Uniwersytet Ekonomiczny w Katowicach

Studia Ekonomiczne. Zeszyty Naukowe
Uniwersytetu Ekonomicznego w Katowicach
ISSN 2083-8611 Nr 288 · 2016

Informatyka i Ekonometria 5

Dominik Krężołek
Uniwersytet Ekonomiczny w Katowicach
Wydział Informatyki i Komunikacji
Katedra Demografii i Statystyki Ekonomicznej
dominik.krezolek@ue.katowice.pl

MODEL REGRESJI GRZBIETOWEJ I JEGO WYKORZYSTANIE DO OCENY RYZYKA INWESTYCYJNEGO – PRZYPADEK RYNKU METALI

Streszczenie: Modele regresji są powszechnie wykorzystywanymi narzędziami statystyki, a także innych nauk ilościowych, służącymi do wykrywania związków w obrębie analizowanych danych oraz do prowadzenia predykcji wielkości zestawu zmiennych objaśnianych na podstawie realizacji zestawu zmiennych objaśniających. Istnieje wiele metod umożliwiających szacowanie nieznanych parametrów modeli regresji, m.in. MNK, MNW, MM, jednakże nie zawsze uzyskane estymatory spełniają wymagane założenia (co do swoich własności oraz co do własności modelu). Istotny stopień współliniowości uniemożliwia właściwe wnioskowanie na podstawie modeli klasycznych. W artykule podjęto próbę wykorzystania regresji grzbietowej do modelowania ryzyka inwestycji na rynku metali. Model wykorzystuje parametr kary, umożliwiający redukcję zmiennych współliniowych, a tym samym uzyskanie prostszej postaci funkcji regresji. Dodatkowo zmniejsza się obciążenie oraz wariancja estymatorów parametrów modelu.

Słowa kluczowe: model regresji, regresja grzbietowa, rynek metali, współliniowość, ryzyko.

Wprowadzenie

Modelowanie zjawisk gospodarczych jest jednym z głównych obszarów zastosowania narzędzi nauk ilościowych w praktyce. Poszukiwanie pewnych formalnych relacji w zbiorze zjawisk ekonomicznych prowadzi niejednokrotnie do istotnego ich uproszczenia. Model, jako uogólniona reprezentacja rzeczywistości, nie jest strukturą idealną, gdyż z reguły oprócz tego, że zazwyczaj zbytnio symplifikuje rzeczywistość, to ponadto niejednokrotnie zdecydowana większość jego teoretycznych założeń nie jest spełniona [Maddala, 2006]. Poprawny model winien realnie oraz rzetelnie aproksymować analizowane zjawisko. Nauka ekonomii, w której szeroko wykorzystywane są narzędzia modelowania, rozróżnia, w sposób jednoznaczny, model ekonomiczny oraz model ekonometryczny. Model ekonomiczny jest definiowany jako zbiór określonych założeń i aksjomatów, opisujących funkcjonowanie gospodarki jako całości, natomiast model ekonometryczny zdefiniować można jako swoiste relacje formalne zapisane za pomocą określonych równań matematycznych. Relacje te, co jest oczywiste, powinny być zgodne z teorią ekonomii [Welfe, 2003].

Jedną z funkcji modelu ekonometrycznego, oprócz formalnej deskrypcji zjawisk gospodarczych, jest funkcja predykcyjna. Umożliwia ona prowadzenie prognoz zjawisk ekonomicznych. Stąd też poprawna konstrukcja teoretyczna powinna bazować na pełnej informacji, dotyczącej takiego zjawiska w przeszłości. Niejednokrotnie obserwowalne są zakłócenia i różnego rodzaju szumy informacyjne, zniekształcające model. Ekonometria określa je terminem „składnik losowy". Wprowadzanie zaburzeń losowych do konstruktów teoretycznych jest istotne, gdyż niejednokrotnie nie ma możliwości wyspecyfikowania wszystkich możliwych charakterystyk, opisujących analizowany problem. Ponadto, dużą rolę w funkcjonowaniu podmiotów gospodarczych, generujących informacje ekonomiczne odgrywa człowiek, który często podejmuje decyzje, niekoniecznie zgodne z ogólnie przyjętymi dogmatami.

Odchodząc od kwestii merytorycznych definiujących model ekonomiczny i ekonometryczny, istotną rolę odgrywa jego wykorzystanie. W prezentowanej pracy podjęto próbę opisu jednego z liniowych modeli regresyjnych, rozpoznawanego w literaturze pod pojęciem „regresji grzbietowej" oraz jego zastosowania w ocenie ryzyka inwestycyjnego na rynku metali.

1. Ogólny model regresji liniowej

Modelowanie regresyjne oraz analiza regresji to grupa metod statystyczno-ekonometrycznych, opisująca relacje pomiędzy pewnymi mierzalnymi wielkościami, związanymi z badanymi zjawiskami, które to wielkości mają charakter stochastyczny. Niech dany będzie model ogólny postaci:

$$y = f(x_1, x_2, \dots, x_k) + \varepsilon \tag{1}$$

gdzie y oznacza zmienną objaśnianą (regresant, zmienna zależna), x oznacza zbiór k zmiennych objaśniających (regresorów, zmienne niezależne), $f(\cdot)$ jest pewną funkcją spełniającą zadość odwzorowaniu $R^k \to R$, natomiast ε jest realizacją składnika losowego. Zakłada się ponadto, że wartość oczekiwana składnika losowego jest równa zero.

Niejednokrotnie reprezentacja matematyczna postaci $f(\cdot)$ jest znana, z wyjątkiem współczynników kierunkowych $\beta_i, (i = 1, \dots, k)$ przy poszczególnych regresorach, nazywanych parametrami strukturalnymi modelu. W pracy ograniczono się do szczególnej postaci funkcji $f(\cdot)$, a mianowicie do postaci funkcji liniowej względem parametrów β_k.

Analizowanie zjawisk ekonomicznych przy wykorzystaniu narzędzi regresji determinuje konieczność określenia zmiennej zależnej oraz zmiennej niezależnej. Jednakże koncentrowanie się na podejściu uwzględniającym tylko jedną zmienną objaśniającą (model regresji liniowej prostej), zdaje się dalece idącym uproszczeniem, abstrakcyjnym w kontekście obserwowanych relacji pomiędzy zjawiskami gospodarczymi. Tym samym zdefiniowano tzw. liniowy model regresji wielorakiej, którego reprezentacja jest następująca:

$$y_i = f(x_{i1}, x_{i2}, \dots, x_{ik}; \beta_1, \beta_2, \dots, \beta_k) + \varepsilon_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i \quad (2)$$

gdzie $y_i \in R$, x_{ik} jest k-wymiarowym wektorem zmiennych objaśniających, β_k jest wektorem nieznanych parametrów strukturalnych modelu, natomiast ε_i jest składnikiem losowym, spełniającym postulat i.i.d.

Popularną jest reprezentacja macierzowa postaci $\mathrm{y} = \mathrm{X}\beta + \varepsilon$. Problem analityczny polega na oszacowaniu nieznanych parametrów β_k na podstawie realizacji zmiennej zależnej i zmiennych niezależnych. Istnieje wiele metod estymacji parametrów liniowego modelu regresji, jednakże największą popularność zyskały Metoda Najmniejszych Kwadratów (MNK) oraz Metoda Największej Wiarygodności (MNW). Większą efektywnością estymatorów cechuje się metoda MNW, jednakże wymaga ona znajomości funkcji gęstości prawdopodobieństwa analizowanych zmiennych, co w praktyce jest rzadkością.

Główne założenia modelu regresji liniowej postulują iż: (a) wartość oczekiwana składnika losowego $E[\varepsilon_i] = 0$, (b) wariancja składnika losowego $D^2[\varepsilon_i] = \sigma^2$ oraz że (c) składniki losowe są niezależne. Wprowadzając dodatkowe założenie, że ów składnik losowy posiada rozkład normalny, estymacja parametrów strukturalnych modelu (2) w formie macierzowej sprowadza się do rozwiązania zagadnienia optymalizacji, minimalizującego wyrażenie:

$$\theta(\beta) = (\mathrm{y} - \mathrm{X}\beta)^T (\mathrm{y} - \mathrm{X}\beta) \quad (3)$$

W przypadku spełnienia powyższych założeń, wyrażenie (3) wyznacza takie same estymatory nieznanych parametrów modelu, niezależnie od przyjętej metody estymacji MNK czy MNW. Rozwiązaniem jest wyrażenie:

$$\hat{\beta} = (\mathrm{X}^T \mathrm{X})^{-1} \mathrm{X}^T \mathrm{y} \quad (4)$$

Powyższy wzór pozwala na oszacowanie wektora nieobciążonych estymatorów nieznanych parametrów modelu, cechujących się ponadto minimalną wariancją. Należy jednoznacznie wskazać, kiedy wariancję estymatora należy uznać za relatywnie małą. Rodzi się problem struktury macierzy X^TX, która może być macierzą „okołoosobliwą”[1], a tym samym niemożliwą do odwrócenia. Stąd wyznaczenie ocen parametrów β_k jest nieosiągalne. Rozwiązaniem jest wprowadzenie pewnego dodatkowego parametru w modelu, zwanego parametrem kary, definiując tym samym nową strukturę modelu regresji, znanego w literaturze pod nazwą „regresji grzbietowej” (ang. *ridge regression*).

2. Model regresji grzbietowej

Prezentowany w poprzednim podrozdziale model regresji wielorakiej zakłada, iż składnik losowy ε_i ma wartość oczekiwaną równą zero oraz wariancję równą σ^2. Ostatni warunek można zapisać za pomocą macierzowej struktury kowariancyjnej jako $E[\varepsilon^T\varepsilon] = \sigma^2 I_n$, gdzie I_n jest macierzą jednostkową o wymiarze $n \times n$. Procedura estymacji parametrów modelu (2) poprzez oszacowanie wektora (4), spełniającego warunek nieobciążoności i największej efektywności, znana jest pod nazwą procedury Gaussa-Markova. Jednakże stosowanie jej wymaga spełnienia zadość założeniu, iż macierz korelacji X^TX jest macierzą jednostkową (lub prawie jednostkową). Jeżeli to założenie nie jest spełnione, to oszacowania najmniejszych kwadratów są obciążone błędem składnika losowego modelu.

A.E. Hoerl i R.W. Kennard [1970] zaproponowali rozwiązanie tego problemu poprzez uzupełnienie komponentu $(X^TX)^{-1}$ wektora parametrów $\hat{\beta}$ parametrem kary λ, usuwając osobliwość macierzy X^TX. Stąd uzyskano następujący estymator grzbietowy modelu regresji $y = X\beta + \varepsilon$ postaci:

$$\tilde{\beta} = (X^TX - \lambda I_n)^{-1}X^Ty \tag{5}$$

będący rozwiązaniem problemu optymalizacyjnego postaci:

$$\min\{\theta(\beta)\} = \min\{(y - X\beta)^T(y - X\beta) + \lambda X^TX\} \tag{6}$$

Parametr kary określony jest na zbiorze $R_+ \cup \{0\}$, przy czym wartość $\lambda = 0$ generuje estymatory parametrów modelu zgodne z MNK. Dodatkowo, wraz ze wzrostem wartości parametru kary zmniejsza się suma wariancji komponentów estymatora grzbietowego $\tilde{\beta}_k$, co powoduje, iż jego średni błąd kwadratowy generuje niższe wartości niż ten sam, uzyskany za pomocą MNK. Tym samym pojawia się

[1] O wyznaczniku w otoczeniu wartości zero [przyp. aut.].

problem ustalenia optymalnej wielkości parametru λ, silnie zależnego od ilości parametrów wejściowego modelu. A.E. Hoerl i R.W. Kennard sugerują, by jego wartość szacować za pomocą estymatora $\hat{\sigma}^2$ wariancji resztowej modelu, która jest nieobciążonym estymatorem nieznanej wariancji σ^2 oraz za pomocą ocen MNK dla parametrów strukturalnych $\hat{\beta}_k$ badanego modelu.

W literaturze istnieje wiele alternatywnych metod szacowania parametru kary jak również wiele alternatywnych zapisów modeli regresji grzbietowej [Becker, Fried, Kuhnt, 2013]. Najpopularniejsze z nich zaprezentowano w tab. 1.

Tabela 1. Najpopularniejsze estymatory parametru kary λ

	Estymator Hoerla i Kennarda	**Estymator Hoerla, Kennarda i Baldwina**	**Estymator Lawlessa i Wanga**
Estymator parametru λ	$\hat{\lambda}_{HK} = \frac{\hat{\sigma}^2}{\hat{\beta}^2_{max}}$	$\hat{\lambda}_{HKB} = \frac{k\hat{\sigma}^2}{\sum_{i=1}^{k}\hat{\beta}_i^2}$	$\hat{\lambda}_{LW} = \frac{k\hat{\sigma}^2}{\hat{\beta}^T X^T X\hat{\beta}}$

Źródło: [Hoerl, Kennard, 1970, s. 64; Hoerl, Kennard, Baldwin, 1975, s. 111; Lawless, Wang, 1976, s. 314].

B.M.G. Kibria [2003] zaproponował, by parametr kary szacować przy wykorzystaniu średniej arytmetycznej (7), geometrycznej (8) bądź mediany (9). Tym samym formuły są następujące:

$$\hat{\lambda}_{AM} = \frac{\sum_{i=1}^{k}\frac{\hat{\sigma}^2}{\hat{\beta}_i^2}}{k} \tag{7}$$

$$\hat{\lambda}_{GM} = \hat{\sigma}^2\left(\prod_{i=1}^{k}\hat{\beta}_i^2\right)^{-\frac{1}{k}} \tag{8}$$

$$\hat{\lambda}_{Me} = Me\{\frac{\hat{\sigma}^2}{\hat{\beta}_i^2}\} \tag{9}$$

Jak wynika z powyższych wzorów estymatory parametru kary λ zależne są od liczby regresorów, uwzględnionych w wejściowym modelu. Ponadto, można wykazać, iż przy założeniu jednakowych wartości parametrów strukturalnych modelu wszystkie prezentowane estymatory są zbieżne do wartości $\frac{\hat{\sigma}^2}{\hat{\beta}^2}$, pierwotnie zaproponowanej przez A.E. Hoerla i R.W. Kennarda.

3. Regresyjny model rynkowy

Analiza regresji znajduje zastosowanie w wielu dziedzinach nauki, jednakże szczególne znaczenie ma ona w analizach rynkowych, w tym dla danych pochodzących z rynków finansowych. W obrębie tej kategorii rynków najpopularniejsze są modele ogólnie definiowane jako modele wskaźnikowe, estymujące wielkości wybranych aktywów finansowych w zależności od szeregu różnych czynników, które istotnie mogą kształtować ich poziom. W prezentowanej pracy skoncentrowano się na modelach określanych jako modele arbitrażowe, które zostały zaproponowane przez S.A. Rossa [1976]. Model ten przyjmuje formalną postać:

$$r_i = \beta_0 + \beta_1 F_{i1} + \cdots + \beta_k F_{ik} + \varepsilon_i \tag{10}$$

gdzie r oznacza zmienną ryzyka, której wartości zależne są od zestawu pewnych k-czynników F_k.

Czytelnik bez problemu rozpozna w modelu (10) reprezentację zapisaną wzorem (2). W analizie ryzyka model arbitrażowy umożliwia pomiar wrażliwości zmiennej ryzyka na kształtowanie się determinujących ją czynników [Jajuga, 2009]. Przy założeniu zasady *ceteris paribus* można interpretować wpływ poszczególnych determinant na poziom zmiennej zależnej. Ograniczając wzór (10) do wariantu jednoczynnikowego i określając ten czynnik jako indeks rynkowy, definiuje się tzw. jednoczynnikowy model Sharpe'a. Współczynnik kierunkowy w modelu Sharpe'a, określany jako współczynnik beta, informuje o wrażliwości zmian zmiennej ryzyka zależnie od kierunku zmian rynku. W prezentowanej pracy wykorzystano podejście wielowskaźnikowe.

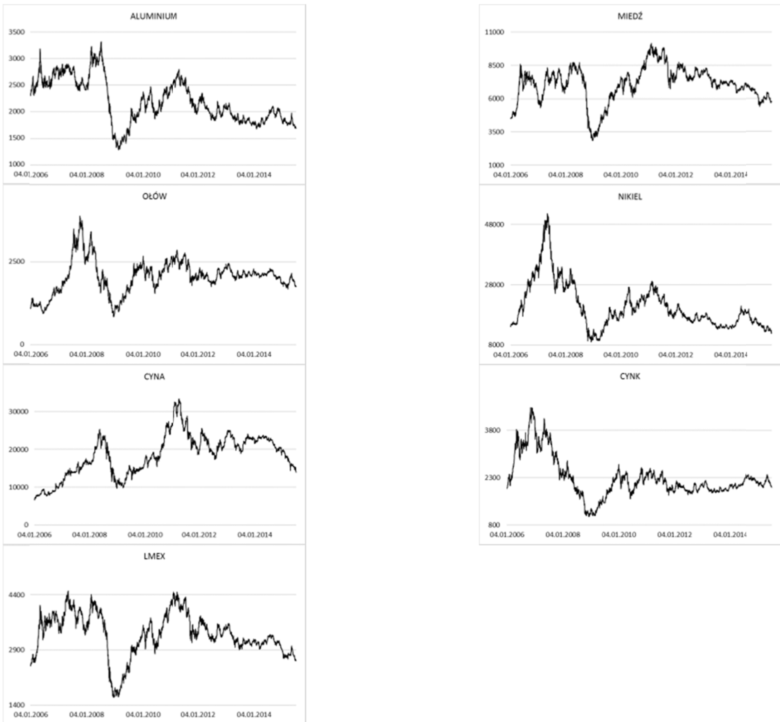
4. Zastosowanie modelu regresji grzbietowej w ocenie ryzyka inwestycji na rynku metali

Poziom rozwoju współczesnych gospodarek zarówno tych wschodzących, jak i o ugruntowanej międzynarodowej pozycji, determinowany jest bardzo wieloma czynnikami. W prezentowanej pracy wybrano jedną z gałęzi gospodarki, istotną z towarowego oraz finansowego punktu widzenia – rynek metali. W kontekście towarowym rozważania dotyczą gospodarczego wykorzystania metali. Znajdują one szerokie zastosowanie w przemyśle konstrukcyjnym, motoryzacyjnym, produkcji małego i dużego AGD, przemyśle wojskowym, lotniczym itd. Można także wskazać bardziej subtelny obszar zastosowań – medycynę. Z kolei finansowy charakter rynku metali, powiązany jest z możliwością lokowania środków finansowych w tego rodzaju aktywa, gdyż są notowane na giełdzie (przykładowo London Metal Exchange, LME). Jednakże metale to nie tylko towar „produkcyjny" – metale to także dobro jubilerskie. Inwestorzy kupujący szlachetne kruszce, traktują je jako alternatywę dla klasycznych inwestycji w papiery wartościowe. Ciekawym przykładem jest sytuacja na rynku złota w okresie kryzysu finansowego na świecie w latach 2007-2010. Ceny złota osiągały w tym okresie ponadprzeciętne poziomy, podczas gdy na rynku akcji panowała silna bessa.

Przechodząc do kwestii zastosowania omawianego modelu regresji, zaproponowano jego praktyczną aplikację do oceny ryzyka inwestycji, podejmowanych na rynku metali. Wykorzystano metale zaliczane do grupy nieżelaznych: aluminium, miedź, ołów, nikiel, cynę oraz cynk. Jako benchmark rynkowy wykorzystano indeks LMEX (London Metal Exchange Index). Okresem badawczym są lata 2006-2015. Analizę prowadzono dla dziennych kursów zamknięcia notowań metali oraz indeksu. Dane pochodzą z giełdy londyńskiej LME. Zaproponowano model o strukturze autoregresyjnej pierwszego rzędu następującej postaci:

$$p_t = \beta_0 + \beta_1 p_{t-1} + \beta_2 v_{M,t} + \beta_3 v_{M,t-1} + \varepsilon_t \tag{11}$$

gdzie p_t oznacza cenę metalu w okresie t (lub $t-1$), v_t – wartość benchmarku w okresie t (lub $t-1$), β_i, $(i = 1,2,3)$ oznacza parametry strukturalne modelu, β_0 reprezentuje wyraz wolny, natomiast ε_t – realizację składnika losowego. Model (13) zakłada, że obecny kurs notowań metalu zależy od wartości benchmarku w badanym okresie oraz kursu tego metalu i wartości benchmarku w okresie poprzednim. Na rys. 1 przedstawiono kursy zamknięcia dla metali i benchmarku w całym analizowanym okresie.



Rys. 1. Kursy notowań metali oraz indeksu LMEX w latach 2006-2015

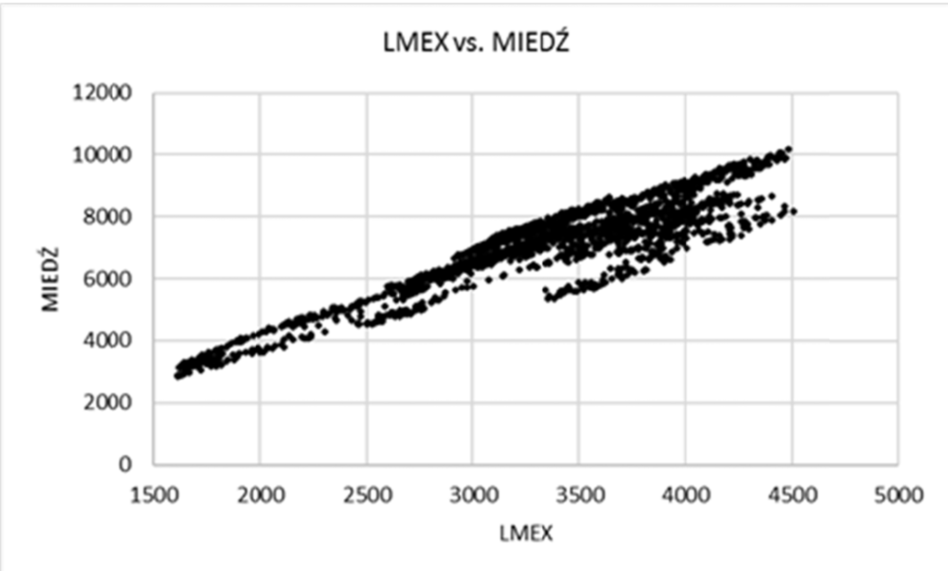
Wykresy sugerują zgodny kierunek zmian kursów notowań poszczególnych metali z kierunkiem zmian benchmarku, jakim jest indeks LMEX. Istnieje zatem przypuszczenie, że występuje dodatnia korelacja pomiędzy tymi wielkościami. Wyniki analizy zależności przedstawiono w tab. 2.

Tabela 2. Współczynniki korelacji dla analizowanych zmiennych

	ALUMINIUM	MIEDŹ	OŁÓW	NIKIEL	CYNA	CYNK	LMEX
ALUMINIUM	1,0000	0,5353	0,2851	0,7628	-0,0138*	0,6786	0,8035
MIEDŹ	0,5353	1,0000	0,6194	0,4187	0,6634	0,3268	0,9017
OŁÓW	0,2851	0,6194	1,0000	0,3656	0,5759	0,1307	0,6017
NIKIEL	0,7628	0,4187	0,3656	1,0000	-0,0683	0,7905	0,7448
CYNA	-0,0138*	0,6634	0,5759	-0,0683	1,0000	-0,3020	0,4179
CYNK	0,6786	0,3268	0,1307	0,7905	-0,3020	1,0000	0,6288
LMEX	0,8035	0,9017	0,6017	0,7448	0,4179	0,6288	1,0000

* współczynnik korelacji statystycznie nieistotny

Wszystkie współczynniki korelacji okazały się statystycznie istotne poza parą aluminium i cyna. Interpretując związek kursów metali z kursem indeksu LMEX, zaobserwowano zależność dodatnią silną bądź umiarkowanie silną. Najsilniejszą korelację przedstawiono na rys. 2.



Rys. 2. Wykres rozrzutu pomiędzy kursem miedzi a kursem LMEX

W kolejnym etapie badania przeprowadzono analizę regresji. Estymowano model regresji wielorakiej, którego parametry strukturalne wyznaczane są zgodnie ze wzorem (6) oraz model regresji grzbietowej, dla którego wektor parametrów estymowany jest zgodnie ze wzorem (7). Dla modelu regresji grzbietowej przyjęto wartości parametru kary λ zgodnie z formułami $\hat{\lambda}_{HKB}$ oraz $\hat{\lambda}_{LW}$ z tab. 1.

Wyniki oszacowania modelu regresji wielorakiej za pomocą estymatora MNK przedstawiono w tab. 3.

Tabela 3. Parametry modelu regresji – estymator MNK

Parametr	**Aluminium**				**Miedź**			
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$
Estymator	-2,1584	0,9953*	0,5076*	-0,5040*	2,4533	0,9966*	2,2034*	-2,1967*
Błąd standardowy	2,4595	0,0017	0,0076	0,0077	5,4850	0,0016	0,0174	0,0177
Parametr	Ołów				Nikiel			
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$
Estymator	0,6969	0,9961*	0,5987*	-0,5964*	24,0203	0,9976*	6,7763*	-6,7691*
Błąd standardowy	4,1881	0,0018	0,0131	0,0132	51,2503	0,0016	0,1509	0,1511
Parametr	Cyna				Cynk			
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$
Estymator	-42,0681	0,9978*	3,8046*	-3,7789*	5,6407	0,9978*	0,7625*	-0,7627*
Błąd standardowy	34,3772	0,0012	0,1072	0,1074	3,7310	0,0013	0,0118	0,0118

* Parametr statystycznie istotny na poziomie < 0,0001

Wyniki estymacji modeli liniowych, opisujących kursy giełdowe analizowanych metali wskazują na statystycznie istotne wartości parametrów strukturalnych dla wszystkich analizowanych zmiennych niezależnych. Ceny obecne wykazują jednokierunkowe zmiany, porównując je z cenami w dniu poprzednim. Podobne wnioski można odnieść w stosunku do relacji z indeksem LMEX, jednak wskazują one na wyższe ryzyko niż w przypadku inwestycji w sam indeks. Kursy metali reagują agresywnie w stosunku do zmian obserwowanych na poziomie benchmarku. Wyjątkiem są inwestycje w aluminium oraz cynk. Biorąc pod uwagę wartości indeksu LMEX z okresu poprzedniego, ceny metali są przeciwnie skorelowane, z równie wysokim poziomem ryzyka. Można zatem wysunąć przypuszczenie, że ryzyko w porównaniu z indeksem w dniu obecnym i poprzednim bilansuje się.

Wyniki oszacowania modelu regresji wielorakiej przy wykorzystaniu estymatora grzbietowego przedstawiono w tab. 4.

Tabela 4. Parametry modelu regresji – estymator grzbietowy

Parametr	**Aluminium**				**Miedź**			
	$\tilde{\beta}_0$	$\tilde{\beta}_1$	$\tilde{\beta}_2$	$\tilde{\beta}_3$	$\tilde{\beta}_0$	$\tilde{\beta}_1$	$\tilde{\beta}_2$	$\tilde{\beta}_3$
1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
Estymator	-2,1634	0,9961*	0,4992*	-0,5011*	2,4534	0,9979*	2,2047*	-2,1896*
Błąd standardowy	2,4571	0,0018	0,0069	0,0072	5,4571	0,0015	0,0170	0,0153
Parametr	Ołów				Nikiel			
	$\tilde{\beta}_0$	$\tilde{\beta}_1$	$\tilde{\beta}_2$	$\tilde{\beta}_3$	$\tilde{\beta}_0$	$\tilde{\beta}_1$	$\tilde{\beta}_2$	$\tilde{\beta}_3$

cd. tabeli 4

1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
Estymator	0,6756	0,9788*	0,6124*	-0,5988*	24,0127	0,9988*	6,7542*	-6,7428*
Błąd standardowy	4,0945	0,0016	0,0101	0,0133	51,8412	0,0016	0,1436	0,1423
Parametr	Cyna				Cynk			
	$\tilde{\beta}_0$	$\tilde{\beta}_1$	$\tilde{\beta}_2$	$\tilde{\beta}_3$	$\tilde{\beta}_0$	$\tilde{\beta}_1$	$\tilde{\beta}_2$	$\tilde{\beta}_3$
Estymator	-42,1347	0,9979*	3,7913*	-3,8217	5,6654	0,9963	0,7698	-0,7543
Błąd standardowy	34,3924	0,0012	0,1127	0,1023	3,7129	0,0012	0,0116	0,0109

* Parametr statystycznie istotny na poziomie < 0,0001

Interpretując wyniki uzyskane za pomocą estymatora grzbietowego, stwierdzono, że, podobnie jak w przypadku estymatora MNK, statystyczną istotność wykazują wszystkie zmienne niezależne. Wyrazy wolne w modelach są statystycznie nieistotne. Porównując błędy standardowe, zauważono nieznaczne ich obniżenie, jednakże nie jest to zmiana znacząca z punktu widzenia walidacji modelu. Wnioski te potwierdzają oszacowania parametru kary dla każdego modelu, które zaprezentowano w tab. 5.

Tabela 5. Oszacowania parametru kary

Parametr	**Aluminium**			**Miedź**		
	$\hat{\lambda}_{HKB}$	$\hat{\lambda}_{LW}$	$\hat{\lambda}_{OPT}$	$\hat{\lambda}_{HKB}$	$\hat{\lambda}_{LW}$	$\hat{\lambda}_{OPT}$
Estymator	0,0012	0,0024	0,0020	0,0004	0,0012	0,0010
Parametr	Ołów			Nikiel		
	$\hat{\lambda}_{HKB}$	$\hat{\lambda}_{LW}$	$\hat{\lambda}_{OPT}$	$\hat{\lambda}_{HKB}$	$\hat{\lambda}_{LW}$	$\hat{\lambda}_{OPT}$
Estymator	0,0025	0,0048	0,0050	0,0018	0,0028	0,0050
Parametr	Cyna			Cynk		
	$\hat{\lambda}_{HKB}$	$\hat{\lambda}_{LW}$	$\hat{\lambda}_{OPT}$	$\hat{\lambda}_{HKB}$	$\hat{\lambda}_{LW}$	$\hat{\lambda}_{OPT}$
Estymator	0,0020	0,0027	0,0080	0,0012	0,0024	0,0030

* Optymalna wartość parametru kary w sensie metody uogólnionej walidacji krzyżowej (ang. *Generalized Cross-Validation*, GCV)

Oszacowane parametry kary dla każdego z modeli, opisujących ceny kursów metali przyjmują wartości bliskie zeru. Wnioskować można zbieżność ocen estymatora grzbietowego w ocenami uzyskanymi za pomocą estymatora MNK. Wynik jest zgodny z teorią, która mówi, że dla $\hat{\lambda} \to 0$ estymator grzbietowy jest zbieżny z estymatorem MNK.

Podsumowanie

W pracy podjęto próbę zastosowania modelu regresji grzbietowej do oceny ryzyka inwestycyjnego na rynku metali. Modele regresyjne znajdują szerokie zastosowanie w opisie zjawisk gospodarczych, bez względu na rynek, na którym są stosowane. Istotna jest taka konstrukcja modelu, by realnie opisywał rzeczywistość. Teoria ekonometrii i statystyki przedkłada pewne założenia formalne, jakie powinien spełnić model, jednakże w praktyce ich restrykcje są zazwyczaj łagodzone.

W artykule zaprezentowano model regresji grzbietowej. Jego zastosowanie jest zasadne w przypadku modelowania zjawisk ekonomicznych, cechujących się silną współliniowością oraz autokorelacją składnika losowego. Porównując oceny estymatorów grzbietowych z estymatorami MNK, można wskazać na poprawę jakości modelu oraz zmniejszanie średnich błędów szacunku poszczególnych parametrów strukturalnych. Regresja grzbietowa wykorzystuje pewien parametr kary, który umożliwia spełnienie zadość założeniu odwracalności macierzy $X^T X$, która w praktyce często jest osobliwa. W pracy przywołano wybrane metody szacowania parametru kary, część z nich zastosowano w praktyce.

Analiza empiryczna wykorzystania modelu regresji grzbietowej została zaprezentowana na przykładzie inwestycji, podejmowanych na rynku metali. Założono, że kurs określonego metalu w dniu obecnym zależny jest od kursu benchmarku w tym dniu oraz dodatkowo od kursu metalu i kursu benchmarku w dniu poprzednim. Wyniki porównawcze dla estymatorów MNK oraz grzbietowych, wykazały zbieżność ocen poszczególnych parametrów strukturalnych modelu. Przesłanką ku temu są także oceny parametru kary, zbliżone do wartości zero. Wskazano na silny wpływ wielkości indeksu na obserwowane kursy metali. Ceny reagują agresywnie w tym samym kierunku, co benchmark dla okresu badanego oraz w przeciwnym kierunku niż benchmark dla okresu poprzedniego.

Biorąc pod uwagę przesłanki zastosowania modeli regresji grzbietowej w kontekście zmniejszenia średnich błędów szacunku parametrów strukturalnych zaobserwowano, iż w przypadku prezentowanych modeli ta poprawa jest znikoma. Bardzo możliwe jest uzyskanie odmiennych wyników w przypadku analizy bazującej na stopach zwrotu, co jest przedmiotem dalszych badań.

Literatura

Hoerl A.E., Kennard R.W. (1970), *Ridge Regression: Biased Estimation for Nonorthogonal Problems*, „Technometrics", Vol. 12, No. 1, s. 55-67.

Hoerl A.E., Kennard R.W., Baldwin K.F. (1975), *Ridge Regression: Some Simulations*, "Communications in Statistics – Simulation and Computation", Vol. 4, No. 2, s. 105-123.

Jajuga K. (2009), *Zarządzanie ryzykiem*, Wydawnictwo Naukowe PWN, Warszawa.

Kibria B.M.G. (2003), *Performance of Some New Ridge Regression Estimators*, „Communication in Statistics – Simulation and Computation", Vol. 32, No. 2, s. 419-435.

Lawless J.F., Wang P. (1976), *A Simulation Study of Ridge and Other Regression Estimators*, "Communications in Statistics – Theory and Methods", Vol. 5, No. 4, s. 307-323.

Maddala G.S. (2006), *Ekonometria*, Wydawnictwo Naukowe PWN, Warszawa.

Ross, S.A. (1976), *The Arbitrage Theory of Capital Asset Pricing*, „Journal of Economic Theory", Vol. 13, s. 341-360.

Welfe A. (2003), *Ekonometria*, Polskie Wydawnictwo Ekonomiczne, Warszawa.

RIDGE REGRESSION MODEL AND ITS APPLICATION IN INVESTMENT RISK ASSESSMENT – THE CASE OF METALS MARKET

Summary: Regression models are commonly used statistical tool (and other quantitative sciences), which allows for modelling relations within analyzed datasets. One of their most important features is the prediction property. In the literature there are a lot of methods for estimating unknown parameters of regression models, e.g. Maximum Likelihood, Ordinary Least Squares etc., but estimated parameters not always meet the required assumptions (regarding their properties itself and the properties of selected model). Co-linearity between data prevents from correct inferring using classical models. The aim of the article is the application of ridge regression model in risk assessment on the metals market. The model uses so-called penalty parameter allowing for elimination of co-linear variables and makes the model more simpler. Additionally, the bias of the estimators and their variances reduce.

Keywords: regression model, ridge regression, metals market, co-linearity, risk.