

Daniel Papla, Krzysztof Piontek

Akademia Ekonomiczna we Wrocławiu

WYKORZYSTANIE MODELU DCC-GARCH ORAZ FUNKCJI POWIĄZAŃ FGM W ANALIZIE ZALEŻNOŚCI MIĘDZY STOPAMI ZWROTU Z AKCJI, INDEKSÓW GIEŁD ŚWIATOWYCH ORAZ KURSAMI WALUT. PRZYKŁAD EMPIRYCZNY

1. Wstęp

Celem artykułu jest analiza oszacowań parametrów funkcji powiązań Farliego-Gumbela-Morgensterna (FGM) dla danych, które wcześniej poddano filtracji za pomocą modelu dynamicznej warunkowej korelacji (DCC – *dynamic conditional correlation*) Engle’a [3]. Oszacowanie takie może być podstawą do wyciągnięcia wniosków o naturze zależności między analizowanymi szeregami. Ze względu na czasochłonność i złożoność obliczeń ograniczono się jedynie do wybranych spółek z polskiej giełdy, kilku indeksów i kilku walut. Z tego samego powodu w obliczeniach wykorzystano jedynie portfele składające się z trzech elementów.

Jedną z cech funkcji powiązań, którą niektórzy autorzy uznają za szczególnie przydatną w modelowaniu zależności między finansowymi szeregami czasowymi, jest ich nieliniowość [5]. Estymatory współczynników funkcji powiązań można traktować jako swego rodzaju nieliniowe rozwinięcie współczynnika korelacji liniowej. W artykule weryfikuje się dla wybranych szeregów hipotezę badawczą, którą można sformułować w następujący sposób: zależność występująca w finansowych szeregach czasowych ma charakter nieliniowy. Jej potwierdzeniem będzie statystyczna istotność współczynników funkcji powiązań po wyeliminowaniu korelacji liniowej. Aby zweryfikować tę hipotezę, autorzy wykorzystali model dynamicznej korelacji DCC-GARCH Engle’a i funkcję powiązań FGM. Analiza składała się z dwóch etapów. W pierwszym dla każdego z trójelementowych portfeli (od-

powiednio akcji, indeksów lub walut) estymowano model DCC. W drugim etapie dla reszt modelu DCC estymowano współczynniki funkcji powiązań FGM. Istotne różne od zera współczynniki tej funkcji mogą potwierdzać hipotezę badawczą, ponieważ wskazują one na istnienie w resztach zależności, które nie udało się wy-modelować za pomocą DCC.

Praca składa się z trzech głównych części. W pierwszej omówiono zastosowany do modelowania korelacji model DCC Engle'a oraz przedstawiono funkcje powiązań (*copula*), jako przykład podając wielowymiarową funkcję FGM. Druga część zawiera omówienie metodologii wybranej do badań empirycznych oraz danych wykorzystanych w badaniach. W ostatniej części przedstawiono wyniki badań wraz z ich interpretacją.

2. Modelowanie zależności

Punktem wyjścia do dalszych rozważań będą pojęcia warunkowych wartości oczekiwanych (μ_t), warunkowej macierzy wariancji-kowariancji (H_t) wyznaczanych na podstawie informacji dostępnej w chwili $t - 1$ oraz pojęcie postaci warunkowego rozkładu standaryzowanych reszt modelu [2; 6; 10].

Aby zilustrować DCC, przedstawmy k -wymiarowy wektor r_t obserwacji w momencie t w równaniu warunkowej wartości oczekiwanej jako zredukowany model wektorowej autoregresji (VAR):

$$A(L)r_t = \varepsilon_t, \text{ gdzie } \varepsilon_t \sim N(0, H_t), \quad (1)$$

gdzie: $A(L)$ jest macierzą wielomianową operatora opóźnienia L , ε_t to wektor reszt modelu VAR z warunkową macierzą wariancji-kowariancji

$$H_t = \{h_{ij}\}_t = \begin{bmatrix} h_{11} & h_{12} & \dots & h_{1k} \\ h_{21} & h_{22} & \dots & h_{2k} \\ \dots & \dots & \dots & \dots \\ h_{k1} & h_{k2} & \dots & h_{kk} \end{bmatrix}_t, i, j = 1, 2, \dots, k.$$

Zapiszmy warunkową macierz wariancji-kowariancji w postaci:

$$H_t = D_t R_t D_t, \quad (2)$$

gdzie: $D_t = \text{diag} \{ \sqrt{h_{ij}} \}_t$ jest diagonalną macierzą zmiennych w czasie odchylen standardowych jednowymiarowych procesów GARCH poszczególnych składników portfela, $R_t = \{\rho_{ij}\}_t$, $i, j = 1, 2, \dots, k$ to macierz korelacji zawierająca warunkowe współczynniki korelacji pomiędzy składnikami portfela. Elementy macierzy D_t można opisać za pomocą jednowymiarowych modeli GARCH (p, q) w następujący sposób:

$$h_{it} = \omega_i + \sum_{p=1}^{P_i} \alpha_{ip} \varepsilon_{it-p}^2 + \sum_{q=1}^{Q_i} \beta_{iq} h_{it-q}, \quad (3)$$

gdzie: $h_{it} = \{h_{it}\}_t$, czyli uproszczony zapis elementów diagonalnych macierzy $\mathbf{D}_t$.

Model DCC(m, n) Engle'a [2002] opisuje strukturę macierzy $\mathbf{R}_t$ w następujący sposób

$$\mathbf{R}_t = \mathbf{Q}_t^{*-1} \mathbf{Q}_t \mathbf{Q}_t^{*-1}, \quad (4)$$

gdzie:

$$\mathbf{Q}_t = \left(1 - \sum_{m=1}^M a_m - \sum_{n=1}^N b_n\right) \bar{\mathbf{Q}} + \sum_{m=1}^M a_m (\xi_{t-m} \xi'_{t-m}) \sum_{n=1}^N b_n \mathbf{Q}_{t-n}, \quad (5)$$

$\xi_{it} = \frac{\varepsilon_{it}}{\sqrt{h_{it}}}$ – elementy wektora standaryzowanych reszt, $\mathbf{Q}_t = \{q_{ij}\}_t$ – macierz

wariancji-kowariancji standaryzowanych reszt, której bezwarunkowa (niezależna od czasu) macierz wariancji-kowariancji $\bar{\mathbf{Q}}$ otrzymywana jest w pierwszym kroku estymacji. $\mathbf{Q}_t^*$ to diagonalna macierz zawierająca pierwiastki kwadratowe elementów leżących na przekątnej macierzy $\mathbf{Q}_t$:

$$\mathbf{Q}_t^* = \begin{bmatrix} \sqrt{q_{11}} & 0 & 0 \\ 0 & \sqrt{q_{22}} & 0 \\ \dots & \dots & \dots \\ 0 & 0 & \dots \sqrt{q_{kk}} \end{bmatrix}_t. \quad (6)$$

Model DCC-GARCH można estymować za pomocą metody największej wiarygodności, w której logarytm funkcji wiarygodności można opisać następującym wzorem:

$$L = -\frac{1}{2} \sum_{t=1}^T \{2 \log(2\pi) + 2 \log|\mathbf{D}_t| + \log|\mathbf{R}_t| + \xi'_t \mathbf{R}_t^{-1} \xi_t\}. \quad (7)$$

Dla naszych rozważań znaczenie ma to, że proces estymacji można przeprowadzić dwuetapowo [3]. W pierwszym etapie macierz $\mathbf{R}_t$ jest zastępowana przez k -wymiarową macierz jednostkową, co redukuje równanie (7) do sumy logarytmów funkcji wiarygodności jednowymiarowych modeli GARCH. W drugim etapie estymuje się parametry modelu DCC w równaniu (5), wykorzystując pełny logarytm funkcji wiarygodności (7) przy warunku wykorzystania parametrów jednowymiarowych modeli GARCH otrzymanych w pierwszym etapie.

Innym sposobem modelowania zależności jest zastosowanie funkcji powiązań. Definicja funkcji powiązań wygląda następująco [5]): k -wymiarową funkcję $C: [0, 1]^k \rightarrow [0, 1]$ nazywamy *copula*, jeśli spełnia następujące warunki:

- $C(u_1, u_1, \dots, u_k)$ jest funkcją k -rosnącą,
- $C(u_1, u_2, \dots, u_{i-1}, 0, u_{i+1}, \dots, u_k) = 0$,
- $C(1, \dots, 1, u_i, 1, \dots, 1) = u_i$.

Znaczenie funkcji powiązań w analizie zależności wielowymiarowych wynika z twierdzenia Sklara:

Niech H będzie dystrybuantą łączną rozkładu wielowymiarowego, którego rozkłady brzegowe oznaczmy odpowiednio przez F_i . Wtedy istnieje funkcja powiązań C taka, że

$$H(x_1, x_2, \dots, x_k) = C(F_1(x_1), F_2(x_2), \dots, F_k(x_k)). \quad (8)$$

Jeśli F_i są ciągłe, to C jest jednoznacznie określona. Ponadto, jeśli F_i są dystrybuantami, to funkcja H określona równaniem (8) jest dystrybuantą łączną rozkładu wielowymiarowego.

Dowód dwuwymiarowego przypadku tego twierdzenia można znaleźć w pracy Sklara z 1959 r. Schweizer i Sklar w pracy z 1974 r. przedstawili nieco prostszą postać. Dowody na przypadek dwu- i wielowymiarowy można znaleźć również w pracy Nelsena [5]. O tym, jak sobie radzić w przypadku, kiedy dystrybuanty F_i nie są ciągłe, traktuje m.in. praca Joego [4].

Jeśli oznaczmy $u_i = F_i(x_i)$, to z twierdzenia Sklara wypływa następujący wniosek:

$$C(u_1, u_2, \dots, u_k) = H(F_1^{-1}(u_1), F_2^{-1}(u_2), \dots, F_k^{-1}(u_k)). \quad (9)$$

Z twierdzenia tego wynika, że funkcją powiązań możemy przybliżać nieznany wielowymiarowy rozkład łączny. Oczywiście problemem jest znalezienie odpowiednio dobrze dopasowanej funkcji.

O ile w przypadku dwuwymiarowym w literaturze występuje bardzo wiele funkcji powiązań (por. np. spis zamieszczony w pracy [5]), o tyle w przypadku, gdy mamy do czynienia z więcej niż dwoma wymiarami, takich funkcji jest znacznie mniej. Bardzo popularna w analizie dwuwymiarowej rodzina funkcji archimedesowskich [5], choć możliwa do uogólnienia na przypadek wielowymiarowy, jest rzadko stosowana, ponieważ stosując ją, należy założyć, że powiązanie każdej z par zmiennych jest takie samo, co, oczywiście, z reguły odbiega od rzeczywistości [5].

Takiej wady nie posiada funkcja powiązań Farliego-Gumbela-Morgensterna, którą w przypadku wielowymiarowym można przedstawić w następujący sposób:

$$C(u_1, u_2, \dots, u_k) = u_1 u_2 \dots u_k P(u_1, u_2, \dots, u_k), \quad (10)$$

gdzie P to wielomian o następującej postaci:

$$P(u_1, u_2, \dots, u_k) = 1 + \sum_{i=2}^k \sum_{1 \leq j_1 < \dots < j_{i-1} \leq k} \theta_{j_1 j_2 \dots j_{i-1}} (1 - u_{j_1})(1 - u_{j_2}) \dots (1 - u_{j_{i-1}}). \quad (11)$$

Przykładowo dla przypadku, gdy $k = 3$, wielomian P funkcji powiązań FGM ma następującą postać:

$$P(u_1, u_2, u_3) = 1 + \theta_{12}(1 - u_1)(1 - u_2) + \theta_{13}(1 - u_1)(1 - u_3) + \theta_{23}(1 - u_2)(1 - u_3) + \theta_{123}(1 - u_1)(1 - u_2)(1 - u_3). \quad (12)$$

Wykorzystując wielomian P , bardzo łatwo również można przedstawić funkcję gęstości funkcji powiązań FGM:

$$c(u_1, u_2, \dots, u_k) = P(2u_1, 2u_2, \dots, 2u_k), \quad (13)$$

Dla trójwymiarowej FGM wzór ten przybiera postać:

$$c(u_1, u_2, u_3) = 1 + \theta_{12}(1 - 2u_1)(1 - 2u_2) + \theta_{13}(1 - 2u_1)(1 - 2u_3) + \theta_{23}(1 - 2u_2)(1 - 2u_3) + \theta_{123}(1 - 2u_1)(1 - 2u_2)(1 - 2u_3). \quad (14)$$

Estymacja parametrów takiej funkcji przebiega w dwóch etapach [1; 5]:

1. W pierwszym etapie estymuje się parametry poszczególnych rozkładów brzegowych. Te parametry służą następnie do wyznaczenia wartości dystrybuant dla obserwacji z każdego rozkładu brzegowego. W badaniach empirycznych w niniejszym artykule jako rozkłady brzegowe wykorzystano jednowymiarowe rozkłady normalne o parametrach estymowanych osobno dla poszczególnych składników.

$$u_{it} = \Phi(x_{it}),$$

gdzie: u_{it} – wartość dystrybuanty szeregu i w momencie t ,

x_{it} – wartość obserwacji i w momencie t ,

Φ – dystrybuanta rozkładu normalnego o parametrach estymowanych dla szeregu i .

2. Wyznaczone w ten sposób wartości dystrybuant wykorzystuje się następnie w estymacji metodą największej wiarygodności parametrów funkcji powiązań.

3. Dane i algorytm badań empirycznych

Dane podzielono na trzy grupy: akcje notowane na GPW w Warszawie, indeksy (jeden polski, pozostałe zagraniczne) i waluty. Do grupy akcji wybrano (losowo) 12 następujących spółek BZWBK, Dębica, Forte, Jelfa, Kęty, Mostostal Export, Optimus, PGF, Provimrol, Softbank, Swarzędz i Wistula. Z kolei w skład indeksów wchodziły: WIG, CAC40 (Paryż), DAX (Frankfurt), FT-SE100 (Londyn), DJIA, NASDAQ, SP500 (USA). Waluty obejmowały notowania w złotych następujących walorów: frank szwajcarski (CHF), euro (EUR), funt brytyjski (GBP), jen (JPY) i dolar (USD). Ostatnim dniem, z którego pochodzą analizowane dane, jest w przypadku akcji 2 lutego 2006 r., a w przypadku indeksów i walut – 3 marca 2006 r. Dane w każdej grupie posłużyły do utworzenia wszystkich możliwych kombinacji składających się z trzech elementów z danej grupy. Dla akcji takich trójek (portfeli) było 220, dla indeksów – 35, a dla walut – jedynie 10.

Wszystkie prezentowane pracy wyniki badań empirycznych uzyskano na podstawie procedur napisanych w środowisku MATLAB 6.0.

Algorytm zastosowany w badaniach empirycznych składał się z kilku etapów:

1. Dla każdej trójki szeregów czasowych estymowano model VAR (1) z opóźnieniem równym 1.

2. Dla reszty modelu VAR otrzymanego w poprzednim etapie estymowano model DCC(1, 1) GARCH .
3. Dla standaryzowanych reszt modelu DCC otrzymanego w poprzednim etapie estymowano parametry funkcji powiązań FGM. W estymacji metodą największej wiarygodności wykorzystano funkcję gęstości opisaną wzorem (12).

4. Wyniki i ich interpretacja

Po estymacji parametrów funkcji powiązań (trzeci etap algorytmu przedstawionego w poprzednim punkcie) otrzymano wyniki, które w formie zbiorczej przedstawiono w tab. 1.

Tabela 1. Liczba istotnych parametrów funkcji FGM w poszczególnych grupach

Parametry \ Grupa	Akcje (liczba portfeli 220)		Indeksy (liczba portfeli 35)		Waluty (liczba portfeli 10)	
θ_{12}	22	10,00%	2	5,71%	0	0,0%
θ_{13}	16	7,27%	1	2,86%	2	20,0%
θ_{23}	23	10,45%	0	0,00%	0	0,0%
θ_{123}	13	5,91%	5	14,29%	1	10,0%

Źródło: obliczenia własne.

Jeśli założymy, że DCC GARCH prawidłowo modeluje korelację pomiędzy składnikami portfela oraz że korelacja ta oddaje w pełni związki pomiędzy tymi składnikami, to możemy oczekiwać, że estymowane parametry funkcji FGM będą nieistotnie różne od zera, zwłaszcza pierwsze trzy parametry. Czwarty parametr oddaje niejako związek łączny pomiędzy wszystkimi trzema składnikami, czego DCC nie modeluje. Niektórzy autorzy określają czwarty parametr mianem statystyki trzypunktowej, a pierwsze trzy statystykami dwupunktowymi (podobnie jak współczynniki korelacji liniowej) [1].

Analizując wyniki w tab. 1, można uznać, że DCC na tyle dobrze modeluje związki pomiędzy zmiennymi, że jedynie w ok. 10% przypadków wartości parametrów funkcji powiązań FGM zdaje się wskazywać, że pomiędzy szeregami istnieją związki, których DCC nie wymodelował.

Ciekawe jest to, że ostatni parametr nie był częściej istotny niż pozostałe, przynajmniej w przypadku akcji i walut. W przypadku indeksów ten parametr okazywał się najczęściej istotny, należy jednak też zauważyć, że dla tej grupy parametry FGM najrzadziej były istotnie różne od zera, czyli w przypadku tej grupy model DCC niejako „spisywał się” najlepiej.

W tabeli 2 przedstawiono estymatory parametrów funkcji FGM wybranych portfeli akcji. Są to portfele, dla których 3 parametry okazały się istotne (2 przypadki) oraz 2 wybrane portfele, dla których 2 parametry były istotne (13 przypadków).

Tabela 2. Estymatory parametrów funkcji FGM i ich istotność

Spółki Parametry	Provimrol Optimus Dębica		Softbank Optimus Forte		Optimus Dębica BZWBK		Optimus Jelfa Dębica	
	Wartość	Stat. <i>t</i>	Wartość	Stat. <i>t</i>	Wartość	Stat. <i>t</i>	Wartość	Stat. <i>t</i>
θ_{12}	0,1947	2,8081	0,2075	2,1503	-0,0409	-0,6478	0,0528	0,7954
θ_{13}	0,1135	1,6961	0,0245	0,2934	0,2284	3,2597	0,1646	2,3324
θ_{23}	0,2302	3,1700	0,3648	4,4121	0,1900	2,7269	0,0443	0,6159
θ_{123}	-0,3706	-3,0121	0,3815	2,5570	-0,1359	-1,1481	-0,2922	-2,3501

Źródło: obliczenia własne.

To, że w tab. 2 w każdym przypadku jednym ze składników jest spółka Optimus, nie jest przypadkiem. Okazuje się, że spółka ta wykazuje silne powiązania, mierzone wartością parametrów funkcji FGM z większością pozostałych spółek, po usunięciu korelacji liniowej za pomocą modelu DCC. Oczywiście taki wynik może wynikać z ograniczonej próby poddanej badaniom.

5. Zakończenie

Wyniki badań wskazują, że analiza związków za pomocą funkcji powiązań, które mogą modelować inne rodzaje zależności niż korelacja liniowa, wydaje się potrzebna jedynie w ok. 10% przypadków, jednakże zdaniem autorów liczba ta jest na tyle duża, że usprawiedliwia dalsze badania.

Oczywiście wyniki uzyskane jedynie na wybranych danych nie są rozstrzygające. Również zastosowanie tylko jednego wielowymiarowego modelu GARCH oraz jednej funkcji powiązań sprawia, że przedstawione wyniki są raczej poglądowe. Autorzy mieli na celu wykazanie istnienia tego typu zagadnień, a nie ich ostateczne rozstrzygnięcie.

Dalsze kierunki badań to przede wszystkim zastosowanie z jednej strony innych modeli korelacji liniowej (inne wielowymiarowe modele GARCH) oraz zastosowanie innych niż funkcja powiązań FGM, metod analizy zależności nieliniowych. Chodzi tu przede wszystkim o mierniki zależności w ogonach oraz statystyki wielopunktowe oddające równoczesne zależności między więcej niż dwiema zmiennymi [1].

Literatura

- [1] Armstrong M., Galli A., *Sequential Nongaussian Simulations Using the FGM Copula*, Working Paper 2002.
- [2] Bollerslev T., Engle R., Nelson D., *ARCH Models*, [w:] *Handbook of Econometrics*, Vol. IV, Amsterdam 1994, s. 2959-3038.

- [3] Engle R.F., *Dynamic Conditional Correlation – A Simple Class of Multivariate GARCH Models*, „Journal of Business and Economic Statistics” 2002, 20(3), s. 339-50.
- [4] Joe H., *Multivariate Models and Dependence Concepts*, Chapman & Hall, Boca Raton 1997.
- [5] Nelsen R.B., *An Introduction to Copulas*, Springer Verlag, New York 1999.
- [6] Piontek K., *Modelowanie i prognozowanie zmienności instrumentów finansowych*, praca doktorska, AE, Wrocław 2002.
- [7] Piontek K., *Prognozowanie macierzy kowariancji i korelacji finansowych szeregów czasowych*, praca złożona w ramach konferencji „Modelowanie preferencji a ryzyko”, Akademia Ekonomiczna w Katowicach, Ustroń 2005.
- [8] Schweizer B., Sklar A., *Operations on Distributions Functions not Derivable from Operations on Random Variables*, „Studia Mathematica” 1974, 52, s. 43-52.
- [9] Sklar A., *Fonctions de répartition á n dimensions et leurs marges*, „Publications de l’Institut Statistique de l’Université de Paris” 1959, 8, s. 229-231.
- [10] Tsay R., *Analysis of Financial Time Series*, Wiley and Sons, 2002.

USING A DCC-GARCH MODEL AND A FGM COPULA IN THE ANALYSIS OF THE DEPENDENCY BETWEEN RETURNS OF STOCKS, WORLD STOCK MARKETS INDEXES AND EXCHANGE RATES. EMPIRICAL EXAMPLE

Summary

Main goal of this paper is an analysis of the estimation of parameters of Farlie-Gumbel-Morgenstern copula (FGM) for data filtered with dynamic conditional correlation (DCC) of Engle [3]. Such an estimation can help establishing a nature of dependency between time series.

The paper consist of three main parts. The first part presents DCC model used for modeling a linear correlation. Also copula functions are presented here with example – multidimensional FGM copula. Data and methodology are described in second part, results – in third.