarium: *Modelowanie danych przestrzennych*, Warszawa, 2000

ESTYMACJA PARAMETRÓW LINIOWYCH MODELI
ESTIMATION PARAMETERS OF LINEAR MODELS

Józef Czaja, Edward Preweda

Akademia Górniczo-Hutnicza w Krakowie,
Wydział Geodezji Górniczej i Inżynierii Środowiska, Katedra Informacji o Terenie

Streszczenie

W pracy sformułowano kryterium doboru liczby obserwacji do estymacji parametrów modelu. Następnie dokonano analizy uogólnionego modelu liniowego postaci:

$$\begin{bmatrix} \mathbf{S} & \mathbf{A} \\ \mathbf{0} & \mathbf{B} \end{bmatrix} \begin{bmatrix} \boldsymbol{\delta} \\ \mathbf{X} \end{bmatrix} = \begin{bmatrix} \mathbf{U} \\ \mathbf{W} \end{bmatrix},$$

w którym poszczególne macierze mają następującą interpretację:

- $\boldsymbol{\delta}$ - wektor odchyłek losowych w modelu probabilistycznym ($n\times 1$),
- **X** - wektor parametrów modelu występujący w części probabilistycznej i części deterministycznej ($u\times 1$),
- **S** - macierz współczynników przy wektorze losowym $\boldsymbol{\delta}$ ($m\times n$),
- **A** - macierz współczynników przy wektorze parametrów w części probabilistycznej ($m\times u$),
- **B** - macierz współczynników przy wektorze parametrów w części deterministycznej ($r\times u$),
- **U** - wektor zredukowanych wyrazów wolnych w części probabilistycznej ($m\times 1$),
- **W** - wektor zredukowanych wyrazów wolnych w części deterministycznej ($r\times 1$).

Estymacji parametrów tego modelu dokonano MNK przy uwzględnieniu warunków Gaussa - Markowa dla formy kwadratowej zapisanej za pomocą funkcji Lagrange'a. Warunki konieczne na minimum funkcji Lagrange'a prowadzą do układu równań o macierzy kwadratowej. Rozwiązanie tego układu równań przedstawiono za pomocą *g*-odwrotności macierzy blokowej (3×3). Podano wzory na estymatory wektorów niewiadomych oraz estymatora wariancji resztowej i macierzy kowariancji wektorów niewiadomych.

W końcowej części pracy podano procedurę weryfikacji istotności estymowanego modelu przy uwzględnieniu statystyki F- Snedecora.

Abstract

The subject of the paper is the point estimation of the generalised linear model of the form:

$$\begin{bmatrix} \mathbf{S} & \mathbf{A} \\ \mathbf{0} & \mathbf{B} \end{bmatrix} \begin{bmatrix} \boldsymbol{\delta} \\ \mathbf{X} \end{bmatrix} = \begin{bmatrix} \mathbf{U} \\ \mathbf{W} \end{bmatrix},$$

where the individual matrixes have the following interpretation:

- $\boldsymbol{\delta}$ - *vector of the random deviations in the probability model ($n \times 1$),*
- **X** - *vector of the model parameters occurred both in the probability part and in the determined part ($u \times 1$),*
- **S** - *matrix of coefficients by the random vector* $\boldsymbol{\delta}$ *($m \times n$),*
- **A** - *matrix of coefficients by the parameters vector in the probability part ($m \times u$),*
- **B** - *matrix of coefficients by the parameters vector in the determined part ($r \times u$),*
- **U** - *vector of the reduced free terms in the probability part ($m \times 1$),*
- **W** - *vector of the reduced free terms in the determined part ($r \times 1$).*

The model estimation has been made by means of the LSM (Least Square Method) taking into account the Gauss-Markov conditions for the square form written by the Lagrange function. The necessary condition for the minimum of Lagrange function lead to the system of equations of square matrix. The solution of this system of equation has been shown by means of the block matrix g-inverse (3×3). The formulas for the estimators computations of the vector unknown as well as the residual variance estimators and covariance matrix of the vector unknown.

1. Uogólniona postać modelu liniowego

W ogólnym przypadku wielkości obserwowane można aproksymować za pomocą nieliniowych modeli regresji wielorakiej, w których wielkości obserwowane reprezentują zmienną objaśnianą (zależną), zaś parametry modelu stanowią zmienne objaśniające (niezależne. Modele takie występują pod ogólną nazwą stochastycznych modeli hybrydowych, czyli modeli zawierających różne warunki nałożone na estymowane parametry.

Kryterium określające liczbę obserwacji w modelowaniu może być przedział ufności dla nieznanej wartości przeciętnej lub wektora wartości przeciętnych (μ) i nieznanego odchylenia standardowego (σ) z próby losowej, czyli

$$\left|\mu - \hat{X}\right| = t(1-\frac{\alpha}{2},\ n-u)\frac{\hat{\sigma}}{\sqrt{n-u}}, \qquad (1.1)$$

gdzie:

$\hat{X}$ - oznacza wartość przeciętną zmiennej losowej liczoną z próby lub wektor wartości przeciętnych,

$\hat{\sigma}$ - oznacza odchylenie standardowe zmiennej losowej liczone z próby,

$t(1-\frac{\alpha}{2},\ n-u) = t_\alpha$ - oznacza kwantyl rozkładu Studenta o $(n-u)$ stopniach swobody dla poziomu ufności $(1-\alpha)$, przy czym α oznacza prawdopodobieństwo wystąpienia *błędu pierwszego rodzaju*, czyli prawdopodobieństwa odrzucenia weryfikowanej hipotezy statystycznej dotyczącej wartości przeciętnej, podczas gdy jest ona w rzeczywistości prawdziwa. Wartość tego kwantyla określa się z tablic rozkładu Studenta.

Po przekształceniu wzoru (1.1) otrzymuje się formułę na obliczenie minimalnej ilości elementów nadliczbowych próby do estymacji modelu, czyli

$$n-u = (t_\alpha)^2 \frac{(\hat{\sigma})^2}{(\mu - \hat{X})^2}, \tag{1.2}$$

która zależna jest od kwantyla rozkładu Studenta (t_α) oraz od odchylenia standardowego w próbie $\hat{\sigma}$ i niedokładności szacowania wartości przeciętnej $(|\mu - \hat{X}|)$.

Jeżeli przyjmiemy założenie, że niedokładność szacowania wartości przeciętnej będzie na poziomie odchylenia standardowego zmiennej losowej w próbie, wtedy wzór (1.2) upraszcza się do następującej postaci:

$$n-u = (t_\alpha)^2 \quad \rightarrow \quad t_\alpha = \sqrt{n-u}, \tag{1.3}$$

w której kwantyl rozkładu Studenta (t_α) zależy od dwóch parametrów, to znaczy od liczby stopni swobody $(n-u)$ oraz od poziomu ufności $(1-\alpha)$. Ustalając poziom ufności $(1-\alpha) = 0.99$ można z równania (1.3) jednoznacznie określić takie $(n-u)$, dla którego kwantyl rozkładu Studenta $t(0.99,\ n-u) = \sqrt{n-u}$. Rozwiązanie tego równania metodą kolejnych przybliżeń daje wartość $n-u = 10,$ czyli $n = u+10$.

Z powyższej analizy wynika bardzo ważny wniosek praktyczny, że przy przyjętych założeniach, iż niedokładność szacowania wektora wartości przeciętnych będzie na poziomie odchylenia standardowego zmiennej losowej w próbie, minimalna liczba obserwacji będąca podstawą estymacji modelu nie powinna być mniejsza od $u+10$.

Najtrudniejszym problemem w modelowaniu wielkości obserwowanych jest dobieranie matematycznej postaci funkcji, która będzie najdokładniej opisywała badane zjawisko. Miarą tej dokładności powinien być estymator współczynnika regresji wielokrotnej (R), za pomocą którego definiuje się statystykę F-Snedecora w następującej postaci:

$$F_S = \frac{R^2}{1-R^2}\frac{n}{n-u} \tag{1.4}$$

Każdy układ równań obserwacyjnych albo aproksymacyjnych a także układ warunków na wielkości obserwowane może być sprowadzony do postaci liniowej. W tym celu wykorzystuje się rozwinięcia dowolnej postaci funkcji w szereg Taylora. Zatem każdy przypadek estymacji parametrów modelu i wartości stacjonarnych można sprowadzić do postaci modelu liniowego.

W dotychczasowych zastosowaniach wykorzystuje się różne procedury estymacji liniowej, które wynikają z postaci układów równań wiążących obserwacje. Istotne znaczenie poznawcze i utylitarne będzie miało sformułowanie uogólnionego modelu estymacji liniowej, którego szczególne przypadki będą odpowiadały dotychczas stosowanym algorytmom w modelowaniu wyników obserwacji.

Niech dowolny układ równań liniowych (nadliczbowy lub uwarunkowany) będzie zapisany w postaci macierzowej

$$\mathbf{M}\,\mathbf{Z} = \mathbf{U} \tag{1.5}$$

gdzie:

- **M** - oznacza macierz współczynników, stanowiących wartości pierwszych pochodnych dla przybliżonych wartości estymowanych parametrów modelu, o wymiarach $(\mathrm{m} \times \mathrm{n} + \mathrm{u})$,
- **Z** - oznacza macierz niewiadomych estymatorów, stanowiących odchyłki losowe lub przyrosty do przybliżonych wartości estymowanych parametrów modelu, o wymiarach $(\mathrm{n} + \mathrm{u} \times 1)$,
- **U** - oznacza macierz zredukowanych wyrazów wolnych dla poszczególnych równań za pomocą zaobserwowanych wartości zmiennych losowych lub przybliżonych wartości estymowanych parametrów modelu, o wymiarach $(\mathrm{m}\times1)$.

Rozwiązanie układu równań (1.5) może być formułowane dla różnych warunków brzegowych, wynikających ze specyfiki rozważanego zagadnienia. Z kolei różne warunki brzegowe wymagają pionowego lub poziomego podziału układu równań (1.5) i prowadzą do układów równań blokowych.

Dla pionowego podziału układu równań (1.5) można na część niewiadomych $\boldsymbol{\delta}$ narzucić warunek Gaussa - Markowa, a pozostałą część niewiadomych **X** potraktować jako parametry modelu. Wtedy taki układ równań może być zapisany w postaci blokowej

$$\left[\mathbf{S}\,\mathbf{A}\right]\begin{bmatrix}\boldsymbol{\delta}\\ \mathbf{X}\end{bmatrix} = \mathbf{U} \tag{1.6}$$

co jest równoważne układowi równań

$$\mathbf{S}\,\boldsymbol{\delta} + \mathbf{A}\,\mathbf{X} = \mathbf{U} \tag{1.6a}$$

przy czym $\boldsymbol{\delta}$ oznacza wektor odchyłek losowych, zaś $\mathbf{X}$ oznacza wektor parametrów analizowanego modelu.

Dla poziomego podziału układu równań (1.5) wykorzystamy postać (1.6), w której na wartości parametrów modelu narzucamy restrykcje (warunki) postaci

$$\mathbf{B}\,\mathbf{X} = \mathbf{W} \tag{1.7}$$

Uwzględniając powyższy warunek, układ równań (1.6) może być zapisany w formie blokowej

$$\begin{bmatrix} \mathbf{S} & \mathbf{A} \\ \mathbf{0} & \mathbf{B} \end{bmatrix} \begin{bmatrix} \boldsymbol{\delta} \\ \mathbf{X} \end{bmatrix} = \begin{bmatrix} \mathbf{U} \\ \mathbf{W} \end{bmatrix} \tag{1.8}$$

Podział poziomy układu równań (1.5) zawarty w postaci (1.8) może wynikać z warunków brzegowych, w których górna część równań powinna spełniać model probabilistyczny, zaś dolna część równań powinna spełniać model deterministyczny.

Macierze występujące w równaniu (1.8) posiadają następujące oznaczenia:

- **S** - macierz współczynników przy wektorze losowym $\boldsymbol{\delta}$,
- **A** - macierz współczynników przy estymowanym wektorze parametrów modelu (**X**) w części probabilistycznej,
- $\boldsymbol{\delta}$ - wektor odchyłek losowych w modelu probabilistycznym,
- **X** - wektor niewiadomych, czyli estymowanych parametrów (przyrostów do przybliżonych wartości estymowanych parametrów), występująca w modelu probabilistycznym i w modelu deterministycznym,
- **B** - macierz współczynników przy estymowanym wektorze parametrów **X** w modelu deterministycznym,
- **U** - macierz zredukowanych wyrazów wolnych w modelu probabilistycznym,
- **W** - macierz zredukowanych wyrazów wolnych w modelu deterministycznym,

2. Estymacja uogólnionego modelu liniowego

Zgodnie z zasadą MNK estymatory nieobciążone układu równań (1.8) muszą spełniać warunki Gaussa - Markowa wynikające z formy kwadratowej zapisanej za pomocą funkcji Lagrange'a.

Dla układu równań (1.8) zapisanego w formie macierzy blokowych funkcja Lagrange'a będzie następującej postaci:

$$\Phi = \boldsymbol{\delta}^{\mathrm{T}}\mathbf{P}\,\boldsymbol{\delta} + 2\mathbf{K}^{\mathrm{T}}\left(\mathbf{S}\,\boldsymbol{\delta} + \mathbf{A}\,\mathbf{X} - \mathbf{U}\right) + 2\mathbf{J}^{\mathrm{T}}\left(\mathbf{B}\,\mathbf{X} - \mathbf{W}\right) = \text{minimum} \tag{2.1}$$

przy czym **P** stanowi macierz wagową, zaś **K** i **J** oznaczają jednokolumnowe macierze współczynników Lagrange'a.

W powyższej formie kwadratowej występuje wektor losowy $\boldsymbol{\delta}$, który związany jest ze zmienną objaśniającą, czyli z obserwacjami $\mathbf{L}$ i ich wartościami modelowymi $\hat{\mathbf{L}}$ następującym równaniem:

$$\delta = \mathbf{L} - \hat{\mathbf{L}} \tag{2.2}$$

Zgodnie z teorią form kwadratowych minimum funkcji Lagrange'a będzie spełnione, gdy pierwsze pochodne będą równe wektorom zerowym, zaś macierze wynikające z drugich pochodnych powinny być dodatnio określone.

Warunki konieczne na minimum funkcji Lagrange'a (2.1) stanowią wyrażenia

$$\begin{aligned}
\frac{\partial \Phi}{\partial \boldsymbol{\delta}} &= 2\mathbf{P}\,\boldsymbol{\delta} + 2\mathbf{S}^{\mathrm{T}}\mathbf{K} = \mathbf{0} \\
\frac{\partial \Phi}{\partial \mathbf{X}} &= 2\mathbf{A}^{\mathrm{T}}\mathbf{K} + 2\mathbf{B}^{\mathrm{T}}\mathbf{J} = \mathbf{0} \\
\frac{\partial \Phi}{\partial \mathbf{K}} &= 2\left(\mathbf{S}\boldsymbol{\delta} + \mathbf{A}\,\mathbf{X} - \mathbf{U}\right) = \mathbf{0} \\
\frac{\partial \Phi}{\partial \mathbf{J}} &= 2\left(\mathbf{B}\,\mathbf{X} - \mathbf{W}\right) = \mathbf{0}
\end{aligned} \tag{2.3}$$

co prowadzi do układu czterech równań macierzowych następującej postaci:

$$\begin{aligned}
&\mathbf{P}\,\boldsymbol{\delta} + \mathbf{S}^{\mathrm{T}}\mathbf{K} = \mathbf{0} \\
&\mathbf{A}^{\mathrm{T}}\mathbf{K} + \mathbf{B}^{\mathrm{T}}\mathbf{J} = \mathbf{0} \\
&\mathbf{S}\,\boldsymbol{\delta} + \mathbf{A}\,\mathbf{X} = \mathbf{U} \\
&\mathbf{B}\,\mathbf{X} = \mathbf{W}
\end{aligned} \tag{2.4}$$

Uwzględniając zależność z pierwszego równania

$$\delta = -\mathbf{P}^{-1}\mathbf{S}^{T}\mathbf{K}\,, \tag{2.5}$$

układ równań (2.4) można zapisać w następującej formie:

$$\begin{array}{ccccccc}
-\mathbf{S}\,\mathbf{P}^{-1}\mathbf{S}^{\mathrm{T}}\mathbf{K} & + & \mathbf{0} & + & \mathbf{A}\,\mathbf{X} & = & \mathbf{U} \\
\mathbf{0} & + & \mathbf{0} & + & \mathbf{B}\,\mathbf{X} & = & \mathbf{W} \\
\mathbf{A}^{\mathrm{T}}\mathbf{K} & + & \mathbf{B}^{\mathrm{T}}\mathbf{J} & + & \mathbf{0} & = & \mathbf{0}
\end{array} \tag{2.6}$$

Układ równań (2.6) zapisany za pomocą macierzy blokowych przyjmuje następującą postać:

$$\begin{bmatrix} -\mathbf{S}\,\mathbf{P}^{-1}\mathbf{S}^{\mathrm{T}} & \mathbf{0} & \mathbf{A} \\ \mathbf{0} & \mathbf{0} & \mathbf{B} \\ \mathbf{A}^{\mathrm{T}} & \mathbf{B}^{\mathrm{T}} & \mathbf{0} \end{bmatrix} \times \begin{bmatrix} \mathbf{K} \\ \mathbf{J} \\ \mathbf{X} \end{bmatrix} = \begin{bmatrix} \mathbf{U} \\ \mathbf{W} \\ \mathbf{0} \end{bmatrix} \tag{2.7}$$

Warunki dostateczne (dopełniające) minimum funkcji Lagrange'a będą realizowane przy rozwiązywaniu układu równań (2.7), doprowadzając macierze współczynników przy niewiadomych do takiej postaci, w której będą dodatnio określone.

Układ równań (2.7) sformułowany według MNK jest niesprzeczny, gdyż liczba równań odpowiada liczbie niewiadomych, zatem zawsze istnieje ich rozwiązanie, ale przy uwzględnieniu uogólnionej odwrotności (g-odwrotności) macierzy ujmującej współczynniki tych równań.

W uogólnionym modelu (2.7) występuje macierz blokowa o dziewięciu podmacierzach, zatem obliczenie dla niej g-odwrotności wymaga odpowiedniego postępowania. Wprowadzając symboliczne oznaczenie g-odwrotności macierzy blokowej można zapisać następującą formułę

$$\begin{bmatrix} -\mathbf{S}\mathbf{P}^{-1}\mathbf{S}^{\mathrm{T}} & \mathbf{0} & \mathbf{A} \\ \mathbf{0} & \mathbf{0} & \mathbf{B} \\ \mathbf{A}^{\mathrm{T}} & \mathbf{B}^{\mathrm{T}} & \mathbf{0} \end{bmatrix} \times \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_2 & \mathbf{C}_3 \\ \mathbf{C}_2^{\mathrm{T}} & \mathbf{C}_4 & \mathbf{C}_5 \\ \mathbf{C}_3^{\mathrm{T}} & \mathbf{C}_5^{\mathrm{T}} & \mathbf{C}_6 \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{\mathrm{h}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{\mathrm{h}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{\mathrm{h}} \end{bmatrix} \tag{2.8}$$

przy czym I_h oznacza jednostkową macierz hermitowską.

Po zrealizowaniu wzoru (2.8) przy oznaczeniu

$$-\mathbf{S}\mathbf{P}^{-1}\mathbf{S}^{\mathrm{T}} = \mathbf{N}$$

otrzymujemy dziewięć układów równań macierzowych, ale o sześciu macierzach niewiadomych, czyli

$$\begin{array}{ccccccc} \mathbf{N}\mathbf{C}_1 & + & \mathbf{0} & + & \mathbf{A}\mathbf{C}_3^{\mathrm{T}} & = & \mathbf{I}_{\mathrm{h}} \\ \mathbf{0} & + & \mathbf{0} & + & \mathbf{B}\mathbf{C}_3^{\mathrm{T}} & = & \mathbf{0} \\ \mathbf{A}^{\mathrm{T}}\mathbf{C}_1 & + & \mathbf{B}^{\mathrm{T}}\mathbf{C}_2^{\mathrm{T}} & + & \mathbf{0} & = & \mathbf{0} \\ \mathbf{N}\mathbf{C}_2 & + & \mathbf{0} & + & \mathbf{A}\mathbf{C}_5^{\mathrm{T}} & = & \mathbf{0} \\ \mathbf{0} & + & \mathbf{0} & + & \mathbf{B}\mathbf{C}_5^{\mathrm{T}} & = & \mathbf{I}_{\mathrm{h}} \\ \mathbf{A}^{\mathrm{T}}\mathbf{C}_2 & + & \mathbf{B}^{\mathrm{T}}\mathbf{C}_4 & + & \mathbf{0} & = & \mathbf{0} \\ \mathbf{N}\mathbf{C}_3 & + & \mathbf{0} & + & \mathbf{A}\mathbf{C}_6 & = & \mathbf{0} \\ \mathbf{0} & + & \mathbf{0} & + & \mathbf{B}\mathbf{C}_6 & = & \mathbf{0} \\ \mathbf{A}^{\mathrm{T}}\mathbf{C}_3 & + & \mathbf{B}^{\mathrm{T}}\mathbf{C}_5 & + & \mathbf{0} & = & \mathbf{I}_{\mathrm{h}} \end{array} \tag{2.9}$$

Układ równań (2.8) jest symetryczny, stąd jego rozwiązanie będzie jednoznaczne. Wykonując szereg przekształceń macierzowych otrzymamy:

$$
\begin{aligned}
\mathbf{C}_6 &= -\left(\mathbf{A}^{\mathrm{T}}\mathbf{N}^{-}\mathbf{A}\right)^{-} + \left(\mathbf{A}^{\mathrm{T}}\mathbf{N}^{-}\mathbf{A}\right)^{-}\mathbf{B}^{-\mathrm{T}}\left[\mathbf{B}\left(\mathbf{A}^{\mathrm{T}}\mathbf{N}^{-}\mathbf{A}\right)^{-}\mathbf{B}^{\mathrm{T}}\right]^{-}\mathbf{B}\left(\mathbf{A}^{\mathrm{T}}\mathbf{N}^{-}\mathbf{A}\right)^{-},\\
\mathbf{C}_5 &= \left[\mathbf{B}\left(\mathbf{A}^{\mathrm{T}}\mathbf{N}^{-}\mathbf{A}\right)^{-}\mathbf{B}^{\mathrm{T}}\right]^{-}\mathbf{B}\left(\mathbf{A}^{\mathrm{T}}\mathbf{N}^{-}\mathbf{A}\right)^{-},\\
\mathbf{C}_4 &= \left[\mathbf{B}\left(\mathbf{A}^{\mathrm{T}}\mathbf{N}^{-}\mathbf{A}\right)^{-}\mathbf{B}^{\mathrm{T}}\right]^{-},\\
\mathbf{C}_3 &= -\mathbf{N}^{-}\mathbf{A}\,\mathbf{C}_6,\\
\mathbf{C}_2 &= -\mathbf{N}^{-}\mathbf{A}\,\mathbf{C}_5^{\mathrm{T}},\\
\mathbf{C}_1 &= \mathbf{N}^{-} + \mathbf{N}^{-}\mathbf{A}\,\mathbf{C}_6\mathbf{A}^{\mathrm{T}}\mathbf{N}^{-}
\end{aligned}
\tag{2.10}
$$

Uwzględniając najpierw wyrażenie na macierz $\mathbf{C}_4$, wzory (2.10) można zapisać w postaci

$$
\begin{aligned}
\mathbf{C}_4 &= \left[\mathbf{B}\left(\mathbf{A}^{\mathrm{T}}\mathbf{N}^{-}\mathbf{A}\right)^{-}\mathbf{B}^{\mathrm{T}}\right]^{-},\\
\mathbf{C}_5 &= \mathbf{C}_4\mathbf{B}\left(\mathbf{A}^{\mathrm{T}}\mathbf{N}^{-}\mathbf{A}\right)^{-},\\
\mathbf{C}_6 &= -\left(\mathbf{A}^{\mathrm{T}}\mathbf{N}^{-}\mathbf{A}\right)^{-} + \left(\mathbf{A}^{\mathrm{T}}\mathbf{N}^{-}\mathbf{A}\right)^{-}\mathbf{B}^{\mathrm{T}}\mathbf{C}_5,\\
\mathbf{C}_3 &= -\mathbf{N}^{-}\mathbf{A}\,\mathbf{C}_6,\\
\mathbf{C}_2 &= -\mathbf{N}^{-}\mathbf{A}\,\mathbf{C}_5^{\mathrm{T}},\\
\mathbf{C}_1 &= \mathbf{N}^{-} + \mathbf{N}^{-}\mathbf{A}\,\mathbf{C}_6\mathbf{A}^{\mathrm{T}}\mathbf{N}^{-}.
\end{aligned}
\tag{2.11}
$$

Zdefiniowane podmacierze (2.11) spełniają układ równań (2.9), zatem można je wykorzystać do rozwiązania układu równań (2.7).

Nieobciążony estymator wektorów parametrów modelu (1.8) można wyrazić iloczynem następujących macierzy blokowych:

$$
\begin{bmatrix} \hat{\mathbf{K}} \\ \hat{\mathbf{J}} \\ \hat{\mathbf{X}} \end{bmatrix} = \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_2 & \mathbf{C}_3 \\ \mathbf{C}_2^T & \mathbf{C}_4 & \mathbf{C}_5 \\ \mathbf{C}_3^T & \mathbf{C}_5^T & \mathbf{C}_6 \end{bmatrix} \begin{bmatrix} \mathbf{U} \\ \mathbf{W} \\ \mathbf{0} \end{bmatrix}
\tag{2.12}
$$

czyli

$$
\hat{\mathbf{K}} = \mathbf{C}_1\mathbf{U} + \mathbf{C}_2\mathbf{W}, \tag{2.13}
$$

$$
\hat{\mathbf{J}} = \mathbf{C}_2^T\mathbf{U} + \mathbf{C}_4\mathbf{W}, \tag{2.14}
$$

$$
\hat{\mathbf{X}} = \mathbf{C}_3^T\mathbf{U} + \mathbf{C}_5^T\mathbf{W} \tag{2.15}
$$

przy czym macierze $\mathbf{C}_{\mathrm{i}}$ mogą być bezpośrednio określone za pomocą wzorów (2.11).

Estymator wektora losowego δ obliczymy według formuły (2.5), czyli

$$\delta = -\mathbf{P}\mathbf{S}^T\hat{\mathbf{K}}\,, \tag{2.16}$$

zaś nieobciążony estymator wariancji resztowej wyrazi się wzorem

$$\hat{\sigma}_0{}^2 = \frac{\hat{\delta}^T\mathbf{P}\delta}{R\left(\mathbf{S}\mathbf{P}^{-1}\mathbf{S}^T\right) - R\left(\mathbf{A}^T \middle| \mathbf{B}^T\right)}, \tag{2.17}$$

Macierz kowariancji dla wektora $\hat{\mathbf{X}}$ będzie wyrażona za pomocą macierzy $\mathbf{C}_6$, zdefiniowanej zależnościami (2.11), czyli

$$Cov\left(\hat{\mathbf{X}}\right) = \hat{\sigma}^2\mathbf{C}_6, \tag{2.18}$$

Macierz kowariancji dla wektora $\hat{\delta}$ można wyznaczyć według następującego wzoru:

$$Cov\left(\hat{\delta}\right) = -\hat{\sigma}^2\mathbf{P}^{-1}\mathbf{S}^T\mathbf{C}_1\mathbf{S}\mathbf{P}^{-1}, \tag{2.19}$$

przy czym macierz $\mathbf{C}_1$ jest określona zależnościami (2.11).

Macierz kowariancji dla wektora $\hat{\mathbf{L}}$ będzie postaci

$$Cov\left(\hat{\mathbf{L}}\right) = \hat{\sigma}^2\mathbf{P}^{-1} - Cov\left(\hat{\delta}\right). \tag{2.20}$$

Rozwiązanie uogólnionego modelu liniowego (2.7) ilustruje jednolitą teorię estymacji modeli liniowych, które można wykorzystać do opisu zjawisk przyrodniczych lub społecznych.

Na podstawie przedstawionego rozwiązania układu równań (2.7) można analizować szczególne przypadki modeli liniowych, które wynikają z warunków obserwacji i charakteru badanego zjawiska. Dla przykładu rozpatrzymy trzy rodzaje takich modeli.

Wieloparametrowy model $(\mathbf{L}, \mathbf{AX}, \sigma^2\mathbf{G})$ Gaussa-Markowa

Jeżeli w uogólnionym modelu (2.7) przyjmiemy następujące założenia:

$$\begin{aligned} &\mathbf{S} = \mathbf{I} \\ &\mathbf{U} = \mathbf{L} \\ &\mathbf{B} = \mathbf{0} \\ &\mathbf{W} = \mathbf{0} \\ &\det(\mathbf{A}^\mathbf{T}\mathbf{P}\mathbf{A}) \neq \mathbf{0} \end{aligned} \tag{2.21}$$

wtedy estymatory dla wektorów niewiadomych i ich kowariancji przyjmują następującą postać:

$$\begin{aligned} &\hat{\mathbf{X}} = (\mathbf{A}^\mathbf{T}\mathbf{P}\mathbf{A})^{-1}\mathbf{A}^\mathbf{T}\mathbf{P}\mathbf{L}, \\ &\hat{\delta} = \mathbf{L} - \mathbf{A}\hat{\mathbf{X}} \end{aligned} \tag{2.22}$$

$$\begin{aligned} &Cov(\hat{\mathbf{X}}) = \hat{\sigma}_\mathbf{0}^\mathbf{2}(\mathbf{A}^T\mathbf{P}\mathbf{A})^{-1} \\ &Cov(\hat{\delta}) = \hat{\sigma}_0^2(\mathbf{P}^{-1} - \mathbf{A}(\mathbf{A}^\mathbf{T}\mathbf{P}\mathbf{A})^{-1}\mathbf{A}^\mathbf{T}) \end{aligned} \tag{2.23}$$

Wyżej rozpatrywany model w tradycyjnym rachunku wyrównawczym nosi nazwę metody wyrównywania skorelowanych obserwacji pośredniczących.

Uwarunkowany model $(\mathbf{L}, \mathbf{IX}, \sigma^2\mathbf{G}, \mathbf{S}\delta - \mathbf{U} = \mathbf{0})$ Gaussa-Markowa

Jeżeli w uogólnionym modelu (2.7) przyjmiemy następujące założenia:

$$\begin{aligned} \mathbf{A} &= \mathbf{0} \\ \mathbf{B} &= \mathbf{0} \\ \mathbf{W} &= \mathbf{0} \end{aligned} \tag{2.24}$$

wtedy elementy g-odwrotności macierzy blokowej przyjmują następującą postać:

$$\begin{aligned} \mathbf{C_3} &= \mathbf{0} \\ \mathbf{C_2} &= \mathbf{0} \\ \mathbf{C_1} &= -(\mathbf{SP^{-1}S^T})^{-1} \end{aligned} \tag{2.25}$$

skąd estymatory wektorów niewiadomych i ich kowariancji określa się według następujących wzorów:

$$\begin{aligned} \hat{\mathbf{K}} &= \mathbf{C}_1\mathbf{U} \\ \hat{\delta} &= -\mathbf{P}^{-1}\mathbf{S^T}\mathbf{C}_1\mathbf{U} \\ Cov(\hat{\delta}) &= \hat{\sigma}_0^2\mathbf{P}^{-1}\mathbf{S^T}\mathbf{C}_1\mathbf{SP}^{-1} \end{aligned} \tag{2.26}$$

Wyżej rozpatrywany model w tradycyjnym rachunku wyrównawczym nosi nazwę metody wyrównywania skorelowanych obserwacji uwarunkowanych.

Wieloparametrowy model $(\mathbf{L}, \mathbf{AX}, \sigma^2\mathbf{G}, \mathbf{BX} - \mathbf{W} = \mathbf{0})$ z restrykcjami

Jeżeli w uogólnionym modelu (2.7) przyjmiemy następujące oznaczenia:

$$\begin{aligned} \mathbf{S} &= \mathbf{I} \\ \mathbf{U} &= \mathbf{L} \end{aligned} \tag{2.27}$$

oraz założymy, że macierze symetryczne występujące we wzorach (2.10) będą bez defektu, wtedy estymatory wektorów niewiadomych będą postaci:

$$\begin{aligned} \hat{\mathbf{X}} = \{(\mathbf{A^TPA})^{-1} - (\mathbf{A^TPA})^{-1}\mathbf{B^T}[\mathbf{B}(\mathbf{A}^T\mathbf{PA})^{-1}\mathbf{B^T}]^{-1}\mathbf{B}(\mathbf{A^TPA})^{-1}\}\,\mathbf{A^TPL} + \\ + \{[(\mathbf{A^TPA})^{-1}\mathbf{B^T}[\mathbf{B}(\mathbf{A^TPA})^{-1}\mathbf{B^T}]^{-1}\}\,\mathbf{W} \end{aligned} \tag{2.28}$$

$$\hat{\delta} = \mathbf{L} - \mathbf{A}\hat{\mathbf{X}} \tag{2.29}$$

natomiast macierze kowariancji przyjmują postać

$$\begin{aligned} Cov(\hat{\mathbf{X}}) &= \hat{\sigma}_\mathbf{0}^2\mathbf{C_6} \\ Cov(\hat{\delta}) &= \hat{\sigma}_0^2(\mathbf{P}^{-1} - \mathbf{AC_6A^T}) \end{aligned} \tag{2.30}$$

Wyżej rozpatrywany model w tradycyjnym rachunku wyrównawczym nosi nazwę metody wyrównania obserwacji pośredniczących z warunkami na niewiadome.

3. Ocena wiarygodności estymowanych modeli

Miarą zaufania do estymowanego modelu jest kwadrat współczynnika korelacji wielokrotnej, który wyraża się następującym wzorem:

$$R^2 = 1 - \frac{\delta^T \mathbf{P} \delta}{(\mathbf{L} - \mathbf{L}_{\mathbf{ŚR}})^T \mathbf{P} (\mathbf{L} - \mathbf{L}_{\mathbf{ŚR}})} \tag{3.1}$$

gdzie:

δ - wektor odchyłek losowych zdefiniowany ogólnym wzorem (2.16),

$\mathbf{P}$ - macierz wag dokładności,

$\mathbf{L}$ - wektor obserwacji (zmiennej niezależnej),

$\mathbf{L}_{\mathbf{ŚR}}$ - wektor utworzony z wartości przeciętnej (średniej) zmiennej objaśniającej.

Miarą dopasowania estymowanego modelu do wyników eksperymentu (pomiaru) jest statystyka F – Snedecora, która może być zapisana według następującej formuły:

$$F_S = \frac{R^2}{1 - R^2} \frac{n}{n - u} \tag{3.2}$$

Pierwszy ułamek w tym wzorze określa stosunek sumy kwadratów odchyłek losowych z części wyjaśnionej przez model do sumy kwadrtatów odchyłek losowych z części niewyjaśnionej przez model, zaś drugi ułamek określa stopnie swobody.

Na ustalonym poziomie istotności α można weryfikować hipotezę statystyczną, czy estymowany model w sposób zadawalający wyjaśnia (aproksymuje) analizowane zjawisko. W tym celu, dla ustalonego poziomu ufności $(1-\alpha)$ oraz $(n-u)$ i (n) stopni swobody, należy określić kwantyl $F(1-\alpha,\ n-u,\ n)$ rozkładu F – Snedecora.

Jeżeli nierówność

$$F_S > F(1-\alpha,\ n-u,\ n) \tag{3.3}$$

dla estymowanego modelu jest spełniona, to oznacza, że nie ma podstaw do odrzucenia weryfikowanej hipotezy. Na podstawie tej weryfikacji można wnioskować, że na przyjętym poziomie ufności estymowany model w sposób zadawalający aproksymuje (reprezentuje) analizowane zjawisko.

Jeżeli nierówność (3.3) dla estymowanego modelu nie będzie spełniona, to oznacza, że analizowaną hipotezę należy odrzucić na korzyść hipotezy alternatywnej. Na tej podstawie będzie można wnioskować, że na przyjętym poziomie ufności analizowany model w sposób niezadawalający reprezentuje badane zjawisko.

W wielu zagadnieniach dotyczących informacji o terenie a'priori trudno ustalić, jaki model będzie najwięcej wyjaśniał zmienność analizowanego zjawiska oraz, które zmienne powinny mieć charakter losowy, a które zmienne powinny mieć charakter deterministyczny. Wstępna estymacja takich modeli może pomóc w rozdzieleniu tych zmiennych, a następnie można dokonać podziału takich równań na część probabilistyczną i część deterministyczną.

Estymacja układu równań (1.8), przy wykorzystaniu g-odwrotności macierzy blokowych, daje jednoznaczne i przejrzyste procedury szacowania macierzy kowariancji dla wektorów niewiadomych i ich funkcji. Szacowanie to minimalizuje wariancje dowolnych funkcji liniowych wektorów niewiadomych, a jednocześnie daje podstawy do oceny wiarygodności analizowanych modeli w aspekcie badanego zjawiska.

LITERATURA

Rao C.R., Mitra S.K., Extension of a duality theorem concerning g-inverse of matrices. Sankhya, 1975.

Rao C.R., Least squares theory for possibly singular models. Canadian Statistician, 1978.

Rao C.R., Modele liniowe statystyki matematycznej. PWN, Warszawa, 1982.