WYDZIAŁ MATEMATYKI, INFORMATYKI I EKONOMETRII
UNIWERSYTET ZIELONOGÓRSKI

Ewa Synówka-Bejenka

Estymacja dopuszczalna efektów stałych i losowych w modelach liniowych

ROZPRAWA DOKTORSKA NAPISANA POD KIERUNKIEM
dra hab. Stefana Zontka, prof. UZ

Zielona Góra 2006

Serdecznie dziękuję Panu Profesorowi Stefanowi Zontkowi za cenne wskazówki i krytyczne uwagi, a także za życzliwość, wyrozumiałość i poświęcony czas, które okazały się bardzo pomocne przy pisaniu tej rozprawy.

Spis treści

Wstęp

Problem dopuszczalności liniowych estymatorów w ogólnym modelu liniowym był wielokrotnie rozważany w literaturze. Pierwszy rezultat, uzyskany przez Cohena (1966), dotyczył charakterystyki dopuszczalnych liniowych estymatorów dla wektora wartości oczekiwanych w modelu Gaussa-Markowa z macierzą kowariancji $\sigma^2 I$. Został on później uogólniony przez Rao (1976) na przypadek modelu, w którym wektor wartości oczekiwanych należy do pewnej podprzestrzeni liniowej, a macierz kowariancji ma postać $\sigma^2 V$, gdzie V jest pewną znaną, dodatnio określoną macierzą. Dalsze uogólnienia związane z modelem Gaussa-Markowa podali m.in. Hoffman (1977), Klonecki (1982), Mathew, Sinha i Rao (1984), Klonecki i Zontek (1988), Drygas i Zmyślony (1988) oraz Baksalary, Markiewicz i Rao (1989).

Podstawowym, dla rozważań w tej rozprawie, jest wynik LaMotte'a (1982). Jego uniwersalność polega na tym, że warunki konieczne i dostateczne na to, aby liniowy estymator był dopuszczalny są podane dla ogólnego modelu liniowego bez ograniczeń nałożonych na przestrzeń parametrów. LaMotte, poprzez lokalną optymalność (patrz Olsen, Seely i Birkes, 1976) w zstępującym ciągu rozmaitości liniowych, opisuje skończoną, krokową procedurę konstrukcji dowolnego estymatora dopuszczalnego. Otrzymuje się w ten sposób charakterystykę estymatorów dopuszczalnych, ale w bardzo uwikłanej formie. Wykorzystany będzie też inny rezultat LaMotte'a (1997) (patrz także Stępniak, 1987), który pokazuje, że jednoznacznie lokalnie optymalne estymatory i ich granice tworzą zupełną klasę estymatorów. Zastosujemy te wyniki do problemu jednoczesnej dopuszczalnej estymacji wektora efektów stałych i losowych w mieszanym modelu liniowym po wstępnym przejściu do równoważnego problemu estymacji tylko efektów stałych w odpowiednim modelu. Zagadnieniem estymacji efektów stałych i losowych zajmowali się m.in. Harville (1976), Peixoto i Harville (1986), Rao (1987), Robinson (1991) oraz Gross i Markiewicz (1999).

Rozprawa składa się z pięciu rozdziałów. W rozdziale pierwszym została opisana wspomniana już metoda LaMotte'a umożliwiająca charakterystykę liniowych estymatorów dopuszczalnych dla liniowej funkcji wektora wartości oczekiwanych w ogólnym modelu liniowym w pewnej rozmaitości liniowej.

W rozdziale drugim opisano mieszany model liniowy. Dokonano także redukcji problemu łącznej estymacji wektora efektów stałych i losowych do estymacji wektora efektów stałych w odpowiednim modelu liniowym w ograniczeniu do pewnej rozmaitości liniowej. Następnie sformułowano lemat, na mocy którego dopuszczalność estymatorów w odpowiadających sobie modelach jest równoważna. Na koniec rozdziału podane zostały przykłady modeli liniowych.

Rozdział trzeci składa się z trzech podrozdziałów, z których każdy zawiera charakterystykę dopuszczalnych estymatorów liniowych w zrównoważonym modelu losowym k – kierunkowej klasyfikacji hierarchicznej, zrównoważonym modelu losowym k – kierunkowej klasyfikacji krzyżowej oraz zrównoważonym modelu losowym dwukierunkowej klasyfikacji krzyżowej z interakcją.

Rozdział czwarty dotyczy problemu dopuszczalnej estymacji w modelu z dwoma komponentami, w którym zostało opuszczone pewne założenie z poprzedniego rozdziału. Odpowiednie twierdzenia pokazują, jaki wpływ ma to osłabienie założeń na postać klasy estymatorów dopuszczalnych.

W rozdziale piątym podane zostały warunki konieczne i dostateczne na to, aby estymator był dopuszczalny w ogólnym modelu liniowym z dwoma komponentami. Warunki te można postrzegać jako uogólnienie wyniku Rao (1976).

Wykaz ważniejszych oznaczeń

$\mathbb{R}^n$ – zbiór n-wymiarowych wektorów o współrzędnych rzeczywistych

$v_i \in \mathbb{R}^n$ – i-ty n-wymiarowy wersor

$\mathcal{M}_{n\times t}$ – zbiór macierzy wymiaru $n \times t$ o elementach rzeczywistych

$\mathcal{M}_n^{\geq}$ – zbiór macierzy symetrycznych i nieujemnie określonych wymiaru $n \times t$

A' – transpozycja macierzy A

A^{-1} – macierz odwrotna do A

A^{+} – odwrotność Moore'a-Penrose'a macierzy A

$\mathrm{tr}(A)$ – ślad macierzy kwadratowej A

$\mathcal{R}(A)$ – przestrzeń liniowa generowana przez kolumny macierzy A

$\mathcal{N}(A)$ – jądro macierzy A

$\mathrm{diag}(A_1, \ldots, A_n)$ – macierz diagonalna, w której macierze $A_1, \ldots, A_n$ są elementami diagonali

$A \otimes B$ – iloczyn Kroneckera macierzy A i B

T_A^t – liniowe przekształcenie odwzorowujące $\mathcal{M}_{n\times t}$ w $\mathcal{M}_{n\times t}$, określone dla każdej macierzy $B \in \mathcal{M}_{n\times t}$ wzorem $T_A^t(B) = AB$, gdzie $A \in \mathcal{M}_{n\times n}$

$\mathcal{R}(T_A^t)$ – obraz przekształcenia T_A^t

$\mathrm{span}\mathcal{W}$ – przestrzeń liniowa rozpięta na elementach zbioru $\mathcal{W}$

$[\mathcal{W}]$ – najmniejszy wypukły, domknięty stożek zawierający $\mathcal{W}$

$\mathrm{dim}\mathcal{W}$ – wymiar przestrzeni liniowej $\mathcal{W}$

$\mathrm{min}Z$ – najmniejszy element zbioru Z

1. Dopuszczalność w ogólnym modelu liniowym

W tym rozdziale przypomnimy kilka najistotniejszych rezultatów uzyskanych przez LaMotte'a, które dotyczą sposobu wyznaczania estymatorów dopuszczalnych. Jak już było wspomniane we wstępie, podał on warunki konieczne i dostateczne na to, aby estymator liniowy był dopuszczalny w ogólnym modelu liniowym, nie zakładając przy tym żadnych ograniczeń na przestrzeń parametrów modelu.

Niech $\boldsymbol{Y}$ będzie n-wymiarowym wektorem losowym, którego rozkład należy do pewnej rodziny rozkładów $\mathcal{P}$. Zakładamy, że dla każdego rozkładu $P \in \mathcal{P}$ istnieje wektor wartości oczekiwanych $\mathrm{E}_P\boldsymbol{Y}$ oraz macierz kowariancji $\mathrm{cov}_P(\boldsymbol{Y})$. Będziemy zajmowali się dopuszczalną estymacją funkcji $\boldsymbol{K'}\mathrm{E}_P\boldsymbol{Y}$, gdzie $\boldsymbol{K} \in \mathcal{M}_{n\times t}$. Rozważania ograniczymy do liniowych estymatorów postaci $\boldsymbol{L}'\boldsymbol{Y}$, dla których macierz $\boldsymbol{L}$ należy do następującej rozmaitości liniowej

$$\mathcal{L}_o = \{\boldsymbol{L}_o + \boldsymbol{\Pi}_o\boldsymbol{M} : \boldsymbol{M} \in \mathcal{M}_{n\times t}\},$$

gdzie $\boldsymbol{L}_o \in \mathcal{M}_{n\times t}$ oraz $\boldsymbol{\Pi}_o \in \mathcal{M}_{n\times n}$ są pewnymi ustalonymi macierzami.

Rozmaitości liniowe odgrywają ważną rolę w teorii estymacji. Klasycznym przykładem takiej rozmaitości jest zbiór liniowych nieobciążonych estymatorów funkcji $\boldsymbol{K'}\mathrm{E}_P\boldsymbol{Y}$ w modelu, w którym $\mathrm{E}_P\boldsymbol{Y} = \boldsymbol{X}\beta$, gdzie $\boldsymbol{X} \in \mathcal{M}_{n\times p}$ jest znaną macierzą, a β jest p-wymiarowym wektorem parametrów stałych. Estymatory z tego zbioru możemy bowiem przedstawić w postaci

$$\boldsymbol{K'}(\boldsymbol{XX'})(\boldsymbol{XX'})^{+}\boldsymbol{Y}+\boldsymbol{M'}\left[I_n-\boldsymbol{XX'}(\boldsymbol{XX'})^{+}\right]\boldsymbol{Y},$$

gdzie $\boldsymbol{M}$ jest dowolną $(n\times t)$-wymiarową macierzą. W tym przypadku klasa nieobciążonych estymatorów dla $\boldsymbol{K'X}\beta$ jest wyznaczona przez $\mathcal{L}_o=\{\boldsymbol{L}_o+\boldsymbol{\Pi}_o\boldsymbol{M}:\boldsymbol{M}\in\mathcal{M}_{n\times t}\}$, gdzie $\boldsymbol{L}_o=\boldsymbol{XX'}(\boldsymbol{XX'})^{+}\boldsymbol{K}$ oraz $\boldsymbol{\Pi}_o=I_n-\boldsymbol{XX'}(\boldsymbol{XX'})^{+}$.

Do porównania estymatorów użyjemy funkcji ryzyka określonej na $\mathcal{P}$, którą dla estymatora $\boldsymbol{L'Y}$ definiujemy jako $\mathrm{E}[(\boldsymbol{L'Y}-\boldsymbol{K'}\mathrm{E}_P\boldsymbol{Y})'(\boldsymbol{L'Y}-\boldsymbol{K'}\mathrm{E}_P\boldsymbol{Y})]$. Po prostych przekształceniach funkcję tę możemy przedstawić w postaci

$$\mathrm{tr}\left[\boldsymbol{L}'\mathrm{cov}_P(\boldsymbol{Y})\boldsymbol{L}+(\boldsymbol{L}-\boldsymbol{K})'\,\mathrm{E}_P\boldsymbol{Y}\,(\mathrm{E}_P\boldsymbol{Y})'\,(\boldsymbol{L}-\boldsymbol{K})\right].$$

Widzimy, że zależy ona od rozkładu P wektora losowego $\boldsymbol{Y}$ poprzez $\mathrm{cov}_P(\boldsymbol{Y})$ oraz $\mathrm{E}_P\boldsymbol{Y}(\mathrm{E}_P\boldsymbol{Y})'$. LaMotte wykorzystał ten fakt i potraktował parę ($\mathrm{cov}_P(\boldsymbol{Y})$, $\mathrm{E}_P\boldsymbol{Y}(\mathrm{E}_P\boldsymbol{Y})'$) jako nowy parametr przebiegający zbiór postaci

$$\mathcal{T}=\left\{(\mathrm{cov}_P(\boldsymbol{Y}),\mathrm{E}_P\boldsymbol{Y}(\mathrm{E}_P\boldsymbol{Y})'):P\in\mathcal{P}\right\}.$$

Zauważył też, że przy badaniu własności estymatorów przydaje się rozszerzenie pojęcia funkcji ryzyka ze zbioru $\mathcal{T}$ na zbiór $\mathcal{W}=\mathrm{span}\mathcal{T}$. Tę rozszerzoną funkcję ryzyka dla każdego $(W_1,W_2)\in\mathcal{W}$ zdefiniował wzorem

$$\boldsymbol{R}(\boldsymbol{L'Y};(W_1,W_2))=tr\left[\boldsymbol{L}'W_1\boldsymbol{L}+(\boldsymbol{L}-\boldsymbol{K})'W_2(\boldsymbol{L}-\boldsymbol{K})\right].$$

Niech $\mathcal{L}$ będzie podrozmaitością $\mathcal{L}_o$. Wtedy $\mathcal{L}$ można opisać następująco

$$\{\boldsymbol{L}_1+\boldsymbol{\Pi}\boldsymbol{M}:\boldsymbol{M}\in\mathcal{M}_{n\times t}\},$$

gdzie $\boldsymbol{L}_1\in\mathcal{L}_o$, natomiast $\boldsymbol{\Pi}$ jest macierzą wymiaru $(n\times n)$ taką, że $\mathcal{R}(\boldsymbol{\Pi})\subset\mathcal{R}(\boldsymbol{\Pi}_o)$. Zauważmy, że funkcję ryzyka estymatora $\boldsymbol{L'Y}$, gdzie $\boldsymbol{L}\in\mathcal{L}$, możemy przedstawić w postaci

$$\boldsymbol{R}(\boldsymbol{L'Y};(W_1,W_2))=tr[\boldsymbol{M'T}_1(W)\boldsymbol{M}]+2tr[\boldsymbol{M'T}_2(W)]+\boldsymbol{R}\left(\boldsymbol{L}_1'\boldsymbol{Y};(W_1,W_2)\right),$$

gdzie

$$\boldsymbol{T}_1(W) = \mathbf{\Pi}'(W_1 + W_2)\mathbf{\Pi} \text{ oraz } \boldsymbol{T}_2(W) = \mathbf{\Pi}' W_2 \boldsymbol{K} - \mathbf{\Pi}'(W_1 + W_2)\boldsymbol{L}_1.$$

Definicja 1.1. *Estymator $\boldsymbol{L}'\boldsymbol{Y}$ jest lokalnie optymalny w punkcie $W \in \mathcal{W}$ w klasie $\mathcal{L}$, jeżeli $R(\boldsymbol{L}'\boldsymbol{Y}; W) \leq R(\boldsymbol{L}_*'\boldsymbol{Y}; W)$ dla każdego $\boldsymbol{L}_* \in \mathcal{L}$.*

Poniższe twierdzenie (LaMotte, 1982) podaje pewne rezultaty dotyczące lokalnej optymalności.

Twierdzenie 1.1.

(i) *W klasie $\mathcal{L}$ istnieje co najmniej jeden estymator lokalnie optymalny w punkcie $W = (W_1, W_2) \in \mathcal{W}$ wtedy i tylko wtedy, gdy macierz $\mathbf{\Pi}'(W_1 + W_2)\mathbf{\Pi}$ jest macierzą nieujemnie określoną oraz $\mathcal{R}(\mathbf{\Pi}' W_2 \boldsymbol{K} - \mathbf{\Pi}'(W_1 + W_2)\boldsymbol{L}_1) \subseteq \mathcal{R}(\mathbf{\Pi}'(W_1 + W_2)\mathbf{\Pi})$.*

(ii) *Estymator $\boldsymbol{L}'\boldsymbol{Y}$ jest estymatorem lokalnie optymalnym w punkcie $W \in \mathcal{W}$ w klasie $\mathcal{L}$ wtedy i tylko wtedy, gdy $\mathbf{\Pi}'(W_1 + W_2)\mathbf{\Pi}$ jest macierzą nieujemnie określoną oraz $\mathbf{\Pi}'(W_1 + W_2)\boldsymbol{L} = \mathbf{\Pi}' W_2 \boldsymbol{K}$.*

(iii) *W klasie $\mathcal{L}$ istnieje dokładnie jeden estymator optymalny w punkcie W wtedy i tylko wtedy, gdy $\mathbf{\Pi}'(W_1 + W_2)\mathbf{\Pi}$ jest macierzą nieujemnie określoną taką, że $\mathcal{R}(\mathbf{\Pi}'(W_1 + W_2)\mathbf{\Pi}) = \mathcal{R}(\mathbf{\Pi}')$.*

Oznaczmy przez $\mathcal{B}(W|\mathcal{L})$ zbiór tych macierzy $\boldsymbol{L} \in \mathcal{L}$, dla których estymator $\boldsymbol{L}'\boldsymbol{Y}$ jest lokalnie optymalny w punkcie W w klasie $\mathcal{L}$. Zauważmy, że dla danego punktu $W \in \mathcal{W}$ może nie istnieć estymator $\boldsymbol{L}'\boldsymbol{Y}$, który jest w nim lokalnie optymalny w klasie $\mathcal{L}$, może istnieć dokładnie jeden taki estymator albo wreszcie może być wiele estymatorów lokalnie optymalnych w tym punkcie. Estymator $\boldsymbol{L}'\boldsymbol{Y}$, dla którego $\mathcal{B}(W|\mathcal{L}) =\{\boldsymbol{L}\}$, będziemy nazywać jednoznacznie lokalnie optymalnym w punkcie W w klasie $\mathcal{L}$. Jednoznacznie lokalnie optymalne estymatory w klasie $\mathcal{L}$ są scharakteryzowane w podpunkcie (iii) twierdzenia 1.1.

Zauważmy, że niepusty zbiór $\mathcal{B}(W|\mathcal{L})$ jest także rozmaitością liniową. Oczywiście $\mathcal{B}(W|\mathcal{L})$ jest podrozmaitością $\mathcal{L}$, a w przypadku, gdy nie wszystkie estymatory w klasie $\mathcal{L}$ są lokalnie optymalne w punkcie W, jest właściwym podzbio- rem $\mathcal{L}$. Punkty, w których każdy estymator $\boldsymbol{L}'\boldsymbol{Y}$ jest lokalnie optymalny w klasie $\mathcal{L}$, LaMotte nazwał trywialnymi. A zatem punkt $W \in \mathcal{W}$ nazywamy *trywialnym* dla $\mathcal{L}$, jeżeli $\mathcal{B}(W|\mathcal{L}) = \mathcal{L}$. Przykładem takiego punktu jest $W = (\mathbf{0}, \mathbf{0})$. Zbiór wszystkich punktów trywialnych dla $\mathcal{L}$ oznaczać będziemy przez $\mathcal{S} = \mathcal{S}(\mathcal{L})$. LaMotte udowodnił, że jeżeli przedstawimy $\mathcal{L}$ w postaci $\{\boldsymbol{L}_1 + \boldsymbol{\Pi}\boldsymbol{M} : \boldsymbol{M} \in \mathcal{M}_{n\times t}\}$, to zbiór $\mathcal{S}$ możemy zapisać następująco

$$\mathcal{S} = \mathcal{S}(\mathcal{L}) = \{(W_1, W_2) \in \mathcal{W} : \boldsymbol{\Pi}'(W_1 + W_2)\boldsymbol{\Pi} = 0,\ \boldsymbol{\Pi}'W_1\boldsymbol{L}_1 = \boldsymbol{\Pi}'W_2(\boldsymbol{K} - \boldsymbol{L}_1)\}\,.$$

Stosując jako kryterium porównawcze funkcję ryzyka możemy w klasie $\mathcal{L}_o$ wprowadzić wśród rozważanych estymatorów następującą relację częściowego porządku.

Definicja 1.2. *Estymator $\boldsymbol{L}'\boldsymbol{Y}$ jest lepszy niż estymator $\boldsymbol{L}_*'\boldsymbol{Y}$ jeżeli*

$$R(\boldsymbol{L}'\boldsymbol{Y}; W) \leq R(\boldsymbol{L}_*'\boldsymbol{Y}; W) \ \text{ dla } \textit{każdego}\ W \in \mathcal{T}$$

i jeżeli ostra nierówność jest spełniona dla co najmniej jednego z tych punktów.

Definicja 1.3. *Estymator $\boldsymbol{L}'\boldsymbol{Y}$, gdzie $\boldsymbol{L} \in \mathcal{L}_o$, jest estymatorem dopuszczalnym dla $\boldsymbol{K}'\mathrm{E}_P\boldsymbol{Y}$ na zbiorze parametrów $\mathcal{T}$, jeżeli w klasie $\mathcal{L}_o$ nie istnieje estymator lepszy niż $\boldsymbol{L}'\boldsymbol{Y}$.*

Podobnie definiujemy pojęcia *lepszy niż* oraz *dopuszczalny* w klasie $\mathcal{L}$ na zbiorze parametrów $\mathcal{T}$, a także na każdym zbiorze $\mathcal{C} \subseteq \mathcal{W}$. LaMotte wykazał, że pojęcia te są równoważne na zbiorach $\mathcal{T}$, $[\mathcal{T}]$ oraz $[\mathcal{T} + \mathcal{S}]$, gdzie $\mathcal{S}$= $\mathcal{S}(\mathcal{L})$.

Punktem wyjścia w jego rozważaniach nad dopuszczalnością było rozszerzenie wyniku Olsena, Seely'ego i Birkesa (1976) do twierdzenia, że jeżeli $\boldsymbol{L}'\boldsymbol{Y}$ jest dopuszczalny w klasie $\mathcal{L}_o$ na zbiorze parametrów $\mathcal{T}$, to istnieje niezerowy punkt $W \in [\mathcal{T}]$, w którym $\boldsymbol{L}'\boldsymbol{Y}$ jest lokalnie optymalny w klasie $\mathcal{L}_o$. Następnie zauważył, że jeżeli $\boldsymbol{L}'\boldsymbol{Y}$ jest lokalnie optymalny w punkcie $W \in [\mathcal{T}]$ w klasie $\mathcal{L}_o$, to dopuszczalność $\boldsymbol{L}'\boldsymbol{Y}$ w klasie $\mathcal{L}_o$ na zbiorze $\mathcal{T}$ jest równoważna dopuszczalności $\boldsymbol{L}'\boldsymbol{Y}$ w klasie $\mathcal{L}_1 = \mathcal{B}(W|\mathcal{L}_o)$ na zbiorze $\mathcal{T}$, co z kolei sugerowało, że chcąc uzyskać charakterystykę estymatorów dopuszczalnych w klasie $\mathcal{L}_o$ należy powtórzyć tok rozumowania z jednoczesną redukcją wymiaru kolejno uzyskiwanych podrozmaitości aż do momentu uzyskania zbioru składającego się tylko z jednej macierzy. Niestety, w przypadku, gdy ograniczymy się jedynie do zbioru $[\mathcal{T}]$, redukcja wymiaru nie musi nastąpić. Stąd też propozycja LaMotte'a, aby rozszerzyć zbiór $[\mathcal{T}]$ do zbioru $[\mathcal{T} + \mathcal{S}_o]$. Oczywiście, jeżeli $\mathcal{T} \subset \mathcal{S}_o$, to wówczas każdy estymator $\boldsymbol{L}'\boldsymbol{Y}$ jest dopuszczalny w klasie $\mathcal{L}_o$ na $\mathcal{T}$.

Twierdzenie 1.2. *Estymator $\boldsymbol{L}'\boldsymbol{Y}$ jest dopuszczalny dla $\boldsymbol{K}'\mathrm{E}_P\boldsymbol{Y}$ w klasie $\mathcal{L}_o$ na zbiorze parametrów $\mathcal{T}$ wtedy i tylko wtedy, gdy $\mathcal{T} \subset \mathcal{S}_o$ albo, gdy istnieje nietrywialny punkt $W \in [\mathcal{T} + \mathcal{S}_o]$ taki, że $\boldsymbol{L}$ należy do klasy $\mathcal{B}(W|\mathcal{L}_o)$, w której $\boldsymbol{L}'\boldsymbol{Y}$ jest dopuszczalny na $\mathcal{T}$.*

Twierdzenie 1.2 jest podstawowym dla dalszych naszych rozważań. Pozwala ono sprawdzić nie tylko, czy dany estymator $\boldsymbol{L}'\boldsymbol{Y}$ jest dopuszczalny w klasie $\mathcal{L}_o$ na $\mathcal{T}$, ale również umożliwia konstrukcję takich estymatorów.

Niech

$$\mathcal{A}(\mathcal{L}) = \{W \in [\mathcal{T} + \mathcal{S}]\backslash\mathcal{S} : \mathcal{B}(W|\mathcal{L}) \neq \emptyset\}.$$

Zauważmy, że dla dowolnego $\mathcal{L} \subseteq \mathcal{L}_o$ mamy spełniony warunek $[\mathcal{T}]\backslash\mathcal{S} \subseteq \mathcal{A}(\mathcal{L})$.

Na bazie twierdzenia 1.2 można zaproponować następujący algorytm sprawdzania, czy estymator $\boldsymbol{L}'\boldsymbol{Y}$ jest dopuszczalny dla $\boldsymbol{K}'\mathrm{E}_P\boldsymbol{Y}$ w klasie $\mathcal{L}_o$.

(0) Podstawmy $i = 0$.

(1) Jeżeli $\mathcal{T} \subset \mathcal{S}_i$, to estymator $\boldsymbol{L}'\boldsymbol{Y}$ jest dopuszczalny w klasie $\mathcal{L}_o$ na $\mathcal{T}$.

(2) Jeżeli $\mathcal{T} \not\subset \mathcal{S}_i$, to mamy dwie możliwości:

(2.1) nie istnieje punkt $W_i \in \mathcal{A}(\mathcal{L}_i)$, dla którego $\boldsymbol{L} \in \mathcal{L}_{i+1} = \mathcal{B}(W_i|\mathcal{L}_i)$, tym samym $\boldsymbol{L}'\boldsymbol{Y}$ jest niedopuszczalny w klasie $\mathcal{L}_o$ na $\mathcal{T}$,

(2.2) istnieje pewien punkt $W_i \in \mathcal{A}(\mathcal{L}_i)$, dla którego $\boldsymbol{L} \in \mathcal{L}_{i+1} = \mathcal{B}(W_i|\mathcal{L}_i)$, wtedy oznaczając przez $\mathcal{S}_{i+1}$ zbiór punktów trywialnych dla $\mathcal{L}_{i+1}$ i dokonując podstawienia $i+1$ w miejsce i przechodzimy do punktu (1).

Zauważmy, że powyższa procedura kończy się po wykonaniu skończonej liczby kroków. Ostatecznie bowiem mamy, że $\boldsymbol{L}'\boldsymbol{Y}$ jest dopuszczalny w klasie $\mathcal{L}_o$ na $\mathcal{T}$ wtedy i tylko wtedy, gdy $\mathcal{T} \subset \mathcal{S}_{i^*} = \mathcal{S}(\mathcal{L}_{i^*})$ dla pewnego $0 \leq i^* \leq \mathrm{dim}\mathcal{W} - \mathrm{dim}\mathcal{S}_\mathrm{o}$. Oczywiście $\mathcal{S}_{i^*} \supset \mathcal{S}_{i^*-1} \supset \ldots \supset \mathcal{S}_o$, a co za tym idzie $\mathcal{L}_{i^*} \subset \mathcal{L}_{i^*-1} \subset \ldots \subset \mathcal{L}_o$.

Twierdzenie 1.2 daje także możliwość konstrukcji dopuszczalnych estymatorów w klasie $\mathcal{L}_o$.

(0) Podstawmy $i = 0$.

(1) Weźmy dowolny punkt $W_i \in \mathcal{A}(\mathcal{L}_i)$.

(1.1) Jeżeli $\mathcal{L}_{i+1} = \mathcal{B}(W_i|\mathcal{L}_i) = \{\boldsymbol{L}\}$, to estymator $\boldsymbol{L}'\boldsymbol{Y}$ jest dopuszczalny w klasie $\mathcal{L}_o$ na zbiorze parametrów $\mathcal{T}$.

(1.2) Jeżeli $\mathcal{L}_{i+1} = \mathcal{B}(W_i|\mathcal{L}_i) \neq \{\boldsymbol{L}\}$, to za i podstawiamy $i+1$ i przechodzimy do punktu (1).

Opisany sposób postępowania jest dosyć skomplikowany. Wymaga w każdym kroku wyznaczania nowych zbiorów $\mathcal{L}$, $\mathcal{S}$ i $\mathcal{A}(\mathcal{L})$, co na ogół nie jest rzeczą prostą. Pomocnym może być tutaj wynik LaMotte'a (1997) (patrz także Stępniak, 1987 i Zontek, 1988) pokazujący związek pomiędzy dopuszczalnymi esty- matorami a granicami lokalnie optymalnych estymatorów wyznaczonych w sposób jednoznaczny. Rezultat ten w terminach ogólnego modelu liniowego przedstawia się następująco.

Twierdzenie 1.3. *Każdy dopuszczalny liniowy estymator dla $\boldsymbol{K'}\mathrm{E}_P\boldsymbol{Y}$ w klasie estymatorów $\mathcal{L}_o$ jest granicą jednoznacznie lokalnie optymalnych estymatorów w punktach należących do $[\mathcal{T}]$ w klasie $\mathcal{L}_o$.*

Powyższe twierdzenie pozwala na ograniczenie uwagi do domknięcia zbioru macierzy $\boldsymbol{L} \in \mathcal{L}_o$ takich, że $\boldsymbol{L}'\boldsymbol{Y}$ jest jednoznacznie lokalnie optymalnym estymatorem w punktach ze zbioru $[\mathcal{T}]$ w klasie $\mathcal{L}_o$. Są modele, w których aby wykazać, że estymator graniczny jest dopuszczalny wystarczy ograniczyć się w każdym kroku procedury LaMotte'a tylko do zbioru $[\mathcal{T}]$. W rozważanych w rozprawie modelach tak właśnie jest.

W dalszej części pracy, chcąc podać warunki konieczne i dostateczne na to, aby liniowy estymator pewnej liniowej funkcji był dopuszczalny, użyjemy modyfikacji lematu Shinozakiego w ograniczeniu do klasy estymatorów $\mathcal{L}_o$ (patrz Klonecki i Zontek, 1988).

Lemat 1.1. *Jeżeli $\boldsymbol{L}'\boldsymbol{Y}$ jest dopuszczalny dla $\boldsymbol{K'}\mathrm{E}_P\boldsymbol{Y}$ w rozmaitości liniowej $\mathcal{L}_o = \{\boldsymbol{L}_o + \boldsymbol{\Pi}_o\boldsymbol{M} : \boldsymbol{M} \in \mathcal{M}_{n\times t}\}$, to dla każdej macierzy $\boldsymbol{C} \in \mathcal{M}_{t\times s}$ estymator $\boldsymbol{C'}\boldsymbol{L}'\boldsymbol{Y}$ jest dopuszczalny dla $\boldsymbol{C'}\boldsymbol{K'}\mathrm{E}_P\boldsymbol{Y}$ w klasie $\{\boldsymbol{L}_o\boldsymbol{C} + \boldsymbol{\Pi}_o\boldsymbol{N} : \boldsymbol{N} \in \mathcal{M}_{n\times s}\}$.*

2. Estymacja w mieszanym modelu liniowym

W dalszej części rozprawy nasze rozważania dotyczące dopuszczalności estymatorów odnosić się będą do modeli liniowych, które są szczególnymi przypadkami tzw. *mieszanego modelu liniowego.* Model ten możemy przedstawić w postaci

$$Y = X\beta + Z_1u_1 + Z_2u_2 + \ldots + Z_ku_k + e,$$

gdzie β jest p-wymiarowym wektorem nieznanych parametrów (tzw. efektów stałych), $X \in \mathcal{M}_{n\times p}, Z_1 \in \mathcal{M}_{n\times m_1}, \ldots, Z_k \in \mathcal{M}_{n\times m_k}$ są macierzami o znanych elementach, natomiast $u_1 \in \mathbb{R}^{m_1}, \ldots, u_k \in \mathbb{R}^{m_k}$ oraz $e \in \mathbb{R}^n$ są nieobserwowalnymi wektorami losowymi. Będziemy zakładać, że $u_1, \ldots, u_k$ i e są wzajemnie nieskorelowane, mają zerowe wartości oczekiwane, a ich macierze kowariancji wynoszą odpowiednio $\sigma_1^2 I_{m_1}, \ldots, \sigma_k^2 I_{m_k}$ oraz $\sigma_{k+1}^2 I_n$, gdzie $\sigma_1^2 \geq 0, \ldots, \sigma_k^2 \geq 0$, $\sigma_{k+1}^2 > 0$. W teorii modeli liniowych nieznane parametry $\sigma_1^2, \ldots, \sigma_{k+1}^2$ nazywane są *komponentami wariancyjnymi.*

W bardziej zwartej postaci powyższy model można zapisać następująco

$$Y = X\beta + Zu + e,$$

gdzie $Z = (Z_1, \ldots, Z_k)$ oraz $u = (u_1', u_2', \ldots, u_k')'$.

Przy poczynionych założeniach otrzymujemy, że $\mathrm{E}Y = X\beta$ i $\mathrm{cov}(Y) = ZDZ' + \sigma_{k+1}^2 I_n$, gdzie $D = \mathrm{cov}(u) = \mathrm{diag}(\sigma_1^2 I_{m_1}, \sigma_2^2 I_{m_2}, \ldots, \sigma_k^2 I_{m_k})$, co będziemy zapisywać symbolicznie jako

$$Y \sim (X\beta,\ ZDZ' + \sigma_{k+1}^2 I_n). \tag{2.1}$$

Istnieje szeroka klasa modeli liniowych o powyższej strukturze, które mają duże zastosowanie praktyczne. Służą one bowiem do modelowania wielu zjawisk, w których pojawia się źródło zmienności, np. do opisu danych pochodzących z eksperymentów genetycznych, medycznych, rolniczych, czy astronomicznych. Właśnie do opisu tych ostatnich astronom Airy (1861) użył jako pierwszy tzw. modelu jednokierunkowej klasyfikacji, który to model wraz z innymi przykładami zostanie omówiony pod koniec tego rozdziału.

Teraz zajmiemy się dopuszczalną estymacją wektora

$$[(K'X\beta)', (Q_1'Z_1u_1)', \ldots, (Q_k'Z_ku_k)']' \tag{2.2}$$

w klasie liniowych estymatorów $L'Y = (L_0, L_1, \ldots, L_k)'Y$, gdzie $K, L_0 \in \mathcal{M}_{n\times t_0}$, $\ldots, Q_k, L_k \in \mathcal{M}_{n\times t_k}$. Jako kryterium porównawcze estymatorów użyjemy funkcji ryzyka

$$\begin{aligned}
R(L'Y) = E&\left\{\begin{bmatrix} L_0'Y - K'X\beta \\ L_1'Y - Q_1'Z_1u_1 \\ \vdots \\ L_k'Y - Q_k'Z_ku_k \end{bmatrix}' \begin{bmatrix} L_0'Y - K'X\beta \\ L_1'Y - Q_1'Z_1u_1 \\ \vdots \\ L_k'Y - Q_k'Z_ku_k \end{bmatrix}\right\} \\
&= tr\{(L_0 - K)'X\beta\beta'X'(L_0 - K) + L_0'cov(Y)L_0 \\
&+ \sum_{i=1}^{k} L_i'X\beta\beta'X'L_i + \sum_{i=1}^{k} L_i'cov(Y)L_i \\
&+ \sum_{i=1}^{k} \sigma_i^2 (L_i - Q_i)'Z_iZ_i'(L_i - Q_i) - \sum_{i=1}^{k} \sigma_i^2 L_i'Z_iZ_i'L_i\}.
\end{aligned}$$

Zauważmy, że ryzyko to zależy liniowo od macierzy $Z\mathrm{cov}(u)Z'$, $\mathrm{cov}(e)$ oraz $\mathrm{E}Y(\mathrm{E}Y)'$. Nie można go przedstawić jako funkcji $\mathrm{cov}(Y)$ i $\mathrm{E}Y(\mathrm{E}Y)'$. Między innymi dlatego na tym etapie nie jesteśmy w stanie bezpośrednio wykorzystać techniki LaMotte'a. Podobnie jak poprzednio zdefiniujmy w rozważanej klasie estymatorów następującą relację.

Definicja 2.1. *Estymator $L'Y$ jest lepszy niż L'_*Y, jeżeli*

$$R(L'Y) \leq R(L'_*Y)$$

dla każdych $Z\mathrm{cov}(u)Z'$, $\mathrm{cov}(e)$ oraz $\mathrm{E}Y(\mathrm{E}Y)'$ i jeżeli ostra nierówność jest spełniona dla co najmniej jednej kombinacji tych macierzy.

Definicja 2.2. *Estymator $L'Y$ nazywa się dopuszczalny w klasie estymatorów $\{(L_0, \ldots, L_k) : L_0 \in \mathcal{M}_{n\times t_0}, \ldots, L_k \in \mathcal{M}_{n\times t_k}\}$, jeżeli nie istnieje estymator lepszy niż $L'Y$.*

Aby móc zastosować teorię dopuszczalnej estymacji liniowej funkcji wartości oczekiwanej w ogólnym modelu liniowym do rozważanego problemu, należałoby w funkcji straty ,,przerzucić" efekty losowe z estymowanej funkcji do części ,,modelowej", tak by funkcja funkcja ryzyka w wyjściowym modelu pokrywała się z funkcją ryzyka odpowiedniego estymatora liniowego w modelu dodatkowo poszerzonym o efekty losowe. W ten sposób w obu modelach zbiór funkcji ryzyka pozostałby bez zmian. Zauważmy, że

$$\begin{bmatrix} L_0'Y - K'X\beta \\ L_1'Y - Q_1'Z_1u_1 \\ \vdots \\ L_k'Y - Q_k'Z_ku_k \end{bmatrix} = \begin{bmatrix} L_0 & L_1 & \cdots & L_k \\ \mathbf{0} & -Q_1 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & -Q_k \end{bmatrix}' \boldsymbol{Y} - \boldsymbol{K'X}\beta, \tag{2.3}$$

gdzie

$$\boldsymbol{Y} = \begin{pmatrix} Y \\ Z_1u_1 \\ \vdots \\ Z_ku_k \end{pmatrix}, \ \boldsymbol{K} = \begin{bmatrix} K & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix} \text{ oraz } \boldsymbol{X} = \begin{pmatrix} X \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix}.$$

Zatem możemy rozważać funkcję ryzyka estymatora $(L_0, \ldots, L_k)'Y$ jako funkcję ryzyka liniowego estymatora $\boldsymbol{L}'\boldsymbol{Y} = \begin{bmatrix} L_0 & L_1 & \cdots & L_k \\ \mathbf{0} & -Q_1 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & -Q_k \end{bmatrix}' \boldsymbol{Y}$ dla $\boldsymbol{K'X}\beta$ w następującym modelu

$$\boldsymbol{Y} \sim \left(\boldsymbol{X}\beta, \sum_{i=1}^{k+1} \sigma_i^2 \boldsymbol{V}_i \right), \tag{2.4}$$

gdzie

$$\boldsymbol{V}_i = (v_1 + v_{i+1})(v_1 + v_{i+1})' \otimes Z_i Z_i', \quad i = 1, \ldots, k,$$

$$\boldsymbol{V}_{k+1} = v_1 v_1' \otimes I_n$$

oraz v_i jest i-tym wersorem należącym do $\mathbb{R}^{k+1}$. Ponieważ $Q_1, \ldots, Q_k$ są ustalonymi macierzami, więc klasa estymatorów dla $\boldsymbol{K'X}\beta$ w powyższym modelu jest ograniczona do zbioru

$$\mathcal{E}_o = \{\boldsymbol{N}'\boldsymbol{Y} : \boldsymbol{N} \in \mathcal{L}_o\},$$

gdzie

$$\mathcal{L}_o = \{\boldsymbol{L}_o + \boldsymbol{\Pi}_o \boldsymbol{M} : \boldsymbol{M} \in \mathcal{M}_{(k+1)n \times (t_0 + \ldots + t_k)}\},$$

$$\boldsymbol{L}_o = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & -Q_1 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & -Q_k \end{bmatrix}$$

oraz

$$\boldsymbol{\Pi}_o = \begin{bmatrix} I_n & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix}.$$

Oczywiście $\mathcal{L}_o \subset \mathcal{M}_{(k+1)n\times(t_0+\ldots+t_k)}$ jest rozmaitością liniową. Zauważmy, że możemy ją równoważnie zapisać w postaci

$$\mathcal{L}_o = \boldsymbol{L}_o + \mathcal{R}(T^t_{\boldsymbol{\Pi}_o}),$$

gdzie $t = \sum_{i=0}^{k} t_i$. Równość (2.3) jest istotnym spostrzeżeniem do dalszych naszych rozważań. Sprowadza ona problem jednoczesnej estymacji liniowej funkcji wektora efektów stałych i losowych w modelu (2.1) do estymacji odpowiedniej liniowej funkcji wektora wartości oczekiwanych w modelu (2.4) w klasie $\mathcal{L}_o$, a to z kolei pozwala wykorzystać znane już narzędzia opisane w rozdziale 2. Ponieważ funkcja ryzyka estymatora $\boldsymbol{L'Y}$ dla $\boldsymbol{K'X}\beta$ daje się przedstawić w postaci

$$\mathrm{tr}[\boldsymbol{L}'\mathrm{cov}(\boldsymbol{Y})\boldsymbol{L} + (\boldsymbol{L}-\boldsymbol{K})'\mathrm{E}\boldsymbol{Y}(\mathrm{E}\boldsymbol{Y})'(\boldsymbol{L}-\boldsymbol{K})],$$

więc nową przestrzenią parametrów jest zbiór

$$\mathcal{T} = \left\{ (\sum_{i=1}^{k+1} \sigma_i^2 \boldsymbol{V}_i, \boldsymbol{X}\beta\beta'\boldsymbol{X}') : \sigma_1^2 \geq 0, \ldots, \sigma_k^2 \geq 0, \sigma_{k+1}^2 > 0, \beta \in \mathbb{R}^p \right\},$$

podczas gdy $\mathcal{W} = \mathrm{span}\mathcal{T}$ wyraża się wzorem

$$\mathcal{W} = \left\{ (\sum_{i=1}^{k+1} f_i \boldsymbol{V}_i, \boldsymbol{X}F\boldsymbol{X}') : f_1, \ldots, f_{k+1} \in \mathbb{R}, F \in \mathcal{M}_{p\times p}, F = F' \right\}.$$

Zauważmy, że dzięki opisanej ,,redukcji problemu" wystarczy podać warunki konieczne i dostateczne na to, aby estymator był dopuszczalny dla $\boldsymbol{K'X}\beta$ w klasie $\mathcal{L}_o$ w modelu (2.4).

Lemat 2.1. *Liniowy estymator jest dopuszczalny dla* $[(K'X\beta)', (Q_1'Z_1u_1)', \ldots, (Q_k'Z_ku_k)']'$ *w modelu* (2.1) *wtedy i tylko wtedy, gdy odpowiadający mu estymator jest dopuszczalny dla* $\boldsymbol{K'X}\beta$ *w klasie* $\mathcal{L}_o$ *w modelu* (2.4).

Dowód. Lemat jest natychmiastową konsekwencją równości (2.3) oraz tego, że

$$
\begin{aligned}
\boldsymbol{R}(\boldsymbol{L'Y};W) &= E\{[\boldsymbol{L'Y}-\boldsymbol{K'X}\beta]'[\boldsymbol{L'Y}-\boldsymbol{K'X}\beta]\}\\
&= E\left\{\begin{bmatrix} L_0'Y-K'X\beta \\ L_1'Y-Q_1'Z_1u_1 \\ \vdots \\ L_k'Y-Q_k'Z_ku_k \end{bmatrix}'\begin{bmatrix} L_0'Y-K'X\beta \\ L_1'Y-Q_1'Z_1u_1 \\ \vdots \\ L_k'Y-Q_k'Z_ku_k \end{bmatrix}\right\}.
\end{aligned}
$$

□

Do charakterystyki estymatorów dopuszczalnych dla $\boldsymbol{K'X}\beta$ w klasie $\mathcal{L}_o$ w modelu (2.4) wykorzystamy twierdzenia 1.2 oraz 1.3.

2.1. Przykłady liniowych modeli mieszanych

W tym podrozdziale opiszemy modele, które są szczególnymi przypadkami modelu (2.1), a dla których, w kolejnych rozdziałach rozprawy, podana zostanie charakterystyka estymatorów dopuszczalnych dla wektora (2.2).

Dokładny opis klasy modeli mieszanych wraz z licznymi przykładami zastosowań znajduje się między innymi w książkach Gnota (1991) oraz Searle'a i in. (1992).

Zrównoważony model losowy k-kierunkowej klasyfikacji hierarchicznej

Niech $Y_{i_1\ldots i_{k+1}}$, gdzie $i_j=1,2,\ldots,n_j$ dla $j=1,\ldots,k+1$, będzie zmienną losową o następującej strukturze

$$Y_{i_1\ldots i_{k+1}}=\beta+u_{1i_1}+u_{2i_1i_2}+\ldots+u_{ki_1\ldots i_k}+e_{i_1\ldots i_{k+1}},$$

gdzie β jest nieznanym parametrem, $u_{1i_1},\ldots,u_{ki_1\ldots i_k}$ oraz $e_{i_1\ldots i_{k+1}}$ są nieskorelowanymi zmiennymi losowymi z zerowymi wartościami oczekiwanymi i wariancjami odpowiednio $\sigma_1^2,\ldots,\sigma_k^2$ i σ_{k+1}^2. Przy tych założeniach, ustawiając zmienne $Y_{i_1 \ldots i_{k+1}}$ w porządku leksykograficznym w n-wymiarowy wektor Y, gdzie $n = n_1 \cdot n_2 \cdot \ldots \cdot n_{k+1}$, dostajemy *zrównoważony model k-kierunkowej klasyfikacji hierarchicznej* będący szczególnym przypadkiem modelu (2.1), dla którego $X = \mathbf{1}_{n_1} \otimes \ldots \otimes \mathbf{1}_{n_{k+1}} = \mathbf{1}_n$ oraz $Z_i = I_{n_1} \otimes \ldots \otimes I_{n_i} \otimes \mathbf{1}_{n_{i+1}} \otimes \ldots \otimes \mathbf{1}_{n_{k+1}}$, tzn.

$$Y \sim \left(\mathbf{1}_n \beta,\ \sum_{i=1}^{k} \sigma_i^2 (I_{n_1} \otimes \ldots \otimes I_{n_i} \otimes \mathbf{J}_{n_{i+1}} \otimes \ldots \otimes \mathbf{J}_{n_{k+1}}) + \sigma_{k+1}^2 I_n \right), \tag{2.5}$$

gdzie $\mathbf{1}_a$ oznacza wektor składający się z a jedynek oraz $\mathbf{J}_a = \mathbf{1}_a \mathbf{1}_a'$.

Zrównoważony model losowy k-kierunkowej klasyfikacji krzyżowej

Niech $Y_{i_1 \ldots i_{k+1}}$, gdzie $i_j = 1, 2, \ldots, n_j$ dla $j = 1, \ldots, k+1$, będzie zmienną losową o następującej strukturze

$$Y_{i_1 \ldots i_{k+1}} = \beta + u_{1i_1} + \ldots + u_{ki_k} + e_{i_1 \ldots i_{k+1}},$$

gdzie β jest nieznanym parametrem, $u_{1i_1}, \ldots, u_{ki_k}$ oraz $e_{i_1 \ldots i_{k+1}}$, są nieskorelowanymi zmiennymi losowymi z zerowymi wartościami oczekiwanymi i wariancjami odpowiednio $\sigma_1^2, \ldots, \sigma_k^2$ i σ_{k+1}^2. Przy tych założeniach, ustawiając zmienne $Y_{i_1, \ldots, i_{k+1}}$ w porządku leksykograficznym w n-wymiarowy wektor Y, gdzie $n = n_1 \cdot n_2 \cdot \ldots \cdot n_{k+1}$, uzyskujemy tzw. *zrównoważony model k-kierunkowej klasyfikacji krzyżowej*, który możemy wyrazić w następujący sposób

$$Y \sim \left(\mathbf{1}_n \beta, \sum_{i=1}^{k} \sigma_i^2 (\mathbf{J}_{n_1} \otimes \ldots \otimes \mathbf{J}_{n_{i-1}} \otimes I_{n_i} \otimes \mathbf{J}_{n_{i+1}} \otimes \ldots \otimes \mathbf{J}_{n_{k+1}}) + \sigma_{k+1}^2 I_n \right). \tag{2.6}$$

Zrównoważony model losowy dwukierunkowej klasyfikacji krzyżowej z interakcją

Niech $Y_{i_1 i_2 i_3}$, gdzie $i_j = 1, 2, \ldots, n_j$ dla $j = 1, 2, 3$, będzie zmienną losową o następującej strukturze

$$Y_{i_1 i_2 i_3} = \beta + u_{1i_1} + u_{2i_2} + u_{3i_1 i_2} + e_{i_1 i_2 i_3},$$

gdzie β jest nieznanym parametrem, $u_{1i_1}, u_{2i_2}, u_{3i_1i_2}$ oraz $e_{i_1i_2i_3}$ są nieskorelowanymi zmiennymi losowymi z zerowymi wartościami oczekiwanymi i wariancjami odpowiednio $\sigma_1^2, \sigma_2^2, \sigma_3^2$ i σ_4^2. Postępując ze zmiennymi $Y_{i_1i_2i_3}$ tak, jak w dwóch poprzednich sytuacjach, uzyskujemy dla n-wymiarowego wektora Y następujący model

$$Y \sim \big(\mathbf{1}_n\beta, \sigma_1^2 I_{n_1} \otimes \mathbf{J}_{n_2} \otimes \mathbf{J}_{n_3} + \sigma_2^2 \mathbf{J}_{n_1} \otimes I_{n_2} \otimes \mathbf{J}_{n_3} + \sigma_3^2 I_{n_1} \otimes I_{n_2} \otimes \mathbf{J}_{n_3} + \sigma_4^2 I_n\big), \quad (2.7)$$

gdzie $n = n_1 \cdot n_2 \cdot n_3$. Model ten nazywany jest *zrównoważonym modelem losowym dwukierunkowej klasyfikacji krzyżowej z interakcją*.

Model Rafajłowicza

Gnot, Rafajłowicz i Urbańska-Motyka (2001) użyli następującego modelu do opisu wartości pomiaru wykonanego przez j-ty sensor

$$y_j = N_j v + e_j,$$

gdzie $j = 1, \ldots, n_2$, v oznacza intensywność źródła, z którego został wysłany sygnał, N_j wpływ źródła sygnału na j-ty sensor, natomiast e_j jest błędem losowym tego pomiaru. Przy tych oznaczeniach mamy, że

$$y = Nv + e,$$

gdzie $y = (y_1, \ldots, y_{n_2})'$, $N = (N_1, \ldots, N_{n_2})'$ oraz $e = (e_1, \ldots, e_{n_2})'$. O wektorze N, zwanym wektorem odpowiedzi sensora, zakładamy, że jest dany. Dodatkowo o nieskorelowanych zmiennych losowych v oraz e_j zakładamy, że mają wartości oczekiwane odpowiednio β oraz 0, natomiast wariancje $\sigma_1^2 \geq 0$ i $\sigma_2^2 > 0$. Założenia te implikują, że

$$Ey = N\beta \;\text{ oraz }\; \operatorname{cov}(y) = \sigma_1^2 NN' + \sigma_2^2 I_{n_2}.$$

Zatem do opisu n_1 niezależnych pomiarów wykonanych na każdym spośród n_2 sensorów użyjemy następującego modelu

$$Y = \mathrm{vec}(y^{(1)}, y^{(2)}, \ldots, y^{(n_1)}),$$

gdzie $y^{(i)} = (y_{i1}, \ldots, y_{in_2})'$, y_{ij} oznacza i-ty pomiar j-tego sensora, natomiast $\mathrm{vec}(A)$ oznacza operator macierzowy tworzący wektor kolumnowy z kolumn macierzy A ustawianych kolejno jedna pod drugą. O wektorze losowym $y^{(i)}$ zakładamy, że $Ey^{(i)} = N\beta$ oraz $\mathrm{cov}(y^{(i)}) = \sigma_1^2 NN' + \sigma_2^2 I_{n_2}$, co z kolei implikuje, że

$$Y \sim \left((\mathbf{1}_{n_1} \otimes N)\beta,\ \sigma_1^2(I_{n_1} \otimes NN') + \sigma_2^2 I_n\right), \tag{2.8}$$

gdzie $n = n_1 \cdot n_2$.

Zauważmy, że kiedy $N = \mathbf{1}_{n_2}$, to otrzymujemy *zrównoważony model losowy jednokierunkowej klasyfikacji*, który jest jednocześnie szczególnym przypadkiem modeli (2.5) oraz (2.6) dla $k = 1$. Symbolicznie będziemy go zapisywać w postaci

$$Y \sim \left(\mathbf{1}_n\beta,\ \sigma_1^2(I_{n_1} \otimes \mathbf{1}_{n_2}\mathbf{1}_{n_2}') + \sigma_2^2 I_n\right). \tag{2.9}$$

3. Jawna charakterystyka estymatorów dopuszczalnych w wybranych modelach liniowych

W tym rozdziale podane zostaną w jawnej postaci warunki konieczne i dostateczne na to, aby liniowy estymator $(L_0, L_1, \ldots, L_k)'Y$ był dopuszczalny dla wektora $[(K'X\beta)', (Q_1'Z_1u_1)', \ldots, (Q_k'Z_ku_k)']'$ w modelach (2.5), (2.6) oraz (2.7). Na mocy lematu 2.1 zostaną one podane w terminach macierzy opisujących modele będące odpowiednikami modelu (2.4), w których możemy wykorzystać wyniki LaMotte'a opisane w rozdziale 1. Zauważmy, że dla tych modeli mamy

$$\mathcal{T} = \left\{ (\sum_{i=1}^{k+1} \sigma_i^2 \boldsymbol{V}_i, \beta^2 \boldsymbol{X}\boldsymbol{X}') : \sigma_1^2 \geq 0, \ldots, \sigma_k^2 \geq 0, \sigma_{k+1}^2 > 0, \beta \in \mathbb{R} \right\},$$

natomiast

$$[\mathcal{T}] = \left\{ (\sum_{i=1}^{k+1} s_i \boldsymbol{V}_i, s_0 \boldsymbol{X}\boldsymbol{X}') : s_0 \geq 0, \ldots, s_{k+1} \geq 0 \right\}.$$

3.1. Zrównoważony model losowy k-kierunkowej klasyfikacji hierarchicznej

W tym podrozdziale podamy jawną postać wszystkich liniowych estymatorów dopuszczalnych dla $\boldsymbol{K'X}\beta$ w klasie $\mathcal{L}_o$ na zbiorze parametrów $\mathcal{T}$ w modelu (2.4), w którym postaci macierzy $\boldsymbol{X}$ oraz $\boldsymbol{V}_i$ zależą od odpowiednich macierzy modelu (2.5), tj. $X = \mathbf{1}_n$, natomiast $Z_i = I_{n_1} \otimes \ldots \otimes I_{n_i} \otimes \mathbf{1}_{n_{i+1}} \otimes \ldots \otimes \mathbf{1}_{n_{k+1}}$ dla $i = 1, \ldots, k$. Tym samym scharakteryzujemy liniowe estymatory dopuszczalne dla wektora (2.2) w modelu (2.5). Zaczniemy od podania warunków koniecznych i dostatecznych na to, aby estymator $\boldsymbol{L'Y}$ był jednoznacznie lokalnie optymalnym estymatorem dla $\boldsymbol{K'X}\beta$ w punkcie $W = (W_1, W_2) \in [\mathcal{T}]$.

Dla uproszczenia notacji, niech $Z_0 = \mathbf{1}_n$ oraz $Z_{k+1} = I_n$. Kładąc $p_i = n_{i+1} \cdot \ldots \cdot n_{k+1}$ dla $i = 0, \ldots, k$ oraz $p_{k+1} = 1$, definiujemy następujące macierze

$$E_0 = \frac{1}{p_0} Z_0 Z_0'$$

oraz

$$E_i = \frac{1}{p_i} Z_i Z_i' - \frac{1}{p_{i-1}} Z_{i-1} Z_{i-1}' \ \text{ dla } \ i = 1, \ldots, k+1.$$

Zauważmy, że $E_0, \ldots, E_{k+1}$ są idempotentnymi i ortogonalnymi macierzami takimi, że

$$Z_i Z_i' = p_i \sum_{j=0}^{i} E_j \ \text{ dla } \ i = 0, \ldots, k.$$

Lemat 3.1. *Estymator $\boldsymbol{L'Y}$ jest jednoznacznie lokalnie optymalny w punkcie $(\sum_{i=1}^{k+1} s_i \boldsymbol{V}_i, s_0 \boldsymbol{XX'})$ z $[\mathcal{T}]$ w klasie $\mathcal{L}_o$ w modelu (2.4) odpowiadającym modelowi (2.5) wtedy i tylko wtedy, gdy $s_0 \geq 0, \ldots, s_k \geq 0, s_{k+1} > 0$ oraz*

$$\boldsymbol{L} = \begin{bmatrix} L_0 & L_1 & \cdots & L_k \\ \mathbf{0} & -Q_1 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & -Q_k \end{bmatrix},$$

gdzie

$$
\begin{aligned}
L_0 &= a_0 E_0 K, \\
L_i &= a_i \left\{ E_i + \sum_{j=0}^{i-1} \left[\prod_{l=j}^{i-1} (1-a_l) \right] E_j \right\} Q_i \;\; \textit{dla} \;\; i = 1, \ldots, k, \\
a_i &= \frac{s_i p_i}{\sum\limits_{j=i}^{k+1} s_j p_j} \;\; \textit{dla} \;\; i = 0, \ldots, k.
\end{aligned} \tag{3.1}
$$

Dowód. Estymator $\boldsymbol{L'Y}$ jest lokalnie optymalny w punkcie $(\sum_{i=1}^{k+1} s_i \boldsymbol{V}_i, s_0 \boldsymbol{XX'})$ z $[\mathcal{T}]$ w klasie $\mathcal{L}_o$ wtedy i tylko wtedy, gdy $s_j \geq 0$ dla $j = 0, \ldots, k+1$ oraz

$$
\boldsymbol{\Pi}_o \left(\sum_{i=1}^{k+1} s_i \boldsymbol{V}_i + s_0 \boldsymbol{XX'} \right) \boldsymbol{L} = s_0 \boldsymbol{\Pi}_o \boldsymbol{XX'K}.
$$

Równoważnie równanie to możemy zapisać w postaci

$$
\begin{bmatrix}
ML_0 & ML_1 - s_1 Z_1 Z_1' Q_1 & \cdots & ML_k - s_k Z_k Z_k' Q_k \\
\mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\
\vdots & \vdots & \ddots & \vdots \\
\mathbf{0} & \mathbf{0} & \cdots & \mathbf{0}
\end{bmatrix}
=
\begin{bmatrix}
s_0 XX'K & \mathbf{0} & \cdots & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\
\vdots & \vdots & \ddots & \vdots \\
\mathbf{0} & \mathbf{0} & \cdots & \mathbf{0}
\end{bmatrix},
$$

gdzie $M = \sum_{i=0}^{k+1} s_i Z_i Z_i'$. Oczywiście równanie to ma dokładnie jedno rozwiązanie ze względu na $L_0, \ldots, L_k$ wtedy i tylko wtedy, gdy macierz M jest nieosobliwa, czyli wtedy i tylko wtedy, gdy $s_{k+1} > 0$. Przy użyciu $E_0, \ldots, E_{k+1}$ możemy zapisać M następująco

$$
M = \sum_{i=0}^{k+1} s_i \left(p_i \sum_{j=0}^{i} E_j \right) = \sum_{i=0}^{k+1} \left(\sum_{j=i}^{k+1} s_j p_j \right) E_i = \sum_{i=0}^{k+1} w_i E_i.
$$

Ponieważ $E_0, \ldots, E_{k+1}$ są ortogonalne i idempotentne, więc

$$
M^{-1} = \sum_{i=0}^{k+1} \frac{1}{w_i} E_i.
$$

Zatem

$$L_0 = s_0 M^{-1} X X' K = \frac{s_0 p_0}{\sum_{j=0}^{k+1} s_j p_j} E_0 K$$

oraz

$$\begin{aligned} L_i &= s_i M^{-1} Z_i Z_i' Q_i = s_i p_i M^{-1} \left(\sum_{j=0}^{i} E_j \right) Q_i = \left(\sum_{j=0}^{i} \frac{s_i p_i}{\sum_{l=j}^{k+1} s_l p_l} E_j \right) Q_i \\ &= a_i \left(E_i + \sum_{j=0}^{i-1} \frac{\sum_{l=i}^{k+1} s_l p_l}{\sum_{l=j}^{k+1} s_l p_l} E_j \right) Q_i \text{ dla } i = 1, \ldots, k. \end{aligned}$$

Zauważmy, że

$$\begin{aligned} \frac{\sum_{l=i}^{k+1} s_l p_l}{\sum_{l=j}^{k+1} s_l p_l} &= \frac{\sum_{l=j+1}^{k+1} s_l p_l}{\sum_{l=j}^{k+1} s_l p_l} \cdot \frac{\sum_{l=j+2}^{k+1} s_l p_l}{\sum_{l=j+1}^{k+1} s_l p_l} \cdot \ldots \cdot \frac{\sum_{l=i-1}^{k+1} s_l p_l}{\sum_{l=i-2}^{k+1} s_l p_l} \cdot \frac{\sum_{l=i}^{k+1} s_l p_l}{\sum_{l=i-1}^{k+1} s_l p_l} \\ &= (1 - a_j)(1 - a_{j-1}) \cdot \ldots \cdot (1 - a_{i-2})(1 - a_{i-1}), \end{aligned}$$

co kończy dowód lematu. □

Twierdzenie 3.1. *Estymator* $\boldsymbol{L'Y}$ *jest dopuszczalny dla* $\boldsymbol{K'X}\beta$ *w klasie* $\mathcal{L}_o$ *na* $\mathcal{T}$ *w modelu* (2.4) *odpowiadającym modelowi* (2.5) *wtedy i tylko wtedy, gdy*

$$\boldsymbol{L} \in \left\{ \begin{bmatrix} L_0(a_0) & L_1(a_0, a_1) & \cdots & L_k(a_0, \ldots, a_k) \\ \mathbf{0} & -Q_1 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & -Q_k \end{bmatrix} : a_i \in [0, 1], i = 0, \ldots, k \right\}, \quad (3.2)$$

gdzie

$$L_0(a_0) = a_0 E_0 K$$

oraz

$$L_i(a_0,\ldots,a_i) = a_i \left\{ E_i + \sum_{j=0}^{i-1} \left[\prod_{l=j}^{i-1} (1-a_l) \right] E_j \right\} Q_i \quad dla \ \ i = 1,\ldots,k.$$

Dowód. Warunek konieczny. Zauważmy, że dla każdych ustalonych wartości $s_{i+1} \geq 0, \ldots, s_k \geq 0$ oraz $s_{k+1} > 0$ współczynnik a_i dany wzorem (3.1) przebiega przedział $[0,1)$, gdy $s_i \in [0,+\infty)$ dla $i = 0,\ldots,k$. Korzystając z lematu 3.1 mamy, że zbiór (3.2) jest domknięciem zbioru

$$\{\boldsymbol{L}_* : \boldsymbol{L}_*'\boldsymbol{Y} \text{ jest jednoznacznie lokalnie optymalnym w punkcie z } [\mathcal{T}] \text{ w klasie } \mathcal{L}_o\}.$$

A zatem na mocy twierdzenia 1.3 pierwsza część dowodu jest zakończona.

Chcąc udowodnić warunek dostateczny, użyjemy krokowej metody zaproponowanej przez LaMotte'a. Liczba kroków będzie zależała od tych współczynników spośród $a_0, a_1, \ldots, a_k$, które są równe 1.
Niech $\mathcal{I} = \{i \in \{0,1,\ldots,k\} : a_i = 1\}$,

$$m_1 = \begin{cases} k+1, & \text{gdy } \mathcal{I} = \emptyset, \\ \min \mathcal{I}, & \text{w przeciwnym przypadku} \end{cases}$$

oraz niech $\mathcal{L} = \boldsymbol{L} + \mathcal{R}(T^t_{\boldsymbol{\Pi}_o})$, gdzie $t = \sum_{i=0}^{k} t_i$. Co oznacza, że $\mathcal{L} = \mathcal{L}_o$. Dalej, niech $W = (W_1, W_2)$ będzie punktem z $[\mathcal{T}]$ określonym następująco

$$W = \begin{cases} \left(\boldsymbol{0}, \frac{1}{p_0}\boldsymbol{X}\boldsymbol{X}'\right), & \text{gdy } m_1 = 0, \\ \left(\frac{1}{p_1}\boldsymbol{V}_1, \frac{a_0}{p_0(1-a_0)}\boldsymbol{X}\boldsymbol{X}'\right), & \text{gdy } m_1 = 1, \\ \left(\sum_{i=1}^{m_1-1} \frac{a_i}{p_i \prod_{j=i}^{m_1-1}(1-a_j)}\boldsymbol{V}_i + \frac{1}{p_{m_1}}\boldsymbol{V}_{m_1}, \frac{a_0}{p_0 \prod_{j=0}^{m_1-1}(1-a_j)}\boldsymbol{X}\boldsymbol{X}'\right), & \text{w przeciwnym przypadku.} \end{cases}$$

Dalszą część dowodu przedstawimy w postaci następującego algorytmu.

START:

Ponieważ $\mathbf{\Pi}(W_1+W_2)\boldsymbol{L} = \mathbf{\Pi}W_2\boldsymbol{K}$, więc $\boldsymbol{L}'\boldsymbol{Y}$ jest lokalnie optymalny w punkcie W w klasie $\mathcal{L}$.

Jeżeli $m_1 = k+1$;

W tym przypadku $\mathcal{R}(\mathbf{\Pi}(W_1+W_2)\mathbf{\Pi}) = \mathcal{R}(\mathbf{\Pi})$, czyli $\mathcal{B}(W|\mathcal{L})$ =$\{\boldsymbol{L}\}$. Stąd $\boldsymbol{L}'\boldsymbol{Y}$ jest dopuszczalny w $\mathcal{L}_o$.

W przeciwnym razie;

Niech $\mathcal{B}(W|\mathcal{L})$ będzie nową rozmaitością $\mathcal{L}$, którą możemy przedstawić w postaci

$$\mathcal{L} = \boldsymbol{L} + \mathcal{R}(T^t_{\mathbf{\Pi}}),$$

gdzie

$$\mathbf{\Pi} = \begin{bmatrix} I_n - \sum\limits_{i=0}^{m_1} E_i & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix} \in \mathcal{M}_{(k+1)n\times(k+1)n}.$$

Niech $\mathcal{I}\setminus\{m_1\}$ będzie nowym zbiorem $\mathcal{I}$,

$$m_2 = \begin{cases} k+1, & \text{gdy } \mathcal{I} = \emptyset, \\ \min\mathcal{I}, & \text{w przeciwnym przypadku} \end{cases}$$

oraz niech $W = (W_1, W_2)$ będzie punktem z $[\mathcal{T}]$ określonym w sposób

$$W = \begin{cases} \left(\frac{1}{p_{m_2}}\boldsymbol{V}_{m_2}, \mathbf{0}\right), & \text{gdy } m_2 = m_1+1, \\ \left(\sum\limits_{i=m_1}^{m_2-1} \frac{a_i}{p_i \prod\limits_{j=i}^{m_2-1}(1-a_j)}\boldsymbol{V}_i + \frac{1}{p_{m_2}}\boldsymbol{V}_{m_2}, \mathbf{0}\right), & \text{w przeciwnym razie.} \end{cases}$$

I ostatecznie niech m_2 będzie nowym m_1.

Przechodzimy do START;

□

Zauważmy, że procedura ta musi zakończyć się po skończonej liczbie kroków, ponieważ po każdym wykonanym kroku wymiar kolejnych rozmaitości liniowych maleje. W skrajnym przypadku, gdy każdy ze współczynników $a_0, \ldots, a_k$ jest równy 1, musimy wykonać $(k+1)+1$ kroków.

3.2. Zrównoważony model losowy k-kierunkowej klasyfikacji krzyżowej

W tym podrozdziale podamy charakterystykę liniowych estymatorów dla $\boldsymbol{K'X}\beta$ w klasie $\mathcal{L}_o$ na zbiorze parametrów $\mathcal{T}$ w modelu (2.4), w którym postaci macierzy $\boldsymbol{X}$ oraz $\boldsymbol{V}_i$ są determinowane odpowiednimi macierzami modelu (2.6), tzn. macierzami $X = \mathbf{1}_n$, natomiast $Z_i = \mathbf{1}_{n_1} \otimes \ldots \otimes \mathbf{1}_{n_{i-1}} \otimes I_{n_i} \otimes \mathbf{1}_{n_{i+1}} \otimes \ldots \otimes \mathbf{1}_{n_{k+1}}$ dla $i = 1, \ldots, k$. Tak, jak poprzednio, celem uproszczenia zapisu, niech $p_0 = n$, $Z_0 = \mathbf{1}_n$, $p_{k+1} = 1$ oraz $Z_{k+1} = I_n$. Kładąc $p_i = \prod_{j \neq i} n_j$ dla $i = 1, \ldots, k$ i definiując

$$\begin{aligned} E_0 &= \frac{1}{p_0} Z_0 Z_0', \\ E_i &= \frac{1}{p_i} Z_i Z_i' - E_0 \ \text{ dla } \ i = 1, \ldots, k, \\ E_{k+1} &= Z_{k+1} Z_{k+1}' - (E_0 + \cdots + E_k) \end{aligned}$$

otrzymujemy, że $E_0, \ldots, E_{k+1}$ są idempotentnymi i ortogonalnymi macierzami, a macierze $Z_0 Z_0', \ldots, Z_{k+1} Z_{k+1}'$ są ich liniowymi kombinacjami. Aby móc schara- kteryzować liniowe estymatory dopuszczalne dla $\boldsymbol{K'X}\beta$ w klasie $\mathcal{L}_o$ na zbiorze $\mathcal{T}$ w modelu (2.4), który odpowiada modelowi (2.6), udowodnimy najpierw następujący lemat.

Lemat 3.2. *Estymator* $\boldsymbol{L}'\boldsymbol{Y}$ *jest jednoznacznie lokalnie optymalny w punkcie* $(\sum_{i=1}^{k+1} s_i\boldsymbol{V}_i, s_0\boldsymbol{X}\boldsymbol{X}')$ *z* $[\mathcal{T}]$ *w klasie* $\mathcal{L}_o$ *w modelu* (2.4) *odpowiadającym modelowi* (2.6) *wtedy i tylko wtedy, gdy* $s_0 \geq 0, \ldots, s_k \geq 0, s_{k+1} > 0$ *oraz*

$$\boldsymbol{L} = \begin{bmatrix} L_0 & L_1 & \cdots & L_k \\ \mathbf{0} & -Q_1 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & -Q_k \end{bmatrix},$$

gdzie

$$\begin{aligned} L_0 &= a_0 E_0 K, \\ L_i &= a_i \left[E_i + (1-a_0)\frac{\frac{1}{1-a_i}}{\sum\limits_{j=1}^{k}\frac{1}{1-a_j} - (k-1)} E_0 \right] Q_i \quad \textit{dla} \;\; i = 1, \ldots, k, \\ a_0 &= \frac{s_0 p_0}{\sum\limits_{j=0}^{k+1} s_j p_j}, \qquad (3.3) \\ a_i &= \frac{s_i p_i}{s_i p_i + s_{k+1}} \quad \textit{dla} \;\; i = 1, \ldots, k. \qquad (3.4) \end{aligned}$$

Dowód. Niech $\left(\sum_{i=1}^{k+1} s_i\boldsymbol{V}_i, s_0\boldsymbol{X}\boldsymbol{X}'\right)$ będzie punktem z $[\mathcal{T}]$ takim, że $s_0 \geq 0, \ldots, s_{k+1} \geq 0$. Tak, jak w dowodzie lematu 3.1 estymator $\boldsymbol{L}'\boldsymbol{Y}$ jest jednoznacznie lokalnie optymalnym estymatorem wtedy i tylko wtedy, gdy $s_{k+1} > 0$ oraz

$$\begin{aligned} L_0 &= s_0 M^{-1} X X' K, \\ L_i &= s_i M^{-1} Z_i Z_i' Q_i \;\; \text{dla} \;\; i = 1, \ldots, k, \end{aligned}$$

gdzie $M = \sum_{i=0}^{k+1} s_i Z_i Z_i'$ można wyrazić wzorem

$$M = \left(\sum_{i=0}^{k+1} s_i p_i\right) E_0 + \sum_{i=1}^{k} (s_i p_i + s_{k+1}) E_i + s_{k+1} E_{k+1}.$$

Zatem

$$L_0 = \frac{s_0 p_0}{\sum\limits_{j=0}^{k+1} s_j p_j} E_0 K,$$

natomiast

$$\begin{aligned} L_i &= \left(\frac{s_i p_i}{\sum\limits_{j=0}^{k+1} s_j p_j} E_0 + \frac{s_i p_i}{s_i p_i + s_{k+1}} E_i \right) Q_i \\ &= \frac{s_i p_i}{s_i p_i + s_{k+1}} \left(E_i + \frac{s_i p_i + s_{k+1}}{\sum\limits_{j=0}^{k+1} s_j p_j} E_0 \right) Q_i \ \text{ dla } \ i = 1, \ldots, k. \end{aligned}$$

Zauważmy, że

$$\begin{aligned} \frac{s_i p_i + s_{k+1}}{\sum\limits_{j=0}^{k+1} s_j p_j} &= \frac{\sum\limits_{j=1}^{k+1} s_j p_j}{\sum\limits_{j=0}^{k+1} s_j p_j} \cdot \frac{s_i p_i + s_{k+1}}{\sum\limits_{j=1}^{k} (s_j p_j + s_{k+1}) - k s_{k+1} + s_{k+1}} \\ &= (1 - a_0) \frac{\frac{s_i p_i + s_{k+1}}{s_{k+1}}}{\sum\limits_{j=1}^{k} \frac{s_j p_j + s_{k+1}}{s_{k+1}} - k + 1} \\ &= (1 - a_0) \frac{\frac{1}{1 - a_i}}{\sum\limits_{j=1}^{k} \frac{1}{1 - a_j} - (k - 1)}, \end{aligned}$$

co kończy dowód. $\square$

Twierdzenie 3.2. *Estymator* $\boldsymbol{L}'\boldsymbol{Y}$ *jest dopuszczalny dla* $\boldsymbol{K}'\boldsymbol{X}\beta$ *w klasie* $\mathcal{L}_o$ *na* $\mathcal{T}$ *w modelu* (2.4) *odpowiadającym modelowi* (2.6) *wtedy i tylko wtedy, gdy*

$$\boldsymbol{L} \in \left\{ \begin{bmatrix} L_0(a_0) & L_1(a_1, A_1) & \cdots & L_k(a_k, A_k) \\ \mathbf{0} & -Q_1 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & -Q_k \end{bmatrix} : a_i \in [0, 1], i = 0, \ldots, k \right\}, \quad (3.5)$$

gdzie

$$\begin{aligned} L_0(a_0) &= a_0 E_0 K, \\ L_i(a_i, A_i) &= a_i[E_i + (1-a_0)A_i E_0]Q_i \ \ \textit{dla} \ \ i = 1, \ldots, k \end{aligned}$$

oraz

$$A_i = \frac{\frac{1}{1-a_i}}{\sum\limits_{j=1}^{k} \frac{1}{1-a_j} - (k-1)} \ \ \textit{dla} \ \ a_1 \in [0,1), \ldots, a_k \in [0,1), \tag{3.6}$$

w przeciwnym razie

$A_1, \ldots, A_k$ *są nieujemnymi liczbami takimi, że* $\sum\limits_{i=1}^{k} A_i = \sum\limits_{i \in \mathcal{I}} A_i = 1$, *gdzie* $\mathcal{I} = \{i \in \{1, \ldots, k\} : a_i = 1\}$.

Dowód. Warunek konieczny. Zauważmy, że a_i określone wzorem (3.3) oraz (3.4) przyjmują każdą wartość z przedziału $[0,1)$ dla dowolnych $s_0 \geq 0, \ldots, s_k \geq 0$ oraz $s_{k+1} > 0$. Oczywiście dla A_i danego formułą (3.6) zachodzi $1 < \sum\limits_{i=1}^{k} A_i \leq k$. Ponieważ $\sum\limits_{j=1}^{k} \frac{1}{1-a_j} \to +\infty$, gdy $a_j \to 1$ dla co najmniej jednego j, tj. gdy $\mathcal{I} \neq \emptyset$, więc $\sum\limits_{i=1}^{k} A_i \to 1$ oraz $A_j \to 0$ dla $j \notin \mathcal{I}$. Wszystko to implikuje, że zbiór (3.5) jest domknięciem zbioru

$$\{\boldsymbol{L}_* : \boldsymbol{L}_*'\boldsymbol{Y} \text{ jest jednoznacznie lokalnie optymalnym w punkcie z } [\mathcal{T}] \text{ w klasie } \mathcal{L}_o\}.$$

Zatem, na mocy twierdzenia 1.3, pierwsza część dowodu jest zakończona.

Warunek dostateczny udowodnimy w następujący sposób.

(1) Jeżeli $a_i \in [0,1)$ dla każdego $i = 0, 1, \ldots, k$, to wówczas możemy sprawdzić, że $\boldsymbol{L}'\boldsymbol{Y}$ jest jednoznacznie lokalnie optymalny w punkcie

$$\left(\sum_{i=1}^{k} \frac{a_i}{p_i(1-a_i)} \boldsymbol{V}_i + \boldsymbol{V}_{k+1}, \frac{a_0}{p_0(1-a_0)} \left(\sum_{i=1}^{k} \frac{a_i}{1-a_i} + 1 \right) \boldsymbol{X}\boldsymbol{X}' \right) \in [\mathcal{T}]$$

w klasie $\mathcal{L}_o$ i oczywiście dopuszczalny dla $\boldsymbol{K}'\boldsymbol{X}\beta$ w tej klasie.

(2) Jeżeli $a_0 = 1$, to wtedy $\boldsymbol{X}\boldsymbol{X}'\boldsymbol{L} = \boldsymbol{X}\boldsymbol{X}'\boldsymbol{K}$, czyli $\boldsymbol{L}'\boldsymbol{Y}$ jest lokalnie optymalny w punkcie $(\boldsymbol{0}, \boldsymbol{X}\boldsymbol{X}') \in [\mathcal{T}]$ w klasie $\mathcal{L}_o$, ale nie jest jednoznaczny. Zbiór

$\mathcal{B}((\mathbf{0}, \boldsymbol{X}\boldsymbol{X}')|\mathcal{L}_o)$ możemy przedstawić w postaci

$$\mathcal{L}_1 = \boldsymbol{L} + \mathcal{R}(T^t_{\mathbf{\Pi}_1}),$$

gdzie

$$\mathbf{\Pi}_1 = \begin{bmatrix} I_n - E_0 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix} \in \mathcal{M}_{(k+1)n \times (k+1)n}.$$

Następnie, jeżeli $\mathcal{I} = \emptyset$, to mamy spełnione równanie $\mathbf{\Pi}_1 W_1 \boldsymbol{L} = \mathbf{0}$, gdzie $W_1 = \sum\limits_{i=1}^{k} \frac{a_i}{p_i(1-a_i)} \boldsymbol{V}_i + \boldsymbol{V}_{k+1}$. Stąd $\boldsymbol{L}'\boldsymbol{Y}$ jest lokalnie optymalny w punkcie $(W_1, \mathbf{0}) \in [\mathcal{T}]$, a ponieważ $\mathcal{R}(\mathbf{\Pi}_1 W_1 \mathbf{\Pi}_1) = \mathcal{R}(\mathbf{\Pi}_1)$, więc $\boldsymbol{L}'\boldsymbol{Y}$ jest dopuszczalny w klasie $\mathcal{L}_o$.

Jeżeli $\mathcal{I} \neq \emptyset$, to biorąc $(\sum\limits_{i \in \mathcal{I}} \frac{1}{p_i} \boldsymbol{V}_i, \mathbf{0}) \in [\mathcal{T}]$ mamy $\mathbf{\Pi}_1 (\sum\limits_{i \in \mathcal{I}} \frac{1}{p_i} \boldsymbol{V}_i) \boldsymbol{L} = \mathbf{0}$, co oznacza, że $\boldsymbol{L}'\boldsymbol{Y}$ jest lokalnie optymalny w tym punkcie w klasie $\mathcal{L}_1$, ale nie jest jednoznaczny. Rozmaitość liniowa $\mathcal{B}((\sum\limits_{i \in \mathcal{I}} \frac{1}{p_i} \boldsymbol{V}_i, \mathbf{0})|\mathcal{L})$ może być wyrażona następująco

$$\mathcal{L}_2 = \boldsymbol{L} + \mathcal{R}(T^t_{\mathbf{\Pi}_2}),$$

gdzie

$$\mathbf{\Pi}_2 = \begin{bmatrix} I_n - (E_0 + \sum\limits_{i \in \mathcal{I}} E_i) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix} \in \mathcal{M}_{(k+1)n \times (k+1)n}.$$

Ostatecznie, jeżeli weźmiemy punkt $(W_1, \mathbf{0}) \in [\mathcal{T}]$, gdzie $W_1 = (\sum\limits_{i \notin \mathcal{I}} \frac{a_i}{p_i(1-a_i)} \boldsymbol{V}_i + \boldsymbol{V}_{k+1})$, to możemy sprawdzić, że $\mathbf{\Pi}_2 W_1 \boldsymbol{L} = \mathbf{0}$. Ponieważ $\mathcal{R}(\mathbf{\Pi}_2 W_1 \mathbf{\Pi}_2) = \mathcal{R}(\mathbf{\Pi}_2)$, więc $\boldsymbol{L}'\boldsymbol{Y}$ jest dopuszczalny w klasie $\mathcal{L}_o$.

(3) Załóżmy teraz, że $a_0 \in [0, 1)$ oraz $\mathcal{I} \neq \emptyset$. Niech $\mathcal{I}_1 = \{i \in \mathcal{I} : A_i > 0\}$ i niech $\mathcal{I}_2 = \mathcal{I} \backslash \mathcal{I}_1$. Oczywiście $\sum\limits_{i \in \mathcal{I}_1} A_i = 1$. Biorąc punkt $(\sum\limits_{i \in \mathcal{I}_1} \frac{A_i}{p_i} \boldsymbol{V}_i, \frac{a_0}{p_0(1-a_0)} \boldsymbol{X}\boldsymbol{X}')$

dostajemy, że

$$\boldsymbol{\Pi}_o \left[\sum_{i\in\mathcal{I}_1} \frac{A_i}{p_i}\boldsymbol{V}_i + \frac{a_0}{p_0(1-a_0)}\boldsymbol{X}\boldsymbol{X}'\right]\boldsymbol{L} = \frac{a_0}{p_0(1-a_0)}\boldsymbol{\Pi}_o\boldsymbol{X}\boldsymbol{X}'\boldsymbol{K}.$$

Rzeczywiście,

$$\begin{aligned}
&\left[\sum_{i\in\mathcal{I}_1} \frac{A_i}{p_i} Z_i Z_i' + \frac{a_0}{p_0(1-a_0)} Z_0 Z_0'\right] L_0(a_0) \\
&= \left[\sum_{i\in\mathcal{I}_1} A_i(E_0+E_i) + \frac{a_0}{1-a_0}E_0\right] a_0 E_0 K = \left(\sum_{i\in\mathcal{I}_1} A_i + \frac{a_0}{1-a_0}\right) a_0 E_0 K \\
&= \left(1+\frac{a_0}{1-a_0}\right) a_0 E_0 K = \frac{a_0}{1-a_0} E_0 K = \frac{a_0}{p_0(1-a_0)} X X' K,
\end{aligned}$$

$$\begin{aligned}
&\left[\sum_{i\in\mathcal{I}_1} \frac{A_i}{p_i} Z_i Z_i' + \frac{a_0}{p_0(1-a_0)} Z_0 Z_0'\right] L_j(a_j, A_j) - \frac{A_j}{p_j} Z_j Z_j' Q_j \\
&= \left[\sum_{i\in\mathcal{I}_1} A_i(E_0+E_i) + \frac{a_0}{1-a_0}E_0\right][E_j + (1-a_0)A_j E_0] Q_j - A_j(E_0+E_j)Q_j \\
&= \left[A_j E_j + (1-a_0)A_j \sum_{i\in\mathcal{I}_1} A_i E_0 + a_0 A_j E_0\right] Q_j - A_j(E_0+E_j)Q_j \\
&= [A_j E_j + (1-a_0)A_j E_0 + a_0 A_j E_0]\, Q_j - A_j(E_0+E_j)Q_j = \mathbf{0}, \quad \text{dla } \; j \in \mathcal{I}_1
\end{aligned}$$

oraz

$$\begin{aligned}
&\left[\sum_{i\in\mathcal{I}_1} \frac{A_i}{p_i} Z_i Z_i' + \frac{a_0}{p_0(1-a_0)} Z_0 Z_0'\right] L_j(a_j, A_j) \\
&= \left[\sum_{i\in\mathcal{I}_1} A_i(E_0+E_i) + \frac{a_0}{1-a_0}E_0\right] a_j E_j Q_j = \mathbf{0}, \quad \text{w przeciwnym razie.}
\end{aligned}$$

Stąd wynika, że $\boldsymbol{L}'\boldsymbol{Y}$ jest lokalnie optymalny w punkcie W w klasie $\mathcal{L}_0$ oraz, że zbiór $\mathcal{B}((\sum_{i\in\mathcal{I}_1} \frac{A_i}{p_i}\boldsymbol{V}_i, \frac{a_0}{p_0(1-a_0)}\boldsymbol{X}\boldsymbol{X}')|\mathcal{L}_o)$ możemy przedstawić w następujący sposób

$$\mathcal{L}_1 = \boldsymbol{L} + \mathcal{R}(T^t_{\boldsymbol{\Pi}_1}),$$

gdzie

$$\boldsymbol{\Pi}_1 = \begin{bmatrix} I_n - (E_0 + \sum\limits_{i \in \mathcal{I}_1} E_i) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix} \in \mathcal{M}_{(k+1)n \times (k+1)n}.$$

Biorąc punkt $(\sum\limits_{i \in \mathcal{I}_2} \frac{1}{p_i} \boldsymbol{V}_i, \mathbf{0}) \in [\mathcal{T}]$ dostajemy $\boldsymbol{\Pi}_1(\sum\limits_{i \in \mathcal{I}_2} \frac{1}{p_i} \boldsymbol{V}_i)\boldsymbol{L} = \mathbf{0}$. Zatem $\boldsymbol{L}'\boldsymbol{Y}$ jest lokalnie optymalny w tym punkcie w klasie $\mathcal{L}$. Zbiór $\mathcal{B}((\sum\limits_{i \in \mathcal{I}_2} \frac{1}{p_i} \boldsymbol{V}_i, \mathbf{0})|\mathcal{L}_1) \neq \{\boldsymbol{L}\}$ może być wyrażony w postaci

$$\mathcal{L}_2 = \boldsymbol{L} + \mathcal{R}(T^t_{\boldsymbol{\Pi}_2}),$$

gdzie

$$\boldsymbol{\Pi}_2 = \begin{bmatrix} I_n - (E_0 + \sum\limits_{i \in \mathcal{I}} E_i) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix} \in \mathcal{M}_{(k+1)n \times (k+1)n}.$$

Dalej możemy sprawdzić, że $\boldsymbol{\Pi}_2(\sum\limits_{i \notin \mathcal{I}} \frac{a_i}{p_i(1-a_i)} \boldsymbol{V}_i + \boldsymbol{V}_{k+1})\boldsymbol{L} = \mathbf{0}$ jak również, że $\boldsymbol{L}$ jest jedynym rozwiązaniem tego równania. To z kolei implikuje, że $\boldsymbol{L}'\boldsymbol{Y}$ jest dopuszczalny w klasie $\mathcal{L}_o$, co kończy dowód. $\square$

W modelu (2.6), w odróżnieniu od modelu (2.5), wystarczy wykonać co najwyżej trzy kroki, ponieważ liczba kroków determinowana jest nie tylko liczbą tych spośród współczynników a_i, które są równe 1, ale przede wszystkim postacią współczynników A_i.

3.3. Zrównoważony model losowy dwukierunkowej klasyfikacji krzyżowej z interakcją

W tym podrozdziale opiszemy klasę estymatorów dopuszczalnych dla $\boldsymbol{K'X}\beta$ w klasie $\mathcal{L}_o$ na zbiorze parametrów $\mathcal{T}$ w modelu (2.4) odpowiadającym modelowi (2.7). Estymatory te będą wyrażone jako kombinacje liniowe następujących macierzy

$$\begin{aligned} E_0 &= \frac{1}{p_0} Z_0 Z_0', \\ E_1 &= \frac{1}{p_1} Z_1 Z_1' - E_0, \\ E_2 &= \frac{1}{p_2} Z_2 Z_2' - E_0, \\ E_3 &= \frac{1}{p_3} Z_3 Z_3' - (E_0 + E_1 + E_2), \end{aligned}$$

gdzie $Z_0 = \mathbf{1}_n, p_0 = n, p_1 = n_2 n_3, p_2 = n_1 n_3$, natomiast $p_3 = n_3$. Oczywiście E_0, E_1, E_2 oraz E_3 są macierzami idempotentnymi i ortogonalnymi, a macierze $Z_i Z_i'$ są ich liniowymi kombinacjami dla $i = 0, 1, 2, 3$. Aby skrócić zapis niektórych wzorów przyjmijmy jeszcze, że $Z_4 = I_n, E_4 = I_n - (E_0 + E_1 + E_2 + E_3)$ oraz $p_4 = 1$.

Lemat 3.3. *Estymator $\boldsymbol{L'Y}$ jest jednoznacznie lokalnie optymalny w punkcie $(\sum_{i=1}^{4} s_i \boldsymbol{V}_i, s_0 \boldsymbol{XX'})$ z $[\mathcal{T}]$ w klasie $\mathcal{L}_o$ w modelu* (2.4) *odpowiadającym modelowi* (2.7) *wtedy i tylko wtedy, gdy $s_0 \geq 0, s_1 \geq 0, s_2 \geq 0, s_3 \geq 0, s_4 > 0$ oraz*

$$\boldsymbol{L} = \begin{bmatrix} L_0 & L_1 & L_2 & L_3 \\ \mathbf{0} & -Q_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & -Q_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & -Q_3 \end{bmatrix}, \tag{3.7}$$

gdzie

$$
\begin{aligned}
L_0 &= a_0 E_0 K, \\
L_1 &= a_1 \left[E_1 + (1-a_0) \frac{\frac{1}{1-a_1}}{\frac{1}{1-a_1} + \frac{1}{1-a_2} - 1} E_0 \right] Q_1, \\
L_2 &= a_2 \left[E_2 + (1-a_0) \frac{\frac{1}{1-a_2}}{\frac{1}{1-a_1} + \frac{1}{1-a_2} - 1} E_0 \right] Q_2, \\
L_3 &= a_3 \left[E_3 + (1-a_2)E_2 + (1-a_1)E_1 + (1-a_0) \frac{1}{\frac{1}{1-a_1} + \frac{1}{1-a_2} - 1} E_0 \right] Q_3
\end{aligned}
$$

oraz

$$a_0 = \frac{s_0 p_0}{s_0 p_0 + s_1 p_1 + s_3 p_3 + s_4}, \tag{3.8}$$

$$a_1 = \frac{s_1 p_1}{s_1 p_1 + s_3 p_3 + s_4}, \tag{3.9}$$

$$a_2 = \frac{s_2 p_2}{s_2 p_2 + s_3 p_3 + s_4}, \tag{3.10}$$

$$a_3 = \frac{s_3 p_3}{s_3 p_3 + s_4}. \tag{3.11}$$

Dowód. Niech $\left(\sum_{i=1}^{4} s_i \boldsymbol{V}_i, s_0 \boldsymbol{X}\boldsymbol{X}' \right)$ będzie punktem z $[\mathcal{T}]$ takim, że $s_0 \geq 0, s_1 \geq 0, s_2 \geq 0, s_3 \geq 0$ oraz $s_4 \geq 0$. Estymator $\boldsymbol{L}'\boldsymbol{Y}$, gdzie macierz $\boldsymbol{L}$ dana jest wzorem (3.7), jest jednoznacznie lokalnie optymalny wtedy i tylko wtedy, gdy $s_4 > 0$ oraz

$$
\begin{aligned}
L_0 &= s_0 M^{-1} X X' K, \\
L_i &= s_i M^{-1} Z_i Z_i' Q_i \quad \text{dla} \quad i = 1, 2, 3,
\end{aligned}
$$

gdzie

$$
\begin{aligned}
M = \sum_{i=0}^{4} s_i Z_i Z_i' &= \left(\sum_{i=0}^{4} s_i p_i \right) E_0 + (s_1 p_1 + s_3 p_3 + s_4) E_1 \\
&\quad + (s_2 p_2 + s_3 p_3 + s_4) E_2 + (s_3 p_3 + s_4 p_4) E_3 + s_4 p_4 E_4.
\end{aligned}
$$

Zatem

$$
\begin{aligned}
L_0 &= \frac{s_0 p_0}{\sum\limits_{j=0}^{4} s_j p_j} E_0 K, \\
L_1 &= \frac{s_1 p_1}{s_1 p_1 + s_3 p_3 + s_4} \left(E_1 + \frac{s_1 p_1 + s_3 p_3 + s_4}{\sum\limits_{j=0}^{4} s_j p_j} E_0 \right) Q_1, \\
L_2 &= \frac{s_2 p_2}{s_2 p_2 + s_3 p_3 + s_4} \left(E_2 + \frac{s_2 p_2 + s_3 p_3 + s_4}{\sum\limits_{j=0}^{4} s_j p_j} E_0 \right) Q_2,
\end{aligned}
$$

natomiast

$$
L_3 = \frac{s_3 p_3}{s_3 p_3 + s_4} \left(E_3 + \frac{s_3 p_3 + s_4}{s_2 p_2 + s_3 p_3 + s_4} E_2 + \frac{s_3 p_3 + s_4}{s_1 p_1 + s_3 p_3 + s_4} E_1 + \frac{s_3 p_3 + s_4}{\sum\limits_{j=0}^{4} s_j p_j} E_0 \right) Q_3.
$$

Następnie, w przypadku macierzy L_1, L_2 oraz L_3 stosujemy dla współczynników stojących przy macierzy E_0 przekształcenia, które zostały omówione w dowodzie lematu 3.4. □

Twierdzenie 3.3. *Estymator* $\boldsymbol{L}'\boldsymbol{Y}$ *jest dopuszczalny dla* $\boldsymbol{K}'\boldsymbol{X}\beta$ *w klasie* $\mathcal{L}_o$ *na* $\mathcal{T}$ *w modelu* (2.4) *odpowiadającym modelowi* (2.7) *wtedy i tylko wtedy, gdy*

$$
\boldsymbol{L} \in \left\{ \begin{bmatrix} L_0(a_0) & L_1(a_1, A_1) & L_2(a_2, A_2) & L_3(a_3, A_1, A_2) \\ \mathbf{0} & -Q_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & -Q_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & -Q_3 \end{bmatrix} : a_i \in [0,1], i = 0,1,2,3 \right\}, \tag{3.12}
$$

gdzie

$$
\begin{aligned}
L_0(a_0) &= a_0E_0K,\\
L_1(a_1, A_1) &= a_1[E_1 + (1-a_0)A_1E_0]Q_1,\\
L_2(a_2, A_2) &= a_2[E_2 + (1-a_0)A_2E_0]Q_2,
\end{aligned}
$$

$$L_3(a_3, A_1, A_2) = a_3[E_3+(1-a_2)E_2+(1-a_1)E_1+(1-a_0)(A_1+A_2-1)E_0]Q_3$$

oraz

$$A_1 = \frac{\frac{1}{1-a_1}}{\frac{1}{1-a_1} + \frac{1}{1-a_2} - 1}, \tag{3.13}$$

$$A_2 = \frac{\frac{1}{1-a_2}}{\frac{1}{1-a_1} + \frac{1}{1-a_2} - 1} \tag{3.14}$$

dla $a_1 \in [0,1)$ *i* $a_2 \in [0,1)$,

w przeciwnym razie

A_1 *i* A_2 *są nieujemnymi liczbami takimi, że* $A_1 + A_2 = 1$, *przy czym* $A_1 = 0$, *gdy* $a_1 \in [0,1)$ *oraz* $A_2 = 0$, *gdy* $a_2 \in [0,1)$.

Dowód. Warunek konieczny. Zauważmy, że współczynniki a_0, a_1, a_2 oraz a_3 określone odpowiednio wzorami (3.8), (3.9), (3.10) oraz (3.11) przyjmują każdą wartość z przedziału $[0,1)$ przy odpowiednio dobranych $s_0 \geq 0, s_1 \geq 0, s_2 \geq 0, s_3 \geq 0$ oraz $s_4 > 0$. Dalej, możemy sprawdzić, że dla A_1 i A_2 danych formułami (3.13) i (3.14) mamy, że $1 < A_1 + A_2 \leq 2$. Ponieważ $\frac{1}{1-a_1} + \frac{1}{1-a_2} \to +\infty$, gdy $a_1 \to 1$ lub $a_2 \to 1$, więc $A_1 + A_2 \to 1$ oraz $A_i \to 0$ dla tego wskaźnika i, dla którego $a_i \in [0,1)$. Stąd wynika, że zbiór (3.12) jest domknięciem zbioru

$$\{\boldsymbol{L}_* : \boldsymbol{L}_*'\boldsymbol{Y} \text{ jest jednoznacznie lokalnie optymalnym w punkcie z } [\mathcal{T}] \text{ w klasie } \mathcal{L}_o\}.$$

Zatem na mocy twierdzenia 1.3 pierwsza część dowodu jest zakończona.

Warunek dostateczny. Podobnie, jak w przypadku twierdzeń 3.1 oraz 3.2, stosując metodę LaMotte'a wskażemy punkty należące do $[\mathcal{T}]$, w których estymatory $\boldsymbol{L}'\boldsymbol{Y}$ z klasy (3.12) są lokalnie optymalne w odpowiednich rozmaitościach liniowych $\mathcal{L} \subseteq \mathcal{L}_o$. Tak, jak w modelach (2.5) oraz (2.6) liczba kroków będzie zależała od tego, które i ile spośród współczynników a_i są równe 1. Ponieważ w modelu (2.7) występują cztery takie współczynniki, więc mamy do rozpatrzenia 16 różnych przypadków. Opiszemy 6 z nich, gdyż dla pozostałych dowód będzie analogiczny do dowodu twierdzenia 3.1 bądź 3.2. Ograniczymy się przy tym do podania kolejnych punktów i rozmaitości linowych, ponieważ sposób postępowania został już wielokrotnie wcześniej opisany.

(1) Jeżeli $a_i \in [0,1)$ dla każdego $i = 0,1,2,3$, to wówczas możemy sprawdzić, że $\boldsymbol{L'Y}$ jest jednoznacznie lokalnie optymalny w punkcie $W = (W_1, W_2) \in [\mathcal{T}]$, gdzie

$$\begin{aligned} W_1 &= \frac{a_1}{p_1(1-a_1)(1-a_3)}\boldsymbol{V}_1 + \frac{a_2}{p_2(1-a_2)(1-a_3)}\boldsymbol{V}_2 + \frac{a_3}{p_3(1-a_3)}\boldsymbol{V}_3 + \boldsymbol{V}_4, \\ W_2 &= \frac{a_0}{p_0(1-a_0)(1-a_3)}\left[\frac{a_1}{(1-a_1)} + \frac{a_2}{(1-a_2)} + 1\right]\boldsymbol{XX'}, \end{aligned}$$

w klasie $\mathcal{L}_o$ i oczywiście dopuszczalny dla $\boldsymbol{K'X}\beta$ w tej klasie.

(2) Jeżeli $a_0 = 1, a_1 \in [0,1), a_2 \in [0,1), a_3 = 1$, to

(k0)

$$\begin{aligned} W^{(0)} &= (\mathbf{0}, \boldsymbol{XX'}), \\ \mathcal{L}_1 &= \mathcal{B}(W^{(0)}|\mathcal{L}_o) = \boldsymbol{L} + \mathcal{R}(T^t_{\mathbf{\Pi}_1}), \end{aligned}$$

gdzie

$$\mathbf{\Pi}_1 = \begin{bmatrix} \Pi_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix} \in \mathcal{M}_{(k+1)n \times (k+1)n}, \quad \Pi_1 = I_n - E_0,$$

(k1)

$$\begin{aligned} W^{(1)} &= \left(\frac{a_1}{p_1(1-a_1)} \boldsymbol{V}_1 + \frac{a_2}{p_2(1-a_2)} \boldsymbol{V}_2 + \frac{1}{p_3} \boldsymbol{V}_3, \boldsymbol{0} \right), \\ \mathcal{L}_2 &= \mathcal{B}(W^{(1)}|\mathcal{L}_1) = \boldsymbol{L} + \mathcal{R}(T^t_{\boldsymbol{\Pi}_2}), \\ \Pi_2 &= E_4, \end{aligned}$$

(k2)

$$\begin{aligned} W^{(2)} &= (\boldsymbol{V}_4, \boldsymbol{0}), \\ \mathcal{L}_3 &= \mathcal{B}(W^{(2)}|\mathcal{L}_2) = \{\boldsymbol{L}\}. \end{aligned}$$

(3) Jeżeli $a_0 = 1, a_1 \in [0,1), a_2 = 1, a_3 = 1$, to

(k0)

$$\begin{aligned} W^{(0)} &= (\boldsymbol{0}, \boldsymbol{X}\boldsymbol{X}'), \\ \mathcal{L}_1 &= \mathcal{B}(W^{(0)}|\mathcal{L}_o) = \boldsymbol{L} + \mathcal{R}(T^t_{\boldsymbol{\Pi}_1}), \\ \Pi_1 &= I_n - E_0, \end{aligned}$$

(k1)

$$\begin{aligned} W^{(1)} &= \left(\frac{1}{p_2} \boldsymbol{V}_2, \boldsymbol{0} \right), \\ \mathcal{L}_2 &= \mathcal{B}(W^{(1)}|\mathcal{L}_1) = \boldsymbol{L} + \mathcal{R}(T^t_{\boldsymbol{\Pi}_2}), \\ \Pi_2 &= E_1 + E_3 + E_4, \end{aligned}$$

(k2)

$$\begin{aligned} W^{(2)} &= \left(\frac{a_1}{p_1(1-a_1)} \boldsymbol{V}_1 + \frac{1}{p_3} \boldsymbol{V}_3, \boldsymbol{0} \right), \\ \mathcal{L}_3 &= \mathcal{B}(W^{(2)}|\mathcal{L}_2) = \boldsymbol{L} + \mathcal{R}(T^t_{\boldsymbol{\Pi}_3}), \\ \Pi_3 &= E_4, \end{aligned}$$

(k3)

$$\begin{aligned} W^{(3)} &= (\boldsymbol{V}_4, \mathbf{0}), \\ \mathcal{L}_4 &= \mathcal{B}(W^{(3)}|\mathcal{L}_3) = \{\boldsymbol{L}\}. \end{aligned}$$

(4) Jeżeli $a_0 \in [0,1), a_1 \in [0,1), a_2 \in [0,1), a_3 = 1$, to

(k0)

$$W^{(0)} = \left(W_1^{(0)}, W_2^{(0)}\right),$$

gdzie

$$\begin{aligned} W_1^{(0)} &= \frac{a_1}{p_1(1-a_1)}\boldsymbol{V}_1 + \frac{a_2}{p_2(1-a_2)}\boldsymbol{V}_2 + \frac{1}{p_3}\boldsymbol{V}_3, \\ W_2^{(0)} &= \frac{a_0}{p_0(1-a_0)}(\frac{a_1}{(1-a_1)} + \frac{a_2}{(1-a_2)} + 1)\boldsymbol{X}\boldsymbol{X}', \\ \mathcal{L}_1 &= \mathcal{B}(W^{(0)}|\mathcal{L}_o) = \boldsymbol{L} + \mathcal{R}(T^t_{\boldsymbol{\Pi}_1}), \\ \Pi_1 &= E_4, \end{aligned}$$

(k1)

$$\begin{aligned} W^{(1)} &= (\boldsymbol{V}_4, \mathbf{0}), \\ \mathcal{L}_2 &= \mathcal{B}(W^{(1)}|\mathcal{L}_1) = \{\boldsymbol{L}\}. \end{aligned}$$

(5) Jeżeli $a_0 \in [0,1), a_1 \in [0,1), a_2 = 1, a_3 = 1$, to

(k0)

$$\begin{aligned} W^{(0)} &= \left(\frac{1}{p_2}\boldsymbol{V}_2 +, \frac{a_0}{p_0(1-a_0)}\boldsymbol{X}\boldsymbol{X}'\right), \\ \mathcal{L}_1 &= \mathcal{B}(W^{(0)}|\mathcal{L}_o) = \boldsymbol{L} + \mathcal{R}(T^t_{\boldsymbol{\Pi}_1}), \\ \Pi_1 &= E_1 + E_3 + E_4, \end{aligned}$$

(k1)

$$
\begin{aligned}
W^{(1)} &= \left(\frac{a_1}{p_1(1-a_1)}\boldsymbol{V}_1 + \frac{1}{p_3}\boldsymbol{V}_3, \mathbf{0}\right),\\
\mathcal{L}_2 &= \mathcal{B}(W^{(1)}|\mathcal{L}_1) = \boldsymbol{L} + \mathcal{R}(T^t_{\boldsymbol{\Pi}_2}),\\
\Pi_2 &= E_4,
\end{aligned}
$$

(k2)

$$
\begin{aligned}
W^{(2)} &= (\boldsymbol{V}_4, \mathbf{0}),\\
\mathcal{L}_3 &= \mathcal{B}(W^{(2)}|\mathcal{L}_2) = \{\boldsymbol{L}\}.
\end{aligned}
$$

(6) Jeżeli $a_0 \in [0,1), a_1 = 1, a_2 \in [0,1), a_3 = 1$, to

(k0)

$$
\begin{aligned}
W^{(0)} &= \left(\frac{1}{p_1}\boldsymbol{V}_1 +, \frac{a_0}{p_0(1-a_0)}\boldsymbol{X}\boldsymbol{X}'\right),\\
\mathcal{L}_1 &= \mathcal{B}(W^{(0)}|\mathcal{L}_o) = \boldsymbol{L} + \mathcal{R}(T^t_{\boldsymbol{\Pi}_1}),\\
\Pi_1 &= E_2 + E_3 + E_4,
\end{aligned}
$$

(k1)

$$
\begin{aligned}
W^{(1)} &= \left(\frac{a_2}{p_2(1-a_2)}\boldsymbol{V}_2 + \frac{1}{p_3}\boldsymbol{V}_3, \mathbf{0}\right),\\
\mathcal{L}_2 &= \mathcal{B}(W^{(1)}|\mathcal{L}_1) = \boldsymbol{L} + \mathcal{R}(T^t_{\boldsymbol{\Pi}_2}),\\
\Pi_2 &= E_4,
\end{aligned}
$$

(k2)

$$
\begin{aligned}
W^{(2)} &= (\boldsymbol{V}_4, \mathbf{0}),\\
\mathcal{L}_3 &= \mathcal{B}(W^{(2)}|\mathcal{L}_2) = \{\boldsymbol{L}\}.
\end{aligned}
$$

(7) Jeżeli $a_0 = 1, a_1 = 1, a_2 \in [0,1], a_3 \in [0,1]$, to postępujemy tak, jak w dowodzie twierdzenia 3.1.

(8) W pozostałych przypadkach postępujemy podobnie, jak w dowodzie twierdzenia 3.2. W kolejnych krokach wskazujemy punkty, dla których współczynniki stojące przy macierzach $\boldsymbol{V}_1, \ldots, \boldsymbol{V}_4$ oraz $\boldsymbol{X}\boldsymbol{X}'$ są odpowiednimi modyfikacjami tychże współczynników z dowodu twierdzenia 3.2, określających punkt, w którym $\boldsymbol{L}'\boldsymbol{Y}$ jest jednoznacznie lokalnie optymalny w klasie $\mathcal{L}_o$.

$\square$

4. Jawna charakterystyka estymatorów dopuszczalnych w pewnym modelu liniowym z dwoma komponentami

Nasze rozważania dotyczące dopuszczalności zostaną teraz ograniczone do pewnego modelu z dwoma komponentami. Model ten jest szczególnym przypadkiem modelu (2.1), dla którego $k = 1$ oraz $\beta \in \mathbb{R}$, a macierz $Z_1 Z_1'$ jest macierzą proporcjonalną do macierzy idempotentnej, tzn. że $(Z_1 Z_1')^2 = p_1 Z_1 Z_1'$, gdzie p_1 jest dodatnią liczbą rzeczywistą. Będziemy go oznaczać następująco

$$Y \sim (X\beta,\ \sigma_1^2 Z_1 Z_1' + \sigma_2^2 I_n). \tag{4.1}$$

Przykładami powyższego modelu są modele (2.8) oraz (2.9). Podamy warunki konieczne i dostateczne na to, aby estymator liniowy był dopuszczalny dla wektora $[(K'X\beta)', (Q_1' Z_1 u_1)']'$ w modelu (4.1). Na mocy lematu 2.1 wyrazimy je w terminach macierzy określających model (2.4), który odpowiada rozważanemu modelowi z dwoma komponentami. Rozpatrzymy dwa przypadki, gdy $X \in \mathcal{R}(Z_1)$ oraz $X \notin \mathcal{R}(Z_1)$. Dla wygody wyraźmy wektor X w postaci $X_1 + X_2$, gdzie $X_1 \in \mathcal{R}(Z_1)$ natomiast $X_2 \in \mathcal{N}(Z_1')$.

4.1. Przypadek, gdy $X \in \mathcal{R}(Z_1)$

W tym podrozdziale zajmiemy się dopuszczalnością liniowych estymatorów $\boldsymbol{L'Y}$ dla $\boldsymbol{K'X}\beta$ w klasie $\mathcal{L}_o$ w modelu (2.4), który odpowiada modelowi (4.1) w przypadku, gdy $X = X_1$. Ponieważ zakładamy także, że $(Z_1Z_1')^2 = p_1Z_1Z_1'$, więc możemy przedstawić macierze XX', Z_1Z_1' oraz I_n jako kombinacje liniowe następujących ortogonalnych i idempotentnych macierzy

$$\begin{aligned} E_0 &= \frac{1}{X'X}XX', \\ E_1 &= \frac{1}{p_1}Z_1Z_1' - E_0 \end{aligned}$$

oraz

$$E_2 = I_n - \frac{1}{p_1}Z_1Z_1'.$$

Tak, jak poprzednio, niech $p_0 = X'X, Z_0 = X, p_2 = 1$ oraz $Z_2 = I_n$. Zaczniemy od sformułowania lematu, który podaje postać estymatorów jednoznacznie lokalnie optymalnych, a następnie, powołując się na twierdzenie 1.3, wyrazimy w jawnej postaci klasę estymatorów dopuszczalnych dla wspomnianej już funkcji parametru β.

Lemat 4.1. *Niech* $X = X_1 \neq \mathbf{0}$. *Estymator* $\boldsymbol{L'Y}$ *jest jednoznacznie lokalnie optymalny w punkcie* $(s_1\boldsymbol{V}_1 + s_2\boldsymbol{V}_2, s_0\boldsymbol{XX'})$ *należącym do* $[\mathcal{T}]$ *w klasie* $\mathcal{L}_o$ *w modelu* (2.4) *odpowiadającym modelowi* (4.1) *wtedy i tylko wtedy, gdy* $s_0 \geq 0, s_1 \geq 0,\ s_2 > 0$ *oraz* $\boldsymbol{L} = \begin{bmatrix} L_0 & L_1 \\ \mathbf{0} & -Q_1 \end{bmatrix}$, *gdzie*

$$\begin{aligned} L_0 &= a_0E_0K, \\ L_1 &= a_1(E_1 - (1-a_0)E_0)Q_1 \end{aligned}$$

oraz

$$\begin{aligned} a_0 &= \frac{s_0p_0}{s_0p_0 + s_1p_1 + s_2p_2}, \\ a_1 &= \frac{s_1p_1}{s_1p_1 + s_2p_2}. \end{aligned} \tag{4.2}$$

Dowód. Dowód jest anologiczny do dowodu lematu 3.1 bądź 3.4 dla $k = 1$. $\square$

Twierdzenie 4.1. *Niech $X = X_1 \neq \mathbf{0}$. Estymator $\boldsymbol{L'Y}$ jest dopuszczalny dla $\boldsymbol{K'X}\beta$ w klasie $\mathcal{L}_o$ na $\mathcal{T}$ w modelu* (2.4) *odpowiadającym modelowi* (4.1) *wtedy i tylko wtedy, gdy*

$$\boldsymbol{L} \in \left\{ \begin{bmatrix} L_0(a_0) & L_1(a_0, a_1) \\ \mathbf{0} & -Q_1 \end{bmatrix} : a_0 \in [0,1], a_1 \in [0,1] \right\},$$

gdzie

$$\begin{aligned} L_0(a_0) &= a_0 E_0 K, \\ L_1(a_0, a_1) &= a_1(E_1 - (1 - a_0)E_0)Q_1. \end{aligned}$$

Dowód. Dowód jest anologiczny do dowodu twierdzenia 3.1 bądź 3.2 dla $k = 1$. $\square$

Podstawiając za E_0 i E_1 odpowiednio macierze $\frac{1}{X'X}XX'$ oraz $\frac{1}{p_1}Z_1Z_1' - E_0$, uzyskujemy następujący wniosek.

Wniosek 4.1. *Niech $X = X_1 \neq \mathbf{0}$. Estymator $\boldsymbol{L'Y}$ jest dopuszczalny dla $\boldsymbol{K'X}\beta$ w klasie $\mathcal{L}_o$ na $\mathcal{T}$ w modelu* (2.4) *odpowiadającym modelowi* (4.1) *wtedy i tylko wtedy, gdy*

$$\boldsymbol{L} \in \left\{ \begin{bmatrix} L_0(a_0) & L_1(a_0, a_1) \\ \mathbf{0} & -Q_1 \end{bmatrix} : a_0 \in [0,1], a_1 \in [0,1] \right\},$$

gdzie

$$L_0(a_0) = \frac{a_0}{X'X} XX'K, \tag{4.3}$$

$$L_1(a_0, a_1) = a_1 \left(\frac{1}{p_1} Z_1 Z_1' - \frac{a_0}{X'X} XX' \right) Q_1. \tag{4.4}$$

Dla uzupełnienia naszych rozważań rozpatrzymy także przypadek, gdy $X = \mathbf{0}$.

Twierdzenie 4.2. *Niech* $X = \mathbf{0}$. *Estymator* $\boldsymbol{L}'\boldsymbol{Y}$ *jest dopuszczalny dla* $\mathbf{0}$ *w klasie* $\mathcal{L}_o$ *na* $\mathcal{T}$ *w odpowiednim modelu* (2.4) *wtedy i tylko wtedy, gdy*

$$\boldsymbol{L} \in \left\{ \begin{bmatrix} \mathbf{0} & L_1(a_1) \\ \mathbf{0} & -Q_1 \end{bmatrix} : a_1 \in [0,1] \right\},$$

gdzie

$$L_1(a_1) = \frac{a_1}{p_1} Z_1 Z_1' Q_1.$$

4.2. Przypadek, gdy $X \notin \mathcal{R}(Z_1)$

Jeżeli pominiemy założenie, że $X \in \mathcal{R}(Z_1)$, to uzyskamy wówczas inne postaci estymatorów jednoznacznie lokalnie optymalnych, a tym samym inne postaci estymatorów dopuszczalnych dla $\boldsymbol{K'X}\beta$ w klasie $\mathcal{L}_o$ na $\mathcal{T}$ w modelu (2.4) odpowiadającym modelowi (4.1). Bowiem, gdy $X_2 \neq \mathbf{0}$, to nie możemy macierzy XX', Z_1Z_1' oraz I_n przedstawić w postaci pewnych kombinacji liniowych macierzy ortogonalnych i idempotentnych.

Lemat 4.2. *Niech* $X_2 \neq \mathbf{0}$. *Estymator* $\boldsymbol{L}'\boldsymbol{Y}$ *jest jednoznacznie lokalnie optymalny w punkcie* $(s_1\boldsymbol{V}_1 + s_2\boldsymbol{V}_2, s_0\boldsymbol{XX}')$ *należącym do* $[\mathcal{T}]$ *w klasie* $\mathcal{L}_o$ *w modelu* (2.4) *odpowiadającym modelowi* (4.1) *wtedy i tylko wtedy, gdy* $s_0 \geq 0$,$s_1 \geq 0$, $s_2 > 0$ *oraz* $\boldsymbol{L} = \begin{bmatrix} L_0 & L_1 \\ \mathbf{0} & -Q_1 \end{bmatrix}$, *gdzie*

$$\begin{aligned} L_0 &= \frac{a_0}{X'X}\left[X_1X_1' + X_1X_2' + \frac{1}{1-a_1}(X_2X_1' + X_2X_2')\right]K, \\ L_1 &= a_1\left[\frac{1}{p_1}Z_1Z_1' - \frac{a_0}{X'X}X_1X_1' - \frac{a_0}{(1-a_1)X'X}X_2X_1'\right]Q_1 \end{aligned}$$

oraz

$$a_0 = \frac{s_0(1-a_1)X'X}{s_2 + s_0[(1-a_1)X'X + a_1X_2'X_2]}, \tag{4.5}$$

natomiast a_1 *wyraża się wzorem* (4.2).

Dowód. Tak jak w dowodzie lematów 3.1, 3.4 oraz 3.3 mamy, że estymator $\boldsymbol{L}'\boldsymbol{Y} = \begin{bmatrix} L_0 & L_1 \\ \mathbf{0} & -Q_1 \end{bmatrix}' \boldsymbol{Y}$ jest jednoznacznie lokalnie optymalny w punkcie $(s_1\boldsymbol{V}_1 + s_2\boldsymbol{V}_2, s_0\boldsymbol{X}\boldsymbol{X}')$ ze zbioru $[\mathcal{T}]$ w klasie $\mathcal{L}_o$ wtedy i tylko wtedy, gdy $s_2 > 0$ oraz

$$\begin{aligned} L_0 &= s_0 M^{-1} X X' K, \\ L_1 &= s_1 M^{-1} Z_1 Z_1' Q_1, \end{aligned}$$

gdzie $M = s_0 X X' + s_1 Z_1 Z_1' + s_2 I_n$.

Wykorzystując równości

$$\begin{aligned} Z_1 Z_1' X X' &= p(X_1 X_1' + X_1 X_2'), \\ X X' X_1 X_1' &= X_1' X_1 (X_1 X_1' + X_2 X_1') \end{aligned}$$

oraz

$$X X' X_2 X_2' = X_2' X_2 (X_1 X_2' + X_2 X_2'),$$

można sprawdzić, że

$$M^{-1} = -c\left(X X' - \frac{s_1 p_1}{s_1 p_1 + s_2} X_1 X_1' + \frac{s_1 p_1}{s_2} X_2 X_2'\right) + \frac{-s_1}{s_2(s_1 p_1 + s_2)} Z_1 Z_1' + \frac{1}{s_2} I_n,$$

gdzie

$$c = \frac{s_0}{s_2(s_1 p_1 + s_2) + s_0(s_2 X' X + s_1 p_1 X_2' X_2)}.$$

Mało skomplikowane przekształcenia kończą dowód. □

Twierdzenie 4.3. *Niech* $X_2 \neq \mathbf{0}$. *Estymator* $\boldsymbol{L}'\boldsymbol{Y}$ *jest dopuszczalny dla* $\boldsymbol{K}'\boldsymbol{X}\beta$ *w klasie* $\mathcal{L}_o$ *na* $\mathcal{T}$ *w modelu* (2.4) *odpowiadającym modelowi* (4.1) *wtedy i tylko wtedy, gdy*

$$\boldsymbol{L} \in \left\{ \begin{bmatrix} L_0(a_0) & L_1(a_0, a_1) \\ \mathbf{0} & -Q_1 \end{bmatrix} : a_0 \in \left[0, \frac{(1-a_1)X'X}{(1-a_1)X'X + a_1 X_2' X_2}\right],\ a_1 \in [0,1] \right\}, \tag{4.6}$$

gdzie

$$\begin{aligned} L_0(a_0,a_1) &= \frac{a_0}{X'X}(X_1X_1'+X_1X_2')K + A(X_2X_1'+X_2X_2')K, \\ L_1(a_0,a_1) &= a_1\left(\frac{1}{p_1}Z_1Z_1' - \frac{a_0}{X'X}X_1X_1'\right)Q_1 - a_1AX_2X_1'Q_1 \end{aligned}$$

oraz $A = \frac{a_0}{(1-a_1)X'X}$ *dla* $a_1 \in [0,1)$, *w przeciwnym razie* $A = 0$.

Dowód. Aby udowodnić warunek konieczny, zauważmy, że dla każdych ustalonych $s_1 \geq 0, s_2 > 0$ współczynnik a_0 dany wyrażeniem (4.5) przebiega przedział $\left[0, \frac{(1-a_1)X'X}{(1-a_1)X'X+a_1X_2'X_2}\right)$, gdy $s_0 \in [0,+\infty)$. Co więcej, dla $a_1 \to 1$ mamy $\frac{a_0}{1-a_1} \to 0$. Podobnie jak poprzednio, korzystając z lematu 4.2, otrzymujemy, że zbiór określony wzorem (4.6) jest domknięciem zbioru

$$\{\boldsymbol{L}_* : \boldsymbol{L}_*'\boldsymbol{Y} \text{ jest jednoznacznie lokalnie optymalnym w punkcie z } [\mathcal{T}] \text{ w klasie } \mathcal{L}_o\},$$

co na mocy twierdzenia 1.3 kończy dowód warunku koniecznego.

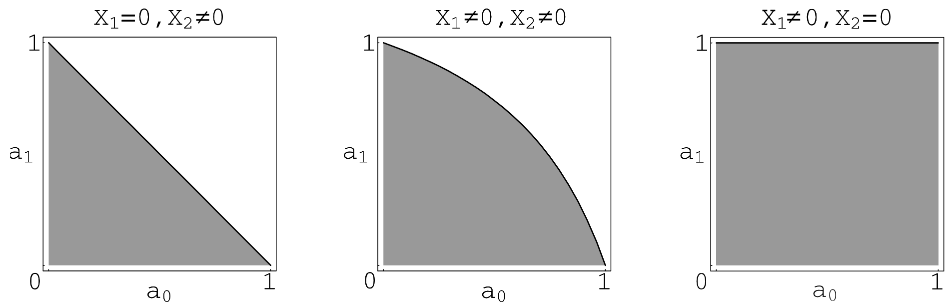
Załóżmy teraz, że $\boldsymbol{L}'\boldsymbol{Y}$ jest estymatorem, dla którego macierz $\boldsymbol{L}$ należy do zbioru (4.6). Jeżeli $a_0 \in \left[0, \frac{(1-a_1)X'X}{(1-a_1)X'X+a_1X_2'X_2}\right)$ oraz $a_1 \in [0,1)$, to wówczas estymator $\boldsymbol{L}'\boldsymbol{Y}$ jest jednoznacznie lokalnie optymalny w punkcie

$$\left(\frac{a_1}{p_1(1-a_1)}\boldsymbol{V}_1 + \boldsymbol{V}_2, \frac{a_0}{(1-a_0)(1-a_1)X'X - a_0a_1X_2'X_2}\boldsymbol{X}\boldsymbol{X}'\right)$$

należącym do $[\mathcal{T}]$ w klasie $\mathcal{L}_o$. Dla $a_0 = \frac{(1-a_1)X'X}{(1-a_1)X'X+a_1X_2'X_2}$ oraz $a_1 \in [0,1)$ dowód przebiega tak samo, jak dowód twierdzenia 3.2, kiedy $a_0 = 1$ i $a_1 \in [0,1)$. Podobnie, w przypadku gdy $a_0 = 0$ oraz $a_1 = 1$, który odpowiada sytuacji kiedy $a_0 = 1$ i $a_1 = 1$ w dowodzie twierdzenia 3.2 z tą tylko różnicą, że $\boldsymbol{\Pi}_2 = \begin{bmatrix} \Pi_2 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}$, natomiast Π_2 jest macierzą rzutu ortogonalnego na przestrzeń $\mathcal{N}(XX'+Z_1Z_1')$. □

Zauważmy, że kładąc $X_2 = \boldsymbol{0}$, uzyskujemy charakterystykę estymatorów dopuszczalnych $\boldsymbol{L}'\boldsymbol{Y}$ podaną we wniosku 4.1.

Wyniki twierdzeń 4.1 oraz 4.3 możemy przedstawić graficznie w sposób następujący.



Pierwszy z rysunków przedstawia obszar tych współczynników a_0 i a_1, dla których liniowe estymatory są dopuszczalne dla $\boldsymbol{K'X}\beta$ w klasie $\mathcal{L}_o$ na $\mathcal{T}$ w modelu (2.4) odpowiadającym modelowi (4.1) w przypadku, gdy $X_1 = \mathbf{0}$ oraz $X_2 \neq \mathbf{0}$. W sytuacji, gdy $X_1 \neq \mathbf{0}$ oraz $X_2 \neq \mathbf{0}$ (rysunek 2), obszar ten jest od góry ograniczony przez fragment hiperboli. Przy ustalonym $X_1 \neq \mathbf{0}$ i $X_2 \to \mathbf{0}$ odpowiednie obszary wypełniają cały kwadrat, dając (po domknięciu) obszar tych a_0 i a_1, dla których rozważane estymatory są dopuszczalne w odpo- wiednim modelu (2.4), dla którego $X \in \mathcal{R}(Z_1)$.

4.3. Zastosowania

Przedstawimy teraz charakterystykę liniowych estymatorów dopuszczalnych dla funkcji $\beta + u_i$ w modelach (2.8) oraz (2.9), gdzie $i = 1, 2, \ldots, n_1$. Wykorzystamy do tego, wynikający z lematu 2.1 oraz wniosku 4.1, następujący wniosek.

Wniosek 4.2. *Estymator* $(L_0, L_1)'Y$ *jest dopuszczalny dla wektora* $[(K'X\beta)', (Q_1'Z_1u_1)']'$ *w modelu* (4.1), *dla którego* $X \in \mathcal{R}(Z_1)$ *wtedy i tylko wtedy, gdy*

$$(L_0, L_1) \in \{[L_0(a_0), L_1(a_0, a_1)] : a_0 \in [0,1], a_1 \in [0,1]\},$$

gdzie macierze $L_0(a_0)$ *oraz* $L_1(a_0, a_1)$ *dane są wzorami* (4.3) *oraz* (4.4).

Zauważmy, że jeżeli we wniosku 4.2 wstawimy w miejsce macierzy K oraz Q_1 odpowiednio $\frac{1}{n_1 N'N}(\mathbf{1}_{n_1} \otimes N)$ i $\frac{1}{N'N}(I_{n_1} \otimes N)v_i$, to dostaniemy w jawnej postaci klasę wszystkich estymatorów dopuszczalnych dla $(\beta, u_i)'$ w modelu (2.8), gdzie v_i jest i-tym wersorem z przestrzeni $\mathbb{R}^{n_1}$.

Wniosek 4.3. *Estymator* $(L_0, L_1)'Y$ *jest dopuszczalny dla* $(\beta, u_i)'$ *w modelu* (2.8) *wtedy i tylko wtedy, gdy*

$$(L_0, L_1) \in \{[L_0(a_0), L_1(a_0, a_1)] : a_0 \in [0, 1], a_1 \in [0, 1]\},$$

gdzie

$$L_0(a_0) = \frac{a_0}{n_1 N'N}(\mathbf{1}_{n_1} \otimes N), \tag{4.7}$$

$$L_1(a_0, a_1) = \frac{a_1}{N'N}\left[I_{n_1} \otimes N - \frac{a_0}{n_1}(\mathbf{1}_{n_1}\mathbf{1}'_{n_1} \otimes N)\right] v_i. \tag{4.8}$$

Stosując dla $C = (1, 1)'$ lemat 4.3 oraz lemat 1.1, uzyskujemy następujący wniosek.

Wniosek 4.4. *Estymator* $l'Y$ *jest dopuszczalny dla* $\beta + u_i$ *w modelu* (2.8) *wtedy i tylko wtedy, gdy*

$$l'Y = [L_0'(a_0) + L_1'(a_0, a_1)]Y,$$

gdzie $L_0(a_0)$ *oraz* $L_1(a_0, a_1)$ *są odpowiednio określone wzorami* (4.7) *i* (4.8).

Jeżeli rozważamy liniowe estymatory dla $\beta + u_i$ w modelu (2.9), to wniosek ten możemy zapisać w postaci

Wniosek 4.5. *Estymator* $l'Y$ *jest dopuszczalny dla* $\beta + u_i$ *w modelu* (2.9) *wtedy i tylko wtedy, gdy*

$$l'Y = a_0\overline{Y} + a_1(\overline{Y}_{i.} - a_0\overline{Y}) \text{ \textit{pod warunkiem, że} } a_0 \in [0, 1] \text{ \textit{i} } a_1 \in [0, 1],$$

gdzie $\overline{Y}_{i.} = \frac{1}{n_2}\sum\limits_{j=1}^{n_2} Y_{ij}$ *oraz* $\overline{Y} = \frac{1}{n}\sum\limits_{i=1}^{n_1}\sum\limits_{j=1}^{n_2} Y_{ij}$.

Zauważmy, że jeżeli weźmiemy $C = (1,0)'$, to otrzymujemy dobrze znaną charakterystykę dopuszczalnych liniowych estymatorów dla β.

Wniosek 4.6. *Estymator* $\ell'Y$ *jest dopuszczalny dla* β *w modelu* (2.9) *wtedy i tylko wtedy, gdy*

$$l'Y = a_0\overline{Y} \ \textit{pod warunkiem}, \quad \textit{że} \ a_0 \in [0, 1].$$

5. Model z dwoma komponentami

W rozdziale tym interesować nas będzie dopuszczalna liniowa estymacja wektora $[(K'X\beta)', (Q_1'Z_1u_1)']'$ w ogólnym modelu z dwoma komponentami, tzn. w modelu (2.1), dla którego $k = 1$. Przypomnijmy, że na mocy lematu 2.1, równoważnie będziemy zajmować się dopuszczalną estymacją $\boldsymbol{K}'\boldsymbol{X}\beta$ w klasie $\mathcal{L}_o$ w modelu (2.4), który odpowiada modelowi z dwoma komponentami. Osłabienie dla tego ostatniego założeń, które zostały postawione w poprzednim rozdziale, tj. $\beta \in \mathbb{R}$ oraz $(Z_1Z_1')^2 = p_1Z_1Z_1'$ dla $p_1 > 0$, powoduje problemy w uzyskaniu jawnych wzorów na estymatory dopuszczalne dla wspomnianego już wektora. Problemy te wiążą się między innymi z tym, że β jest teraz p-wymiarowym wektorem nieznanych parametrów, co wpływa na postaci zbiorów $\mathcal{T}$ oraz $[\mathcal{T}]$. Oba te zbiory wyrazimy za pomocą pewnej, wygodnej dla dalszych rozważań, reparametryzacji, tzn.

$$\begin{aligned}
\mathcal{T} &= \left\{(cov(\boldsymbol{Y}), E\boldsymbol{Y}(E\boldsymbol{Y})') : \beta \in \mathcal{R}^p, \sigma_1^2 \geq 0, \sigma_2^2 > 0\right\} \\
&= \left\{\left(\rho^2 \begin{bmatrix} \lambda Z_1Z_1' + (1-\lambda)I_n & \lambda Z_1Z_1' \\ \lambda Z_1Z_1' & \lambda Z_1Z_1' \end{bmatrix}, \begin{bmatrix} X\beta\beta'X' & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\right) : \right. \\
&\qquad \left. \beta \in \mathcal{R}^p,\ \rho^2 = \sigma_1^2 + \sigma_2^2,\ \lambda = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}\right\},
\end{aligned}$$

natomiast

$$[\mathcal{T}] = \left\{(W_1(s,\lambda), W_2(\Phi)) = (s\boldsymbol{V}_\lambda, \boldsymbol{X}\Phi\boldsymbol{X}') : s \geq 0, \lambda \in [0,1], \Phi \in \mathcal{M}_p^{\geq}\right\},$$

gdzie $\boldsymbol{V}_\lambda = \begin{bmatrix} \lambda Z_1 Z_1' + (1-\lambda) I_n & \lambda Z_1 Z_1' \\ \lambda Z_1 Z_1' & \lambda Z_1 Z_1' \end{bmatrix}$.

Rozważymy dwa przypadki. Pierwszy z nich, to przypadek, kiedy iloraz wariancji efektu losowego do wariancji błędu losowego jest znany. Wówczas $\lambda = \frac{\sigma_1^2}{\sigma_1^2+\sigma_2^2} \in [0,1)$ jest znane i problem redukuje się do charakteryzacji dopuszczalnych estymatorów dla $\boldsymbol{K}'\boldsymbol{X}\beta$ w klasie $\mathcal{L}_o$ w następującym modelu Gaussa-Markova

$$\boldsymbol{Y} \sim \left(\boldsymbol{X}\beta, \rho^2 \boldsymbol{V}_\lambda\right). \tag{5.1}$$

Poniższe twierdzenie podaje warunki konieczne i dostateczne na to, aby estymator był dopuszczalny dla $\boldsymbol{K}'\boldsymbol{X}\beta$ w klasie $\mathcal{L}_o$ w powyższym modelu.

Twierdzenie 5.1. *Estymator $\boldsymbol{L}'\boldsymbol{Y}$ jest dopuszczalny dla $\boldsymbol{K}'\boldsymbol{X}\beta$ w klasie $\mathcal{L}_o$ w modelu* (5.1) *wtedy i tylko wtedy, gdy*

(i) $\mathcal{R}\left(\boldsymbol{\Pi}_o \boldsymbol{V}_\lambda \boldsymbol{L}\right) \subseteq \mathcal{R}(\boldsymbol{X})$,

(ii) $(\boldsymbol{K}-\boldsymbol{L})'\boldsymbol{\Pi}_o \boldsymbol{V}_\lambda \boldsymbol{L} \in \mathcal{M}_{2n}^{\geq}$,

gdzie $\boldsymbol{\Pi}_o = \begin{bmatrix} I_n & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \in \mathcal{M}_{2n\times 2n}$.

Dowód. Załóżmy, że macierz $\boldsymbol{L}$ spełnia warunki (i) oraz (ii). Wówczas istnieje macierz $\Phi_1 \in \mathcal{M}_p^{\geq}$, taka że $\mathcal{R}(\boldsymbol{X}\Phi_1\boldsymbol{X}') = \mathcal{R}(\boldsymbol{X}) \cap \mathcal{N}[(\boldsymbol{K}-\boldsymbol{L})']$. Oczywiście,

$$\boldsymbol{X}\Phi_1\boldsymbol{X}'(\boldsymbol{K}-\boldsymbol{L}) = 0, \tag{5.2}$$

co oznacza, że $\boldsymbol{L}'\boldsymbol{Y}$ jest lokalnie optymalny w punkcie $W = (W_1(0,\lambda), W_2(\Phi_1))$ w klasie $\mathcal{L}_o$. Zauważmy, że $\mathcal{L}_1 = \mathcal{B}(W|\mathcal{L}_o)$ może być przedstawiona następująco

$$\mathcal{L}_1 = \boldsymbol{L} + \mathcal{R}(T^t_{\boldsymbol{\Pi}_1}),$$

gdzie $\boldsymbol{\Pi}_1 = \begin{bmatrix} \Pi_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \in \mathcal{M}_{2n\times 2n}$ oraz $\Pi_1 = I_n - X\Phi_1X'(X\Phi_1X')^+$. W szczególności, jeżeli $\Phi_1 = \mathbf{0}$, to $\mathcal{L}_o = \mathcal{L}_1$ i wówczas ten krok może być pominięty.

Jeżeli natomiast $\dim(\mathcal{R}(\boldsymbol{X})) = n$ oraz jeżeli Φ_1 jest dodatnio określona, to $\boldsymbol{L}'\boldsymbol{Y}$ jest jednoznacznie lokalnie optymalnym estymatorem, a więc dopuszczalnym. W tym przypadku pierwszy krok procedury jest jednocześnie ostatnim.

Teraz pokażemy, że istnieje macierz $\Phi_2 \in \mathcal{M}_p^{\geq}$ taka, że $\boldsymbol{L}'\boldsymbol{Y}$ jest jednoznacznie lokalnie optymalny w punkcie $(W_1(1,\lambda), W_2(\Phi_2))$ w klasie $\mathcal{L}_1$.

Ponieważ $\mathcal{R}(\boldsymbol{\Pi}_1) \subseteq \mathcal{R}(\boldsymbol{\Pi}_o)$, więc z (i) mamy

$$\mathcal{R}(\boldsymbol{\Pi}_1 \boldsymbol{V}_\lambda \boldsymbol{L}) \subseteq \mathcal{R}(\boldsymbol{X}). \tag{5.3}$$

Dalej, ponieważ $\boldsymbol{\Pi}_1\boldsymbol{\Pi}_o = \boldsymbol{\Pi}_1$, więc z (ii) oraz (5.2) wynika, że

$$\begin{aligned}(\boldsymbol{K}-\boldsymbol{L})'\boldsymbol{\Pi}_o\boldsymbol{V}_\lambda\boldsymbol{L} &= (\boldsymbol{K}-\boldsymbol{L})'\boldsymbol{\Pi}_1\boldsymbol{V}_\lambda\boldsymbol{L} + (\boldsymbol{K}-\boldsymbol{L})'\boldsymbol{X}\Phi_1\boldsymbol{X}'(\boldsymbol{X}\Phi_1\boldsymbol{X}')^+\boldsymbol{L} \\ &= (\boldsymbol{K}-\boldsymbol{L})'\boldsymbol{\Pi}_1\boldsymbol{V}_\lambda\boldsymbol{L} \in \mathcal{M}_{2n}^{\geq}. \end{aligned} \tag{5.4}$$

Podstawiając $\Phi_2 = [(\boldsymbol{K}-\boldsymbol{L})'\boldsymbol{X}]^+(\boldsymbol{K}-\boldsymbol{L})'\boldsymbol{\Pi}_1\boldsymbol{V}_\lambda\boldsymbol{L}[\boldsymbol{X}'(\boldsymbol{K}-\boldsymbol{L})]^+$, na mocy (5.4), mamy $\Phi_2 \in \mathcal{M}_p^{\geq}$. Warunek (5.3) implikuje, że istnieje macierz $A \in \mathcal{M}_{p\times 2n}$ taka, że $\boldsymbol{\Pi}_1\boldsymbol{V}_\lambda\boldsymbol{L} = \boldsymbol{X}A$. Stąd

$$\begin{aligned}(\boldsymbol{K}-\boldsymbol{L})'\boldsymbol{X}\Phi_2 &= (\boldsymbol{K}-\boldsymbol{L})'\boldsymbol{X}[(\boldsymbol{K}-\boldsymbol{L})'\boldsymbol{X}]^+(\boldsymbol{K}-\boldsymbol{L})'\boldsymbol{X}A[\boldsymbol{X}'(\boldsymbol{K}-\boldsymbol{L})]^+ \\ &= (\boldsymbol{K}-\boldsymbol{L})'\boldsymbol{\Pi}_1\boldsymbol{V}_\lambda\boldsymbol{L}[\boldsymbol{X}'(\boldsymbol{K}-\boldsymbol{L})]^+.\end{aligned}$$

Ponieważ $(\boldsymbol{K}-\boldsymbol{L})'\boldsymbol{\Pi}_1\boldsymbol{V}_\lambda\boldsymbol{L}$ jest macierzą symetryczną, więc

$$(\boldsymbol{K}-\boldsymbol{L})'\boldsymbol{\Pi}_1\boldsymbol{X}\Phi_2\boldsymbol{X}'(\boldsymbol{K}-\boldsymbol{L}) = (\boldsymbol{K}-\boldsymbol{L})'\boldsymbol{\Pi}_1\boldsymbol{V}_\lambda\boldsymbol{L}. \tag{5.5}$$

Ponadto, ponieważ $\mathcal{N}(\boldsymbol{\Pi}_1\boldsymbol{X}) = \mathcal{N}[(\boldsymbol{K}-\boldsymbol{L})'\boldsymbol{\Pi}_1\boldsymbol{X}]$, więc istnieje macierz $\boldsymbol{B}$ należąca do $\mathcal{M}_{2n\times 2n}$ taka, że $\boldsymbol{B}(\boldsymbol{K}-\boldsymbol{L})'\boldsymbol{\Pi}_1\boldsymbol{X} = \boldsymbol{\Pi}_1\boldsymbol{X}$, czyli

$$\boldsymbol{B}(\boldsymbol{K}-\boldsymbol{L})'\boldsymbol{\Pi}_1\boldsymbol{V}_\lambda\boldsymbol{L} = \boldsymbol{B}(\boldsymbol{K}-\boldsymbol{L})'\boldsymbol{\Pi}_1\boldsymbol{X}A = \boldsymbol{\Pi}_1\boldsymbol{V}_\lambda\boldsymbol{L}.$$

Podobnie

$$\boldsymbol{B}(\boldsymbol{K}-\boldsymbol{L})'\boldsymbol{\Pi}_1\boldsymbol{X}\Phi_2\boldsymbol{X}'(\boldsymbol{K}-\boldsymbol{L}) = \boldsymbol{\Pi}_1\boldsymbol{X}\Phi_2\boldsymbol{X}'(\boldsymbol{K}-\boldsymbol{L}).$$

Warunek (5.5) implikuje, że

$$\mathbf{\Pi}_1 \boldsymbol{V}_\lambda \boldsymbol{L} = \mathbf{\Pi}_1 \boldsymbol{X} \Phi_2 \boldsymbol{X}'(\boldsymbol{K} - \boldsymbol{L}), \tag{5.6}$$

czyli $\boldsymbol{L}'\boldsymbol{Y}$ jest jednoznacznie lokalnie optymalny w punkcie $(W_1(1,\lambda), W_2(\Phi_2))$ w klasie $\mathcal{L}_1$, co kończy pierwszą część dowodu.

Teraz załóżmy, że $\boldsymbol{L}'\boldsymbol{Y}$ jest dopuszczalny dla $\boldsymbol{K}'\boldsymbol{X}\beta$ w klasie $\mathcal{L}_o$. Ponieważ $\boldsymbol{L} \in \mathcal{L}_1$, gdzie rozmaitość liniowa $\mathcal{L}_1$ została zdefiniowana na początku dowodu, więc estymator $\boldsymbol{L}'\boldsymbol{Y}$ jest dopuszczalny dla $\boldsymbol{K}'\boldsymbol{X}\beta$ w klasie $\mathcal{L}_1$. Zatem na mocy twierdzenia 1.2 istnieje punkt $W = (W_1(s,\lambda), W_2(\Phi_2)) \in \mathcal{A}(\mathcal{L}_1)$, gdzie $s > 0$ taki, że

$$s\mathbf{\Pi}_1 \boldsymbol{V}_\lambda \boldsymbol{L} = \mathbf{\Pi}_1 \boldsymbol{X} \Phi_2 \boldsymbol{X}'(\boldsymbol{K} - \boldsymbol{L}).$$

Ponieważ $\mathcal{R}(\mathbf{\Pi}_o - \mathbf{\Pi}_1) \subseteq \mathcal{R}(\boldsymbol{X})$, więc $\mathcal{R}(\mathbf{\Pi}_1 \boldsymbol{X}) \subseteq \mathcal{R}(\boldsymbol{X})$. Stąd

$$\mathcal{R}(\mathbf{\Pi}_o \boldsymbol{V}_\lambda \boldsymbol{L}) = \mathcal{R}((\mathbf{\Pi}_o - \mathbf{\Pi}_1)\boldsymbol{V}_\lambda \boldsymbol{L}) \oplus \mathcal{R}(\mathbf{\Pi}_1 \boldsymbol{V}_\lambda \boldsymbol{L}) \subseteq \mathcal{R}(\boldsymbol{X}),$$

co kończy dowód warunku (i).

Dalej, ponieważ $(\boldsymbol{K} - \boldsymbol{L})'\mathbf{\Pi}_o \boldsymbol{V}_\lambda \boldsymbol{L} = (\boldsymbol{K} - \boldsymbol{L})'\mathbf{\Pi}_1 \boldsymbol{V}_\lambda \boldsymbol{L}$ (co wynika z (5.4)) oraz $\boldsymbol{X}'(\boldsymbol{K} - \boldsymbol{L}) = \boldsymbol{X}'\mathbf{\Pi}_1(\boldsymbol{K} - \boldsymbol{L})$, więc

$$s(\boldsymbol{K} - \boldsymbol{L})'\mathbf{\Pi}_o \boldsymbol{V}_\lambda \boldsymbol{L} = (\boldsymbol{K} - \boldsymbol{L})'\mathbf{\Pi}_1 \boldsymbol{X} \Phi_2 \boldsymbol{X}'\mathbf{\Pi}_1(\boldsymbol{K} - \boldsymbol{L}) \in \mathcal{M}_{2n}^{\geq}.$$

□

Rozważmy teraz ogólniejszą sytuację, kiedy iloraz σ^2 oraz σ_e^2 jest nieznany. Dodatkowo załóżmy, że $\mathcal{R}(\boldsymbol{X}) \subseteq \mathcal{R}(\boldsymbol{Z})$, gdzie $\boldsymbol{Z} = [Z_1', 0]' \in \mathcal{M}_{2n\times m}$.

Twierdzenie 5.2. *Estymator* $\boldsymbol{L}'\boldsymbol{Y}$ *jest dopuszczalny dla* $\boldsymbol{K}'\boldsymbol{X}\beta$ *w klasie* $\mathcal{L}_o$ *wtedy i tylko wtedy, gdy istnieje* $\lambda \in [0,1]$, *dla której*

(i) $\mathcal{R}(\mathbf{\Pi}_o \boldsymbol{V}_\lambda \boldsymbol{L}) \subseteq \mathcal{R}(\boldsymbol{X})$,

(ii) $(\boldsymbol{K} - \boldsymbol{L})'\mathbf{\Pi}_o \boldsymbol{V}_\lambda \boldsymbol{L} \in \mathcal{M}_{2n}^{\geq}$, *oraz dodatkowo, gdy* $\lambda = 1$

(iii) $\mathbf{\Pi}_{Z_1}\boldsymbol{L} = \mathbf{0}$,

gdzie $\mathbf{\Pi}_{Z_1} = \begin{bmatrix} \Pi_{Z_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$, *natomiast* Π_{Z_1} *jest macierzą rzutu ortogonalnego na przestrzeń* $\mathcal{N}(Z_1')$.

Dowód. W przypadku, gdy $\lambda \in [0, 1)$ dowód warunku dostatecznego wynika z twierdzenia 5.1. Jeżeli $\lambda = 1$, to musimy wykonać dodatkowy (trzeci) krok. Ponieważ

$$\mathbf{\Pi}_1\boldsymbol{V}_1\boldsymbol{L} = \mathbf{\Pi}_1\boldsymbol{X}\Phi_2\boldsymbol{X}'(\boldsymbol{K} - \boldsymbol{L}),$$

gdzie

$$\Phi_2 = [(\boldsymbol{K} - \boldsymbol{L})'\boldsymbol{X}]^+(\boldsymbol{K} - \boldsymbol{L})'\mathbf{\Pi}_1\boldsymbol{V}_1\boldsymbol{L}[\boldsymbol{X}'(\boldsymbol{K} - \boldsymbol{L})]^+$$

(patrz (5.6)), więc $\boldsymbol{L}'\boldsymbol{Y}$ jest lokalnie optymalny w punkcie $(W_1(1, 1), W_2(\Phi_2))$ w klasie $\mathcal{L}_1$. Rozmaitość liniowa $\mathcal{L}_2 = \mathcal{B}(W_1(1, 1), W_2(\Phi_2))|\mathcal{L}_1)$ ma następującą postać

$$\mathcal{L}_2 = \boldsymbol{L} + \mathcal{R}(T^t_{\mathbf{\Pi}_2}),$$

gdzie $\mathbf{\Pi}_2 = \begin{bmatrix} \Pi_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \in \mathcal{M}_{2n\times 2n}$ oraz $\Pi_2 = \Pi_1[I_n - \Pi_1 Z_1 Z_1'\Pi_1(\Pi_1 Z_1 Z_1'\Pi_1)^+]$. Macierz Π_2 jest macierzą rzutu ortogonalnego na $\mathcal{N}(X\Phi_1 X') \cap \mathcal{N}(\Pi_1 Z_1 Z_1'\Pi_1)$. Ponieważ $\mathcal{R}(\mathbf{\Pi}_{Z_1}) \subseteq \mathcal{R}(\mathbf{\Pi}_1)$, więc $\mathbf{\Pi}_2 = \mathbf{\Pi}_{Z_1}$ oraz $\mathbf{\Pi}_{Z_1}\boldsymbol{X} = \mathbf{0}$. Korzystając z warunku (iii) mamy, że $\mathbf{\Pi}_{Z_1}\boldsymbol{L} = \mathbf{\Pi}_{Z_1}\boldsymbol{X} = \mathbf{0}$. To implikuje z kolei, że $\boldsymbol{L}'\boldsymbol{Y}$ jest jednoznacznie lokalnie optymalny w $(W_1(1, 0), W_2(0))$ w klasie $\mathcal{L}_2$.

Załóżmy teraz, że $\boldsymbol{L}'\boldsymbol{Y}$ jest dopuszczalny dla $\boldsymbol{K}\boldsymbol{X}\beta$ w klasie $\mathcal{L}_o$. W oparciu o dowód twierdzenia 5.1 pozostał do rozważenia przypadek, kiedy $\boldsymbol{L} \in \mathcal{L}_2 = \mathcal{B}(\,(W_1(1, 1), W_2(\Phi_2))\,|\mathcal{L}_1) \subset \mathcal{L}_1 = \mathcal{B}(\,(W_1(0, \lambda), W_2(\Phi_1))\,|\mathcal{L}_o)$. Z lokalnej optymalności w punkcie $(W_1(1, 1), W_2(\Phi_2))$ nadal wynikają warunki (i) oraz (ii) dla $\lambda = 1$. Ponieważ $\boldsymbol{L}'\boldsymbol{Y}$ jest dopuszczalny dla $\boldsymbol{K}\boldsymbol{X}\beta$ w klasie $\mathcal{L}_2$ i ponieważ

$\mathcal{R}(\boldsymbol{X}) \subseteq \mathcal{R}(\boldsymbol{Z})$, więc korzystając z twierdzenia 1.2 mamy, że istnieje punkt $(W_1(s,0), W_2(0)) \in \mathcal{A}(\mathcal{L}_2)$ z $s > 0$, dla którego $s\boldsymbol{\Pi}_Z\boldsymbol{L} = \boldsymbol{0}$, co kończy dowód. $\square$

W szczególnym przypadku, gdy $\beta \in \mathbb{R}$, $X \in \mathcal{R}(Z_1)$ oraz $(Z_1Z_1')^2 = p_1Z_1Z_1'$ dla $p_1 > 0$ z charakterystyki estymatorów dopuszczalnych dla $\boldsymbol{K}'\boldsymbol{X}\beta$ w klasie $\mathcal{L}_o$ podanej w powyższym twierdzeniu wynika charakterystyka podana we wniosku 4.1 i odwrotnie. Korzystając z lematu 1.1, wystarczy wykazać równoważność pomiędzy tymi charakterystykami dla $K = I_n$ i $Q_1 = I_n$. Wtedy dowód będzie bardziej przejrzysty.

Warunki (i) oraz (ii) twierdzenia 5.2 możemy wyrazić w terminach macierzy L_0 i L_1 następująco

(i') $\mathcal{R}(V_\lambda L_0) \subseteq \mathcal{R}(X)$, $\mathcal{R}(V_\lambda L_1 - \lambda Z_1Z_1') \subseteq \mathcal{R}(X)$,

(ii') $$\begin{bmatrix} (I_n - L_0)'V_\lambda L_0 & (I_n - L_0)'(V_\lambda L_1 - \lambda Z_1Z_1') \\ -L_1V_\lambda L_0 & -L_1(V_\lambda L_1 - \lambda Z_1Z_1') \end{bmatrix} \in \mathcal{M}^{\geq}_{2n},$$

gdzie $V_\lambda = \lambda Z_1Z_1' + (1-\lambda)I_n$.

Załóżmy teraz, że istnieje $\lambda \in [0,1)$, dla której te warunki są spełnione. Na mocy warunku (ii') mamy, że macierz $(I_n - L_0)'V_\lambda L_0$ jest symetryczna, tym samym macierz $V_\lambda L_0$ jest też symetryczna. Zatem pierwszy z warunków (i') implikuje, że istnieje $c \in \mathbb{R}$ takie, że $V_\lambda L_0 = cXX'$. Stąd

$$\begin{aligned} L_0 &= cV_\lambda^{-1}XX' = \left(\frac{-\lambda}{(1-\lambda)(p_1\lambda + 1 - \lambda)}Z_1Z_1' + \frac{1}{1-\lambda}I_n\right)XX' \\ &= \frac{c}{p_1\lambda + 1 - \lambda}XX'. \end{aligned} \tag{5.7}$$

Z symetrii macierzy występującej w warunku (ii'), dotyczących elementów poza diagonalą, wynika również, że

$$V_\lambda L_1 - \lambda Z_1Z_1' = -\lambda L_0'Z_1Z_1'.$$

Zatem z (5.7) otrzymujemy, że

$$\begin{aligned} L_1 &= V_\lambda^{-1}\left(\frac{-cp_1\lambda}{p_1\lambda+1-\lambda}XX' + \lambda Z_1Z_1'\right) \\ &= \frac{p_1\lambda}{p_1\lambda+1-\lambda}\left(\frac{1}{p_1}Z_1Z_1' - \frac{cX'X}{p_1\lambda+1-\lambda}\frac{1}{X'X}XX'\right). \end{aligned} \tag{5.8}$$

Korzystając z (5.7) dostajemy

$$(I_n - L_0)'V_\lambda L_0 = c\left(1 - \frac{cX'X}{p_1\lambda+1-\lambda}\right)XX'.$$

Z nieujemnej nieokreśloności tej macierzy wynika, że $c\left(1 - \frac{cX'X}{p_1\lambda+1-\lambda}\right) \geq 0$. Stąd $c \in [0, \frac{p_1\lambda+1-\lambda}{X'X}]$ lub równoważnie $\frac{cX'X}{p_1\lambda+1-\lambda} \in [0,1]$.

Załóżmy teraz, że $\lambda = 1$. Z warunku (iii) twierdzenia 5.2 dostajemy, że

$$L_0 = \frac{1}{p_1}Z_1Z_1'L_0 \tag{5.9}$$

oraz

$$L_1 = \frac{1}{p_1}Z_1Z_1'L_1. \tag{5.10}$$

Z warunku (i) tego twierdzenia oraz symetrii macierzy $Z_1Z_1'L_0$ wynika, że istnieje $c_* \in \mathbb{R}$ takie, że $Z_1Z_1'L_0 = c_*XX'$. Korzystając z (5.9) mamy, że

$$L_0 = \frac{c_*}{p_1}XX'.$$

Postępując podobnie, jak w przypadku $\lambda \in [0,1)$ otrzymujemy

$$Z_1Z_1'L_1 - Z_1Z_1' = -c_*XX'.$$

Stąd i z (5.10) dostajemy następującą postać macierzy

$$L_1 = \left(\frac{1}{p_1}Z_1Z_1' - \frac{c_*X'X}{p_1}\frac{1}{X'X}XX'\right).$$

Dalej, z nieujemnej określoności macierzy $(I_n - L_0)'V_1L_0$ mamy $c_* \in [0, \frac{p_1}{X'X}]$.

Pokazaliśmy zatem, że jeżeli istnieje $\lambda \in [0,1]$, dla której spełnione są warunki twierdzenia 5.2, to estymator $\boldsymbol{L'Y}$ dla $\boldsymbol{X}\beta$ jest postaci podanej we wniosku 4.1.

Chcąc pokazać, że z charakterystyki podanej we wniosku 4.1 wynika charakterystyka podana w twierdzeniu 5.2, wystarczy podstawić $\lambda = \frac{a_1}{a_1+p_1(1-a_1)}$. Oczywiście, jeżeli $a_1 \in [0,1]$, to $\lambda \in [0,1]$. Wykonując standardowe przekształcenia, wykorzystujące zależność $Z_1Z_1'XX' = p_1XX'$, dostajemy żądane wynikanie.

Podsumowanie

W pracy rozważany był następujący model liniowy

$$Y = X\beta + Z_1u_1 + Z_2u_2 + \ldots + Z_ku_k + e,$$

w którym zajmowano się dopuszczalną liniową estymacją wektora

$$[(K'X\beta)', (Q_1'Z_1u_1)', \ldots, (Q_k'Z_ku_k)']'.$$

Kluczowym w całej rozprawie jest spostrzeżenie, że zagadnienie to można sprowadzić do równoważnego problemu dopuszczalności estymatora dla $\boldsymbol{K'X}\beta$ w klasie $\mathcal{L}_o$ w modelu

$$\boldsymbol{Y} \sim \left(\boldsymbol{X}\beta, \sum_{i=1}^{k+1} \sigma_i^2 \boldsymbol{V}_i\right).$$

Uzyskano w jawnej postaci charakterystykę dopuszczalnych estymatorów dla $\boldsymbol{K'X}\beta$ w klasie $\mathcal{L}_o$ w modelach odpowiadających zrównoważonemu modelowi losowemu k – kierunkowej klasyfikacji hierarchicznej, zrównoważonemu modelowi losowemu k – kierunkowej klasyfikacji krzyżowej oraz zrównoważonemu modelowi losowemu dwukierunkowej klasyfikacji krzyżowej z interakcją. Opisano także jawną postać dopuszczalnych estymatorów $\boldsymbol{L'Y}$ dla $\boldsymbol{K'X}\beta$ w klasie $\mathcal{L}_o$ w modelu odpowiadającym modelowi z dwoma komponentami, w którym $\beta \in \mathbb{R}$, a macierz Z_1Z_1' jest macierzą proporcjonalną do macierzy idempotentnej. Dla ogólnego modelu liniowego z dwoma komponentami przy założeniu, że $\mathcal{R}(X) \subseteq \mathcal{R}(Z_1)$ podano warunki konieczne i dostateczne na to, aby odpowiedni estymator był dopuszczalny. Warunki te zostały wyrażone w terminach podobnych do tych, których użył Rao (1976) w modelu Gaussa-Markowa.

Bibliografia

[1] J.K. Baksalary, A. Markiewicz, C.R. Rao, *Admissible linear estimation in the general Gauss-Markov model with respect to an arbitrary quadratic risk function*, J. Statist. Planning and Inference **44** (1995), 341-347.

[2] A. Cohen, *All admissible linear estimators of the mean vector*, Ann. Math. Statist. **37** (1966), 458-463.

[3] H. Drygas, R. Zmyślony, *On admissible estimation for parametric functions in linear models*, Statistical Papers **29** (1988), 113-123.

[4] S. Gnot, Estymacja Komponentów Wariancyjnych w Modelach Liniowych, WNT, Warszawa 1991.

[5] S. Gnot, E. Rafajłowicz, A. Urbańska-Motyka, *Statistical inference in a linear model for spatially located sensors and random input*, Ann. Inst. statist. Math. **52.2** (2001), 370-379.

[6] D.A. Harville, Matrix Algebra From a Statistician's Perspective, Springer-Verlag, New York 2000.

[7] D.A. Harville, *Extension of the Gauss-Markov theorem to include the estimation of random effects*, Ann. of Statist. **4** (1976), 384-395.

[8] K. Hoffman, *Admissibility of linear estimators with respect to restricted parameter sets*, Math. Operationsforsch. Statist., Ser. Statistics **4** (1977), 425-438.

[9] W. Klonecki, *Linear estimators of the mean vector in linear models, problem of admissibility*, Prob. Math. Statist. **2** (1982).

[10] W. Klonecki, S. Zontek, *On the structure of admissible linear estimators*, J. Multivariate Anal. **24** (1988), 11-30.

[11] L.R. LaMotte, *Admissibility in linear estimation*, Ann. Statist. **10.1** (1982), 245-255.

[12] L.R. LaMotte, *On limits of uniquely best linear estimators*, Metrika **45** (1997), 197-211.

[13] T. Mathew, C.R. Rao, B.K. Sinha, *Admissible linear estimation in singular linear models*, Comm. Statist., Theory Methods **13** (1984), 3033-3045.

[14] J. Gross, A. Markiewicz, *On admissibility of linear estimators with respect to the mean square error matrix criterion under the general mixed linear model*, Statistics **33** (1999), 57-71.

[15] A. Olsen, J. Seely, D. Birkes, *Invariant quadratic estimation for two variance components*, Ann. Statist. **4** (1976), 591-612.

[16] J.L. Peixoto, D.A. Harville, *Comparisons of alternative predictors under the balanced one-way random model*, JASA **81** (1986), 431-436.

[17] C.R. Rao, *Estimation of parameters in a linear model*, Ann. Statist. **4** (1976), 1023-1037.

[18] C.R. Rao, Modele Liniowe Statystyki Matematycznej, PWN, Warszawa 1982.

[19] C.R. Rao, *Estimation in a linear models with mixed effects: a unified theory*, Proceedings Second International Tampere Conference in Statistics., Dept. of Mathematical Sciences, Univ. of Tempere, (1987), 73-98.

[20] G.K. Robinson, *That BLUP is good thing: the estimation of random effects*, Statistical Science **6** (1991), 15-51.

[21] N. Shinozaki, *A Study of Generalized Inverse of Matrix and Estimation with Quadratic Loss*, Ph.D. thesis, Keio University, Japan 1975.

[22] R. Searle, G. Casella, C. McCulloch, Variance Components, Wiley, New York 1992.

[23] C. Stępniak, *A complete class for linear estimation in a general linear model*, Ann. Inst. Statist. Math. **39** (1987), 563-573.

[24] E. Synówka-Bejenka, S. Zontek, *A characterization of admissible linear estimators of fixed and random effects in linear midels*, (praca wysłana do druku).

[25] S. Zontek, *Estymacja parametrów w modelach liniowych*, Rozprawa doktorska, IM PAN 1983.

[26] S. Zontek, *Admissibility of limits of the unique locally best linear estimators with application to variance components models*, Probab. Math. Statist. **9** (1988), 29-44.