

Daniel Papla

Akademia Ekonomiczna we Wrocławiu

KLASYFIKACJA SPÓŁEK NOTOWANYCH NA GPW W WARSZAWIE Z WYKORZYSTANIEM FUNKCJI POWIĄZAŃ (*COPULA FUNCTIONS*)

1. Wstęp

W artykule tym zaprezentowano próbę klasyfikacji spółek notowanych na GPW w Warszawie na podstawie zależności pomiędzy ich stopami zwrotu. Wykorzystano tu związki między zależnością a odległością, omówione m.in. w pracy [Jajuga, Walesiak 2004]. Zależność między dwiema spółkami może być mierzona tradycyjnie, współczynnikiem korelacji liniowej pomiędzy notowaniami tych spółek, lub za pomocą wybranej funkcji powiązań. W opracowaniu wykorzystano jedną z archimedesowskich funkcji powiązań, których parametr theta można uznać za miarę powiązania między notowaniami, analogiczny do współczynnika korelacji liniowej.

Artykuł składa się z trzech części. Pierwsza zawiera omówienie pojęć i metod wykorzystanych w badaniach empirycznych. W drugiej części porównano portfele zbudowane na podstawie klasyfikacji spółek z GPW w Warszawie, wykorzystującej funkcje powiązań, z portfelami zbudowanymi na podstawie klasyfikacji wykorzystującej współczynnik korelacji liniowej. Wyniki badań empirycznych omówiono w ostatniej części artykułu.

2. Metodologia badań empirycznych

Zastosowanie funkcji powiązań pozwala częściowo rozwiązać najpoważniejszy problem związany z wykorzystaniem rozkładów wielowymiarowych, a mianowicie

nieznajomość postaci analitycznej empirycznego łącznego rozkładu stóp zwrotu. Definicja funkcji powiązań (w przypadku dwuwymiarowym, który łatwo uogólnić na przypadek wielowymiarowy) wygląda następująco [Nelsen 1999]:

Dwumiarową funkcję $C: [0, 1]^2 \rightarrow [0, 1]$ nazywamy *copula*, jeśli spełnia następujące warunki:

- a) $C(u, v)$ jest rosnąca i ze względu na u , i ze względu na v ,
- b) $C(u, 0) = C(0, v) = 0$, $C(u, 1) = u$, $C(1, v) = v$,
- c) $\forall u_1, u_2, v_1, v_2 \in [0, 1]$ takich, że $u_1 < u_2$ i $v_1 < v_2$, mamy $C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0$.

Jeśli przyjmiemy, że u i v są wartościami dystrybucyjnymi brzegowej rozkładu np. stóp zwrotu, wtedy własność (a) oznacza, że jeśli jedna z dystrybucyjnych brzegowych jest stała, wtedy wzrost wartości drugiej dystrybucyjnej powoduje wzrost wartości funkcji powiązań. Własność druga upodabnia funkcję powiązań jeszcze silniej do dystrybucyjnej łącznej rozkładów. Wynika z niej, iż jeśli prawdopodobieństwo zajścia jednego ze zdarzeń brzegowych jest równe zero, wtedy wartość funkcji powiązań jest równa zero, jeśli zaś jedno ze zdarzeń marginalnych jest zdarzeniem pewnym, wtedy wartość funkcji powiązań jest równa prawdopodobieństwu zajścia zdarzenia drugiego. Własność trzecia mówi nam, iż jeśli wartości u i v rosną równocześnie, wtedy wartość funkcji powiązań również rośnie.

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

Niech H będzie dystrybucyjną łączną rozkładu dwuwymiarowego, którego rozkłady brzegowe oznaczmy odpowiednio przez F i G . Wtedy istnieje funkcja powiązań C taka, że

$$H(x, y) = C(F(x), G(y)). \quad (1)$$

Jeśli F i G są ciągłe, wtedy C jest jednoznacznie określona. Ponadto, jeśli F i G są dystrybucyjnymi, wtedy funkcja H określona równaniem (1) jest dystrybucyjną łączną rozkładu dwuwymiarowego.

Dowód twierdzenia można znaleźć w pracy Sklara z 1959 r. Nieco prostszą postać Schweizer i Sklar przedstawili w pracy z 1974 r. Z twierdzenia tego wynika, że możemy przybliżać nieznaną dwuwymiarową rozkład łączny funkcją powiązań. Oczywiście problemem jest znalezienie odpowiednio dobrze dopasowanej funkcji. Częściowym rozwiązaniem tego problemu jest dopasowanie do danych empirycznych kilku rodzajów funkcji powiązań i wybranie do dalszej analizy funkcji najlepiej dopasowanej.

Jeden z pełniejszych przeglądów funkcji powiązań można znaleźć w pracy Nelsena (1999). Szczególnym przypadkiem funkcji powiązań są funkcje archimedejskie, które można przedstawić za pomocą ogólnego wzoru:

$$C(u_1, u_2) = \psi^{-1}(\psi(u_1) + \psi(u_2)), \quad (2)$$

$$\psi: [0, 1] \rightarrow [0, \infty), \quad \psi(1) = 0.$$

W badaniach empirycznych, których wyniki zamieszczono w niniejszym artykule wykorzystano jedną z funkcji archimedesowskich, funkcję Franka:

$$C(u, v) = -\frac{1}{\theta} \ln \left(1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1} \right), \quad \theta \in (-\infty, \infty). \quad (3)$$

Współczynnik θ mierzy siłę powiązania między rozkładami brzegowymi. Im jest on większy, tym to powiązanie jest silniejsze.

Wartość współczynnika θ wykorzystano w grupowaniu za pomocą metody aglomeracyjnej (najdalszego sąsiada i Warda) jako miarę podobieństwa, zakładając, że notowania dwóch spółek są tym bardziej do siebie podobne, im wartość współczynnika θ , estymowana na podstawie rozkładów brzegowych ich stóp zwrotu, jest większa.

Jako miarę podobieństwa dla porównania zastosowano również uogólnioną miarę odległości Walesiaka (*general distance measure* – GDM), w zmodyfikowanej formie, wykorzystującej współczynniki korelacji liniowej, a zaprezentowanej w pracy [Jajuga, Walesiak 2004]:

$$d_{jk} = \frac{1}{2} - \frac{-4 + m(r_{jk} + 1) - \sum_{l=1, l \neq j, k}^m r_{jl} - \sum_{l=1, l \neq j, k}^m r_{kl}}{4 \cdot \left[(m - \sum_{l=1}^m r_{jl}) \cdot (m - \sum_{l=1}^m r_{kl}) \right]^{0.5}}, \quad (4)$$

gdzie: m – liczba zmiennych uwzględnionych w analizie,

r_{ij} – współczynnik korelacji liniowej między zmiennymi i i j .

GDM w zaprezentowanej formie posiada kilka charakterystycznych cech [Jajuga, Walesiak 2004]:

- wartość GDM zależy od liczby badanych zmiennych (w tym przypadku jako zmienne traktuje się notowania poszczególnych spółek),
- odległość między dwiema spółkami zależy nie tylko od korelacji między tymi spółkami, ale również od korelacji między pozostałymi spółkami.

Obliczenia, których wyniki zamieszczono w dalszej części artykułu, przeprowadzono według następującego schematu:

- estymowano współczynnik θ , wykorzystując 500 (ok. 2 lata) ostatnich stóp zwrotu dla każdej pary spółek, tworząc macierz 49×49 podobieństw między spółkami,
- wykorzystując tę macierz, dzielono spółki na 10 lub 5 grup, metodą aglomeracji według najdalszego sąsiada i metodą Warda [Anderberg 1973],
- z każdej grupy wybierano jedną spółkę o największej (najmniejszej) średniej stopie zwrotu w ciągu ostatnich 500 notowań,

- z wybranych spółek tworzone portfele 10- lub 5- elementowe o równych udziałach, które kupowano po cenach z następnego notowania i sprzedawano po upływie 20 sesji (ok. miesiąca),
- po sprzedaży obliczano 20-sesyjną stopę zwrotu.

Całą procedurę powtórzono 29 razy, przesuwając się za każdym razem o 20 sesji do przodu.

Dla porównania akcje grupowano, wykorzystując GDM obliczoną na podstawie współczynników korelacji dla ostatnich 500 sesji i postępując tak, jak przy podobieństwie mierzonym θ .

Jako przybliżenie rozkładów brzegowych wykorzystano dystrybuantę empiryczną, rozkład łączny modelowano za pomocą funkcji powiązań Franka. Wybierano akcje o maksymalnej lub o minimalnej stopie zwrotu w celu weryfikacji hipotezy o kontynuacji lub o zmianie trendu.

3. Dane

Do obliczeń wykorzystano dzienne notowania i logarytmiczne stopy zwrotu z okresu od 3 stycznia 2000 r. do 8 września 2005 r. (ok. 1080 obserwacji), dla 49 wybranych spółek z GPW w Warszawie. Kryteriami wyboru były:

- data pierwszego notowania – spółki wybrane do badań powinny być notowane od początku wziętego pod uwagę okresu, czyli co najmniej od 3 stycznia 2000 r.,
- jak najmniejsza liczba brakujących notowań.

Drugie kryterium wymaga nieco szerszego omówienia. W przypadku analizy zależności pomiędzy jednowymiarowymi szeregami czasowymi ważne jest, aby te szeregi były równej długości, a poszczególne obserwacje pochodziły z tych samych okresów. Niestety, jeśli w jednym z szeregów brakuje obserwacji z danego okresu (np. akcje spółki Apator nie były notowane 12 stycznia 2000 r.), okres ten należy usunąć ze wszystkich pozostałych szeregów, aby zachować zgodność danych. Oczywiście zmniejsza to długość szeregów wykorzystanych w obliczeniach, co negatywnie wpływa na jakość estymacji parametrów modeli. Autor wybrał 49 spółek tak, aby zminimalizować utratę danych, przy zachowaniu jak największej liczby spółek wykorzystanych w obliczeniach.

4. Wyniki badań empirycznych

W tabelach 1-2 przedstawiono 20-sesyjne, procentowe stopy zwrotu osiągane przez analizowane portfele.

Tabela 1. Stopy zwrotu z portfeli, w których kryterium wyboru była maksymalna stopy zwrotu z wykorzystaniem metody aglomeracji według najdalszego sąsiada (NS) i metody Warda (w %)

Miara podobieństwa Data	θ Franka dla portfeli				GDM dla portfeli			
	10-elementowe		5-elementowe		10-elementowe		5-elementowe	
	NS	Ward	NS	Ward	NS	Ward	NS	Ward
XI 2002	4,37	0,44	9,08	6,55	6,93	0,11	1,62	9,76
XII 2002	13,31	17,31	9,82	14,05	6,01	16,40	8,12	13,93
II 2003	6,66	4,72	9,77	7,54	9,36	9,20	17,72	7,44
IV 2003	21,30	15,04	17,81	16,86	12,79	16,50	16,21	12,88
V 2003	12,16	15,76	15,25	20,50	11,63	13,98	11,37	19,94
VI 2003	20,83	24,69	17,16	12,67	21,25	16,63	20,82	16,68
VII 2003	24,01	28,71	27,69	23,40	24,33	28,05	18,65	29,49
VIII 2003	-0,70	-0,48	-4,54	0,50	0,24	-1,74	1,19	2,26
IX 2003	13,26	18,33	21,85	23,79	12,58	11,09	6,02	16,68
X 2003	-3,95	-5,61	-3,69	-3,98	-4,32	-4,04	-4,04	-3,69
XII 2003	15,27	15,03	23,05	14,52	10,07	14,82	15,89	24,08
I 2004	30,29	30,95	34,87	36,77	17,96	24,39	20,55	23,44
II 2004	9,38	10,33	12,60	10,31	8,86	8,32	9,35	6,92
III 2004	9,31	5,76	-0,25	11,71	9,68	6,03	0,67	10,72
IV 2004	11,00	12,52	19,00	20,54	10,82	9,20	11,30	14,79
VI 2004	2,88	4,69	1,95	3,17	1,04	1,46	4,86	-0,36
VII 2004	6,00	9,02	11,91	9,31	6,28	9,78	9,07	7,00
VIII 2004	2,59	1,44	-1,29	4,18	3,20	2,40	2,50	0,52
IX 2004	0,65	0,74	-0,40	2,68	-1,36	-1,92	-2,75	-0,51
X 2004	-2,03	-1,97	-2,99	-1,67	-0,85	1,86	-0,79	1,12
XI 2004	-4,53	-2,70	-6,06	-3,09	-3,78	1,08	-3,54	-2,78
XII 2004	-3,78	-5,26	-3,45	-4,31	-4,31	-3,74	-7,29	-5,34
II 2005	1,37	4,22	-2,20	-0,79	-0,59	2,07	-2,45	2,24
III 2005	-0,42	-9,71	-5,05	-3,27	-2,21	0,30	0,37	-4,12
IV 2005	-2,97	-4,83	-2,80	-3,59	-2,67	-7,60	-2,27	-6,20
V 2005	-2,07	-1,15	-2,70	0,39	-3,68	-3,71	-5,64	-5,13
VI 2005	0,32	-0,57	5,50	-9,36	-4,65	-1,37	-14,28	3,52
VII 2005	4,71	2,01	2,10	-1,28	4,22	5,99	2,98	1,16
VIII 2005	-3,32	-0,47	-3,90	5,45	-1,55	0,33	-1,67	-3,37

Źródło: obliczenia własne.

Tabela 2. Stopy zwrotu z portfeli, w których kryterium wyboru była minimalna stopy zwrotu z wykorzystaniem metody aglomeracji według najdalszego sąsiada (NS) i metody Warda (w %)

Miara podobieństwa Data	θ Franka dla portfeli				GDM dla portfeli			
	10-elementowe		5-elementowe		10-elementowe		5-elementowe	
	NS	Ward	NS	Ward	NS	Ward	NS	Ward
I	2	3	4	5	6	7	8	9
XI 2002	-5,67	-3,15	-7,12	-4,74	-7,81	-5,24	-10,30	-6,69
XII 2002	17,05	23,58	40,50	15,73	19,68	20,34	40,46	37,81
II 2003	1,16	2,73	13,59	3,69	9,19	2,75	6,81	9,04
IV 2003	11,01	5,63	-3,62	7,68	10,31	18,51	8,53	20,78
V 2003	-0,70	0,29	-4,59	0,59	2,89	4,46	-1,25	3,25

I	2	3	4	5	6	7	8	9
VI 2003	20,89	28,09	12,23	34,62	16,67	16,57	11,20	13,94
VII 2003	35,16	28,06	40,47	30,45	26,90	21,24	30,55	21,52
VIII 2003	-4,53	-2,10	-8,47	-1,64	-5,67	-3,85	-5,17	-2,50
IX 2003	-1,34	2,05	-8,68	-4,44	-2,13	-2,70	-15,69	-1,87
X 2003	-4,65	-4,78	3,35	-2,46	-7,12	-6,78	-3,16	-7,01
XII 2003	-5,67	-3,15	-7,12	-4,74	-7,81	-5,24	-10,30	-6,69
I 2004	9,99	9,42	16,02	8,46	11,63	15,71	10,86	15,83
II 2004	29,78	29,11	45,38	36,42	27,91	32,19	31,54	42,90
III 2004	23,79	20,82	31,13	14,44	22,20	18,73	22,91	22,71
IV 2004	19,31	15,82	20,11	26,20	12,30	11,41	9,74	24,08
VI 2004	4,11	4,31	8,51	4,09	8,68	2,01	1,08	15,04
VII 2004	0,18	1,43	-5,26	-8,15	-2,50	-3,63	-10,78	-5,93
VIII 2004	0,59	4,27	-1,26	-2,05	-2,96	-3,70	-5,15	-1,44
IX 2004	-6,59	-2,68	-7,07	-4,82	-5,39	-5,74	-4,92	-3,68
X 2004	7,20	8,82	2,16	4,31	4,29	3,36	-0,74	2,60
XI 2004	-4,62	-7,06	-8,87	-13,09	-2,15	-1,30	-1,86	-2,37
XII 2004	-1,11	-1,89	-2,61	-1,58	-2,90	-3,03	-4,52	-4,03
II 2005	0,31	4,54	-0,39	-1,29	4,67	8,18	4,65	4,17
III 2005	-5,05	-10,26	-9,22	-15,85	-5,13	-5,31	1,47	-5,14
IV 2005	-7,03	-5,97	-10,38	-6,89	-5,07	-2,94	-9,56	-8,10
V 2005	-5,82	-1,85	-4,83	-1,52	-0,46	4,38	-0,45	-1,68
VI 2005	7,23	10,07	8,63	8,94	8,10	9,26	8,07	10,16
VII 2005	4,36	-1,05	-1,38	2,87	7,06	4,13	4,77	-0,53
VIII 2005	-0,09	2,33	0,81	3,03	-3,52	-1,27	-0,91	1,81

Źródło: obliczenia własne.

Pogrubiono stopy zwrotu tych portfeli, które osiągnęły najlepsze wyniki w danym okresie. Okazuje się, że w przypadku portfeli utworzonych ze spółek o maksymalnej stopie zwrotu w poprzednim okresie lepsze wyniki wydają się osiągać portfele tworzone z wykorzystaniem współczynnika θ funkcji *copula* Franka niż portfele wykorzystujące GDM. W przypadku portfeli utworzonych ze spółek o minimalnej stopie zwrotu ta zależność raczej nie występuje. Potwierdzają to wyniki zamieszczone w tab. 3, gdzie przedstawiono średnie arytmetyczne i odchylenia standardowe stóp zwrotu osiąganych przez portfele tworzone z wykorzystaniem badanych metod.

Tabela 3. Podsumowanie wyników

Portfele dla maksymalnej stopy zwrotu								
Miara podobieństwa	θ Franka				GDM			
Liczba elementów	10-elementowe		5-elementowe		10-elementowe		5-elementowe	
Metoda aglomeracji	NS	Ward	NS	Ward	NS	Ward	NS	Ward
Średnia	6,41%	6,51%	6,90%	7,36%	5,08%	6,06%	4,64%	6,66%
Odch. std.	9,33%	10,58%	11,31%	10,76%	8,04%	8,87%	9,10%	9,96%
Portfele dla minimalnej stopy zwrotu								
Średnia	5,02%	5,52%	5,39%	4,30%	4,76%	5,03%	3,98%	6,66%
Odch. std.	11,45%	11,02%	16,08%	13,40%	10,34%	10,27%	13,20%	13,55%

Źródło: obliczenia własne.

Jak wynika z tab. 3, wykorzystanie funkcji powiązań pozwala osiągać lepsze wyniki, niestety kosztem większego ryzyka. Dla porównania w tab. 4 przedstawiono średnią 20-sesyjną stopę zwrotu z indeksów giełdowych i z 5-elementowych portfeli utworzonych ze spółek o największej (kolumna Maks.) lub o najmniejszej (kolumna Min.) stopie zwrotu osiągniętej w przeszłości.

Tabela 4. Stopy zwrotu indeksów i portfeli niezdywersyfikowanych (w %)

	WIG	WIG20	Maksimum	Minimum
Średnia	2,97	2,55	5,76	3,99
Odchylenie standardowe	5,18	5,61	9,68	12,22

Źródło: obliczenia własne.

Z tabeli 4 wynika, że zastosowanie metod zaprezentowanych w niniejszym artykule pozwala na osiągnięcie prawie trzykrotnego wzrostu stóp zwrotu w porównaniu z rynkiem (porównując stopę zwrotu z indeksu WIG20 i stopę zwrotu z 5-elementowych portfeli utworzonych z wykorzystaniem *copula* Franka i metody Warda), przy dwukrotnym wzroście ryzyka. Ponieważ wyniki osiągnięte przez portfele zbudowane bez wykorzystania klasyfikacji są gorsze od wyników portfeli ją wykorzystujących, można uznać, że dobre wyniki z tab. 3 nie są wynikiem jedynie wyboru spółek najlepszych w przeszłości.

W tabelach 5 i 6 zamieszczono średnie stopy zwrotu z portfeli otrzymane przy założeniu reinwestowania środków.

Tabela 5. Podsumowanie wyników

Portfele dla maksymalnej stopy zwrotu								
Miara podobieństwa	θ Franka				GDM			
Liczba elementów	10-elementowe		5-elementowe		10-elementowe		5-elementowe	
Metoda aglomeracji	NS	Ward	NS	Ward	NS	Ward	NS	Ward
Stopa zwrotu	15,40%	15,37%	17,08%	20,19%	9,94%	13,84%	8,11%	16,40%
Portfele dla minimalnej stopy zwrotu								
Stopa zwrotu	8,76%	10,74%	8,34%	5,97%	8,22%	9,13%	5,24%	14,66%

Źródło: obliczenia własne.

Tabela 6. Stopy zwrotu indeksów i portfeli niezdywersyfikowanych (w %)

	WIG	WIG20	Maksimum	Minimum
Stopa zwrotu	4,34	3,41	12,17	5,48

Źródło: obliczenia własne.

Oczywiście stopy zwrotu są większe, jednakże ogólny schemat, tzn. przewaga portfeli wykorzystujących klasyfikację z wykorzystaniem *copula* Franka, jest nadal zachowany.

5. Wnioski

Prezentowane wyniki pozwalają wyciągnąć następujące wnioski:

1. Stopy zwrotu dla portfeli, dla których podobieństwo mierzono funkcją powiązań, są nieco wyższe niż stopy zwrotu dla portfeli otrzymanych za pomocą GDM.
2. Ta wyższa stopa zwrotu wiąże się z wyższym ryzykiem.
3. Obydwie metody pozwalają uzyskać o wiele wyższe stopy niż stopy rynkowe, mierzone za pomocą stóp zwrotu z indeksów. Okupione to jest wyższym ryzykiem.
4. Portfele 5-elementowe są bardziej dochodowe niż portfele 10-elementowe.
5. Wyniki dla całego badanego okresu zdają się bardziej przemawiać na korzyść hipotezy o kontynuacji trendu: portfele, w których skład wchodzi spółki o najwyższej stopie zwrotu w danej grupie dają wyższe stopy zwrotu niż portfele utworzone ze spółek o najniższych stopach zwrotu w przeszłości.
6. Jeśli przyjrzymy się dokładniej danym z poszczególnych okresów, widać, że na początku i pod koniec badanego okresu portfele utworzone ze spółek o najniższych stopach zwrotu dają lepsze wyniki, co z kolei przemawiałoby za hipotezą o zmianie kierunku trendu.

Wydaje się, że funkcja powiązań Franka lepiej oddaje związek pomiędzy notowaniami akcji wykorzystanych do budowy portfeli niż klasyczny współczynnik korelacji wykorzystany w GDM, pozwalając osiągnąć wyższe stopy zwrotu. Oczywiście wymagane są dalsze badania, jednakże zdaniem autora za metodą wykorzystywaną w artykule przemawia względna prostota (np. ograniczenie się jedynie do dwuwymiarowych funkcji *copula*) i wykorzystanie jedynie najłatwiej dostępnych danych, czyli notowań giełdowych (a nie danych finansowych, które są czasami trudniejsze do zdobycia i obarczone pewną niejednoznacznością).

Literatura

- Anderberg M.R. (1973), *Cluster Analysis for Applications*, Academic Press, New York.
- Jajuga K., Walesiak M. (2004), *Remarks on the Dependence Measures and the Distance Measures*, *Klasyfikacja i analiza danych – teoria i zastosowania*, Taksonomia 11, Prace Naukowe Akademii Ekonomicznej we Wrocławiu nr 1022, AE, Wrocław.
- Nelsen R.B. (1999), *An Introduction to Copulas*, Springer Verlag, New York.
- Schweizer B., Sklar A. (1974), *Operations on Distributions Functions not Derivable from Operations on Random Variables*, „*Studia Mathematica*” nr 52, s. 43-52.
- Sklar A. (1959), *Fonctions de répartition à n dimensions et leurs marges*, Publications de l’Institut Statistique de l’Université de Paris 8, s. 229-231.

CLASSIFICATION OF THE COMPANIES LISTED ON THE WSE USING A *COPULA FUNCTIONS*

Summary

This paper presents an attempt of classification of the companies listed on the WSE based on dependency between their returns. The relationship between dependency and distance is used here. This relationship is described in [Jajuga, Waleśiak 2004]. Dependency between two companies could be measured in the traditional way, using linear correlation coefficient between returns, or using chosen copula function. In this paper one of the Archimedean copula function was used. The parameter θ of those function can be treated as a dependency measure, analogous to the linear correlation coefficient. The paper consist of three parts. In first part the methodology used in empirical example is presented. Comparison between portfolios based on classification of companies from WSE using copula function and portfolios based on classification using the linear correlation coefficient. Empirical example is discussed in the third part.