

JERZY K. BAKSALARY i RADOSŁAW KALA (Poznań)

Estymowalność liniowych funkcji parametrycznych w jednowymiarowym modelu liniowym z restrykcjami

(Praca przyjęta do druku 12.05.1976)

0. Wstęp. Obszerne omówienie zagadnienia liniowej nieobciążonej estymowalności liniowych funkcji parametrycznych w jednowymiarowym modelu liniowym znajduje się w [2]. Praca niniejsza poświęcona jest temu samemu problemowi, tyle że rozważanemu w sytuacji, gdy na parametry modelu nałożone są restrykcje, wyrażające o nich dodatkową wiedzę *a priori*.

Przedstawione tutaj rozważania są ogólniejsze od znanych w literaturze, dotyczą bowiem modelu, w którym macierz kowariancji wektora błędów losowych może być osobliwa. Pokazano, iż osobliwość ta powoduje jedynie zmianę definicji estymowalności, natomiast nie wywiera wpływu na postać kryteriów pozwalających estymowalność tę weryfikować. Zamieszczono szeroki przegląd takich kryteriów, a dodatkowo podjęto próbę sformalizowania, a przez to uwypuklenia, omawianych w literaturze związków między estymowalnością funkcji parametrycznych w modelu bez restrykcji i w modelu z restrykcjami.

Podobnie jak w [2], w pracy tej stosowane są następujące oznaczenia:

$\mathcal{M}_{m,n}$ — zbiór $m \times n$ -wymiarowych macierzy rzeczywistych,

\mathbf{I} — macierz jednostkowa (stopnia wynikającego z kontekstu),

\mathbf{A}' — transpozycja macierzy \mathbf{A} ,

\mathbf{A}^- — g -odwrotność macierzy \mathbf{A} , tzn. dowolna macierz spełniająca równanie $\mathbf{A}\mathbf{A}^-\mathbf{A} = \mathbf{A}$,

$r(\mathbf{A})$ — rząd macierzy \mathbf{A} ,

$\mathcal{C}(\mathbf{A})$ — podprzestrzeń rozpięta na kolumnach macierzy \mathbf{A} ,

$\mathcal{R}(\mathbf{A})$ — podprzestrzeń rozpięta na wierszach macierzy \mathbf{A} ,

$\mathcal{C}^\perp(\mathbf{A})$ — dopełnienie ortogonalne (w sensie standardowego iloczynu skalarnego) przestrzeni $\mathcal{C}(\mathbf{A})$,

$\mathcal{R}^\perp(\mathbf{A})$ — dopełnienie ortogonalne (w sensie jw.) przestrzeni $\mathcal{R}(\mathbf{A})$.

Ponadto, przy odwoływaniu się do twierdzeń i definicji używana jest litera T zamiast wyrazu *twierdzenie*, a litera D zamiast wyrazu *definicja*.

1. Kryteria estymowalności. Rozważmy jednowymiarowy model liniowy $y = \mathbf{X}\xi + \mathbf{e}$, oznaczany symbolem

$$(1.1) \quad (\mathbf{y}, \mathbf{X}\xi, \sigma^2\mathbf{V}),$$

gdzie $\mathbf{y} \in \mathcal{M}_{N,1}$ jest obserwowanym wektorem losowym, $\mathbf{X} \in \mathcal{M}_{N,p}$ jest znaną macierzą dowolnego rzędu, $\xi \in \mathcal{M}_{p,1}$ jest wektorem nieznanych parametrów będących przedmiotem analizy, natomiast $\mathbf{e} \in \mathcal{M}_{N,1}$ jest wektorem błędów losowych o wartości oczekiwanej $E(\mathbf{e}) = \mathbf{0}$ i macierzy kowariancji $D(\mathbf{e}) = \sigma^2\mathbf{V}$, przy czym σ^2 jest nieznanym parametrem dodatnim, a $\mathbf{V} \in \mathcal{M}_{N,N}$ znaną macierzą określoną nieujemnie. Przyjmujemy ponadto, że na wektor parametrów ξ nałożone są niesprzeczne (jako układ równań liniowych) restrykcje postaci

$$(1.2) \quad \mathbf{R}\xi = \mathbf{s},$$

gdzie $\mathbf{R} \in \mathcal{M}_{q,p}$, a $\mathbf{s} \in \mathcal{M}_{q,1}$. Model liniowy (1.1) z parametrami podlegającymi restrykcjom (1.2) oznaczany będzie symbolem

$$(1.3) \quad (\mathbf{y}, \mathbf{X}\xi | \mathbf{R}\xi = \mathbf{s}, \sigma^2\mathbf{V}).$$

Problemy związane z estymowalnością liniowych funkcji parametrycznych w modelu (1.3) będziemy rozwiązywać wykorzystując możliwość (patrz np. [16] lub [12], str. 136) wyrażenia go w postaci

$$(1.4) \quad \left(\begin{bmatrix} \mathbf{y} \\ \mathbf{s} \end{bmatrix}, \begin{bmatrix} \mathbf{X} \\ \mathbf{R} \end{bmatrix} \xi, \sigma^2 \begin{bmatrix} \mathbf{V} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \right)$$

oraz stosując wyniki podane dla modelu (1.1) w [2].

Zauważmy, że (1.4) jest modelem z osobliwą macierzą kowariancji, bez względu na to, czy \mathbf{V} jest macierzą nieosobliwą czy też osobliwą. Wiadomo (patrz [11], [15], a także [2]), że w takiej sytuacji dalsze rozważania są sensowne jedynie wtedy, gdy model jest niesprzeczny. Przypomnijmy, że dla modelu (1.1) warunkiem niesprzeczności jest relacja

$$(1.5) \quad y_0 \in \mathcal{C}([\mathbf{X} : \mathbf{V}]),$$

gdzie y_0 jest aktualnie zaobserwowaną wartością wektora losowego \mathbf{y} . Wobec (1.4) wynika stąd, iż warunek taki dla modelu (1.3) ma postać

$$(1.6) \quad \begin{bmatrix} y_0 \\ \mathbf{s} \end{bmatrix} \in \mathcal{C} \left(\begin{bmatrix} \mathbf{X} & \mathbf{V} \\ \mathbf{R} & \mathbf{0} \end{bmatrix} \right).$$

Ileć w dalszym ciągu pracy będziemy mówić o modelu (1.3) albo o modelu (1.1), tyleć milcząco będziemy zakładać spełnienie odpowiednio warunku (1.6) albo warunku (1.5). Zauważmy jeszcze, że z (1.6) wynika zarówno niesprzeczność modelu (1.1), jak i niesprzeczność restrykcji (1.2), podczas gdy implikacja odwrotna jest nieprawdziwa, tzn. relacja (1.5) i niesprzeczność restrykcji (1.2) nie wystarczają do zapewnienia niesprzeczności modelu (1.3).

Wiadomo (choćby z [2]), że osobliwość macierzy kowariancji wektora błędów losowych implikuje istnienie restrykcji na wektor parametrów. Dla modelu (1.1) są

one postaci

$$(1.7) \quad \mathbf{K}'\mathbf{X}\xi = \mathbf{d},$$

gdzie \mathbf{K} jest dowolną taką macierzą, dla której $\mathcal{C}(\mathbf{K}) = \mathcal{C}^\perp(\mathbf{V})$, natomiast

$$(1.8) \quad \mathbf{d} = \mathbf{K}'\mathbf{y}_0.$$

Odpowiednikiem macierzy \mathbf{K} dla modelu (1.4) jest taka macierz \mathbf{K}_1 , że

$$\mathcal{C}(\mathbf{K}_1) = \mathcal{C}^\perp\left(\begin{bmatrix} \mathbf{V} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\right).$$

Ponieważ za \mathbf{K}_1 można przyjąć macierz

$$\begin{bmatrix} \mathbf{K}' & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix},$$

więc analogon (1.7) dla modelu (1.3) ma postać

$$\begin{bmatrix} \mathbf{K}' & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ \mathbf{R} \end{bmatrix} \xi = \begin{bmatrix} \mathbf{K}' & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{y}_0 \\ \mathbf{s} \end{bmatrix},$$

a wobec (1.8)

$$(1.9) \quad \begin{bmatrix} \mathbf{K}'\mathbf{X} \\ \mathbf{R} \end{bmatrix} \xi = \begin{bmatrix} \mathbf{d} \\ \mathbf{s} \end{bmatrix}.$$

Związek między niesprzecznością układu restrykcyj (1.9) a niesprzecznością modelu wyrażoną relacją (1.6) podaje

TI.1. *Warunkiem koniecznym i dostatecznym niesprzeczności modelu $(\mathbf{y}, \mathbf{X}\xi|\mathbf{R}\xi = \mathbf{s}, \sigma^2\mathbf{V})$ jest niesprzeczność (jako układu równań) restrykcyj (1.9).*

D o w ó d. TI.1 wynika bezpośrednio z zastosowania do modelu postaci (1.4) TI.2 z [2]. ■

Przechodzimy obecnie do omawiania interesującego nas zagadnienia estymowalności liniowych funkcji parametrycznych w modelu (1.3). Odnosząc D1.2 z [2] do modelu (1.4) uzyskujemy następujące sformułowanie definicji estymowalności.

DEFINICJA 1.1. Liniowe funkcje parametryczne $\mathbf{C}\xi$, gdzie $\mathbf{C} \in \mathcal{M}_{k,p}$, nazywamy *estymowalnymi w modelu $(\mathbf{y}, \mathbf{X}\xi|\mathbf{R}\xi = \mathbf{s}, \sigma^2\mathbf{V})$* , jeżeli istnieją macierze $\mathbf{L} \in \mathcal{M}_{k,n}$ i $\mathbf{M} \in \mathcal{M}_{k,q}$ takie, że dla wszystkich ξ , spełniających układ restrykcyj (1.9),

$$(1.10) \quad \mathbf{E}(\mathbf{L}\mathbf{y}) + \mathbf{M}\mathbf{s} = \mathbf{C}\xi. \quad \blacksquare$$

Zestawimy teraz kryteria estymowalności funkcji $\mathbf{C}\xi$ w modelu (1.3) wykorzystując, podobnie jak powyżej, możliwość jego przedstawienia w postaci (1.4) oraz stosując odpowiednie twierdzenia podane dla modelu (1.1) w [2], §§ 1–3

TI.2. *Każdy z następujących warunków jest konieczny i dostateczny na to, by liniowe funkcje parametryczne $\mathbf{C}\xi$ były estymowalne w modelu $(\mathbf{y}, \mathbf{X}\xi|\mathbf{R}\xi = \mathbf{s}, \sigma^2\mathbf{V})$:*

$$(1.11) \quad \mathcal{R}(\mathbf{C}) \subset \mathcal{R}\left(\begin{bmatrix} \mathbf{X} \\ \mathbf{R} \end{bmatrix}\right),$$

$$(1.12) \quad C(X'X + R'R)^{-1}(X'X + R'R) = C,$$

$$(1.13) \quad C \begin{bmatrix} X \\ R \end{bmatrix}^{-1} \begin{bmatrix} X \\ R \end{bmatrix} = C,$$

$$(1.14) \quad C(I - X^{-1}X)(I - F^{-1}F) = 0, \quad \text{gdzie} \quad F = R(I - X^{-1}X),$$

$$(1.15) \quad r \left(\begin{bmatrix} X \\ R \\ C \end{bmatrix} \right) = r \left(\begin{bmatrix} X \\ R \end{bmatrix} \right),$$

$$(1.16) \quad r \left\{ \begin{bmatrix} X \\ R \end{bmatrix} (I - C^{-1}C) \right\} = r \left(\begin{bmatrix} X \\ R \end{bmatrix} \right) - r(C),$$

$$(1.17) \quad r \left\{ \begin{bmatrix} X \\ R \end{bmatrix} (I - C^{-1}C) \right\} = r\{X(I - R^{-1}R)\} + r(R) - r(C).$$

D o ó d. Warunek (1.11) jest odpowiednikiem kryterium podanego dla modelu bez restrykcji przez Bose'a [3] (patrz również T1.3 w [2]).

Relacja (1.12) wynika z zastosowania do (1.4) warunku podanego przez Rao [10] (patrz również T2.1 w [2]). W przypadku modelu z $V = I$ kryterium (1.12) zostało wprowadzone przez Pringle'a i Raynera w [8].

Kolejne kryterium jest analogonem warunku Rao i Mitry ([12], str. 139). Dla modelu z restrykcjami zostało ono podane przez Nelsona *et al.* [6], tyle że w postaci mniej ogólnej, bo przy użyciu odwrotności Moore'a-Penrose'a (patrz [7]) zamiast g -odwrotności dowolnej.

Równość (1.14) wynika z (1.13) po zastosowaniu wzoru wyrażającego g -odwrotność macierzy podzielonej wierszowo, który otrzymuje się jako odpowiednik wzoru podanego dla macierzy podzielonej kolumnowo przez Pringle'a i Raynera ([8], str. 38). W przypadku macierzy nas interesującej wzór ten ma postać

$$(1.18) \quad \begin{bmatrix} X \\ R \end{bmatrix}^{-1} = [X^{-1} - (I - X^{-1}X)F^{-1}RX^{-1} \quad (I - X^{-1}X)F^{-1}],$$

gdzie $F = R(I - X^{-1}X)$, a X^{-1} i F^{-1} są dowolnymi g -odwrotnościami odpowiednio macierzy X i F . Po podstawieniu (1.18) do (1.13), równość (1.14) uzyskuje się jako wynik prostych przekształceń.

Kryterium (1.15) jest bezpośrednim wynikiem zastosowania T3.1 z [2] do modelu (1.4).

Warunki (1.16) i (1.17) stanowią uogólnienia kryterium Millikena [5]. Zostały one wyprowadzone przez Baksalarego i Kalę w [1]. ■

Przy ocenie przydatności numerycznej powyższych kryteriów można wykorzystać uwagi znajdujące się w [2], § 4.

2. Zależności między estymowalnością w modelu $(y, X\xi, R\xi = s, \sigma^2V)$ a estymowalnością w modelu $(y, X\xi, \sigma^2V)$. W monografii Searle'a ([14], § 5.6) zawarte są rozważania dotyczące zależności między estymowalnością liniowych funkcji parametrycznych w modelu z restrykcjami a estymowalnością tychże funkcji w modelu

bez restrykcji. Poniższe twierdzenia stanowią formalizację zasadniczych wniosków z tych rozważań.

Twierdzenie 2.1. *Jeżeli liniowe funkcje parametryczne $C\xi$ są estymowalne w modelu $(y, X\xi, \sigma^2V)$, to są one estymowalne również w modelu $(y, X\xi | R\xi = s, \sigma^2V)$ przy dowolnych restrykcjach $R\xi = s$, dla których model pozostaje niesprzeczny.*

D o w ó d. Wiadomo (patrz T1.3 w [2]), że estymowalność funkcji $C\xi$ w modelu (1.1) implikuje relację $\mathcal{R}(C) \subset \mathcal{R}(X)$. Wówczas, oczywiście, dla dowolnej macierzy R (o odpowiedniej liczbie kolumn)

$$\mathcal{R}(C) \subset \mathcal{R}\left(\begin{bmatrix} X \\ R \end{bmatrix}\right),$$

skąd na mocy (1.11) wynika estymowalność funkcji $C\xi$ w modelu (1.3). ■

Twierdzenie 2.2. *Jeżeli liniowe funkcje parametryczne $C\xi$ są estymowalne w modelu $(y, X\xi | R\xi = s, \sigma^2V)$, a ponadto funkcje $R\xi$ są estymowalne w modelu $(y, X\xi, \sigma^2V)$, to funkcje $C\xi$ są estymowalne również w modelu $(y, X\xi, \sigma^2V)$.*

D o w ó d. Z założenia estymowalności funkcji $R\xi$ w modelu (1.1) wynika (patrz T1.3 w [2]), iż $\mathcal{R}(R) \subset \mathcal{R}(X)$. Wtedy oczywiście

$$\mathcal{R}\left(\begin{bmatrix} X \\ R \end{bmatrix}\right) = \mathcal{R}(X),$$

a stąd relacja (1.11), spełniona na mocy założenia estymowalności funkcji $C\xi$ w modelu (1.3), przyjmuje postać

$$\mathcal{R}(C) \subset \mathcal{R}(X).$$

Wobec T1.3 z [2] oznacza to, że funkcje $C\xi$ są estymowalne w modelu (1.1). ■

Twierdzenie 2.3. *Jeżeli restrykcje $R\xi = s$ są takie, że $\mathcal{R}(R) \supset \mathcal{R}^\perp(X)$, to w modelu $(y, X\xi | R\xi = s, \sigma^2V)$ liniowe funkcje parametryczne $C\xi$ są estymowalne dla dowolnej macierzy C o p kolumnach.*

D o w ó d. Jest jasne, że

$$\mathcal{R}\left(\begin{bmatrix} X \\ R \end{bmatrix}\right) = \{x = x_1 + x_2 : x_1 \in \mathcal{R}(X), x_2 \in \mathcal{R}(R)\}.$$

Jednakże wobec założenia, że $\mathcal{R}(R) \supset \mathcal{R}^\perp(X)$, mamy

$$(2.1) \quad \mathcal{R}\left(\begin{bmatrix} X \\ R \end{bmatrix}\right) = \mathcal{M}_{1,p}.$$

Niech teraz C będzie dowolną macierzą o p kolumnach. Ponieważ, oczywiście, $\mathcal{R}(C) \subset \mathcal{M}_{1,p}$, więc na mocy (2.1) zachodzi relacja

$$\mathcal{R}(C) \subset \mathcal{R}\left(\begin{bmatrix} X \\ R \end{bmatrix}\right),$$

a zatem spełniony jest warunek (1.11) estymowalności funkcji $C\xi$ w modelu (1.3). ■

Przyjrzyjmy się powyższym twierdzeniom raz jeszcze. T2.1 orzeka w gruncie rzeczy tyle, że nałożenie restrykcji na parametry modelu może klasę funkcji estymowalnych jedynie rozszerzyć. T2.2 powiada natomiast, że restrykcje wyrażone funkcjami estymowalnymi klasy funkcji estymowalnych nie zmieniają. Płyń więc stąd wniosek, że jej rozszerzenie może nastąpić jedynie w wyniku nałożenia na parametry modelu restrykcji określonych funkcjami nieestymowalnymi. W przypadku skrajnym można na tej drodze zapewnić estymowalność wszystkich parametrów modelu bezpośrednio, a więc, w konsekwencji, estymowalność wszystkich liniowych funkcji parametrycznych. Daje się to osiągnąć przez wprowadzenie do modelu takich restrykcji, które spełniają warunek podany w T2.3. To ostatnie podejście jest w istocie wykorzystywane w wielu metodach analizy wyników doświadczeń (patrz np. [13] lub [4]).

Podkreślmy na zakończenie, że włączanie restrykcji do modelu, a więc posługiwanie się modelem (1.3), jest uzasadnione wyłącznie wtedy, gdy restrykcje te wyrażają o parametrach modelu pewną dodatkową wiedzę *a priori*. Błędne jest natomiast interpretowanie jako restrykcji takich związków liniowych między parametrami, które nakładane są ze względów czysto obliczeniowych. Na niebezpieczeństwa wynikające z mechanicznego utożsamiania takich związków z restrykcjami zwrócili uwagę między innymi Searle ([14], str. 209 i 212) oraz Pringle i Rayner ([9], str. 93–98 i 118). Niebezpieczeństwa te stają się oczywiste, jeżeli weźmiemy pod uwagę wpływ, jaki wywierają restrykcje na przebieg i wyniki analizy. Mówiąc najogólniej, jest on następujący: restrykcje wyrażone funkcjami nieestymowalnymi zmieniają interpretację parametrów modelu, podczas gdy restrykcje wyrażone funkcjami estymowalnymi zmieniają estymatory funkcji parametrycznych i — w konsekwencji — także konstruowane dla nich przedziały ufności lub obszary krytyczne.

Prace cytowane

- [1] J. K. Baksalary, R. Kala, *Extensions of Milliken's estimability criterion*, Ann. Statist. 4 (1976), str. 639–641.
- [2] — — *Estymowalność liniowych funkcji parametrycznych w jednowymiarowym modelu liniowym*. Matematyka Stosowana niniejszy tom, str. 133–144.
- [3] R. C. Bose, *The fundamental theorem of linear estimation*, Proceedings of the Thirty-First Indian Science Congress 4 III (1944), str. 2–3.
- [4] W. T. Federer, M. Zelen, *Analysis of multifactor classifications with unequal numbers of observations*, Biometrics 22 (1966), str. 525–552.
- [5] G. A. Milliken, *New criteria for estimability for linear models*, Ann. Math. Statist. 24 (1971), str. 1588–1594.
- [6] D. L. Nelson, T. O. Lewis, T. L. Boullion, *A generalized inverse method for regression analysis with linear restrictions*, Industrial Math. 22 (1972), str. 1–10.
- [7] R. Penrose, *A generalized inverse for matrices*, Proc. Cambridge Philos. Soc. 51 (1955), str. 406–413.
- [8] R. M. Pringle, A. A. Rayner, *Expressions for generalized inverses of a bordered matrix with application to the theory of constrained linear models*, SIAM Rev. 12 (1970), s. 107–115.
- [9] — — *Generalized inverse matrices with applications to statistics*, London 1971.

- [10] C. R. R a o, *A note on a generalized inverse of a matrix with applications to problems in mathematical statistics*, J. R. Statist. Soc. B 24 (1962), str. 152–158.
- [11] — *Representations of best linear unbiased estimators in the Gauss–Markoff model with a singular dispersion matrix*, J. Multivariate Anal. 3 (1973), str. 276–292.
- [12] —, S. K. M i t r a, *Generalized inverse of matrices and its applications*, New York 1971.
- [13] D. H. R e e s, *The analysis of variance of designs with many non-orthogonal classifications*, J. R. Statist. Soc. B 28 (1966), str. 110–117.
- [14] S. R. S e a r l e, *Linear models*, New York 1971.
- [15] G. Z y s k i n d, *Error structures, projections and conditional inverses in linear model theory; w: A survey of statistical design and linear models* (J. N. S r i v a s t a v a, ed.), str. 647–663, Amsterdam 1975.
- [16] —, F. B. M a r t i n, *On best linear estimation and a general Gauss–Markov theorem in linear models with arbitrary nonnegative covariance structure*, SIAM J. Appl. Math. 17 (1969), str. 1190–1202.

ZAKŁAD METOD MATEMATYCZNYCH I STATYSTYCZNYCH
AKADEMII ROLNICZEJ W POZNANIU