

Ekonofizyka: Macierz kowariancji o długich ogonach

Andrzej Görlich

Instytut Fizyki, Uniwersytet Jagielloński



Kraków, 20 kwietnia 2006

Giełda jako układ statystyczny

- Giełda jest złożonym i skomplikowanym systemem.
- Pomijamy dokładne mechanizmy, które nią rządzą.
- Przyjmujemy, że giełdę opisuje układ statystyczny o wielu stopniach swobody.
- Każda notowana spółka odpowiada jednemu stopniowi swobody.

Giełda jako układ statystyczny

- Notowania giełdowe możemy zapisać w formie macierzy.
- Element W_{it} oznacza cenę akcji i -tej spółki w czasie $t = n\Delta t$.

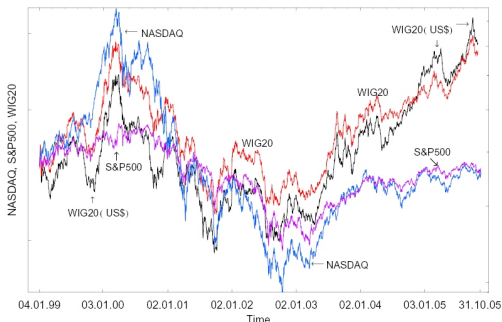
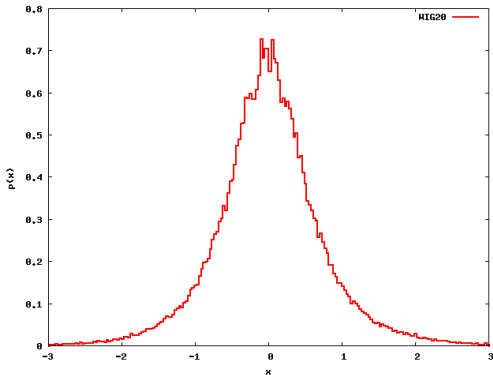


Fig. 1. NASDAQ, S&P500 and WIG20(Warsaw Stock Exchange) index from 1999.01.04 until 2005.10.31.

(R. Rak, S. Drożdż, J. Kwapien, physics/0603071)

Giełda jako układ statystyczny - macierz zwrotów

- Ważne są względne zmiany cen.
- Wygodniej rozpatrywać zwrot ceny akcji: $R_{it} = \text{Ln} \frac{W_{it}}{W_{it-\Delta t}}$.
- Wprowadzamy znormalizowany zwrot: $x_{it} = \frac{R_{it} - \langle R_i \rangle}{\sigma_i}$.
- Wartość średnia x_{it} wynosi zero: $\langle x_{it} \rangle = 0$
- Wariancja x_{it} wynosi jeden: $\langle x_{it}^2 \rangle = 1$



Rozkład zwrotu x dla WIG20 w latach 2000 - 2005 z krokiem czasowym $\Delta t = 1 \text{ min}$.

Korelacje a optymalny portfel

Notowania spółek nie są niezależne, lecz występują między nimi korelacje.

$$\delta_{ij} \neq \langle x_i x_j \rangle_t$$

Korelacje a optymalny portfel

Notowania spółek nie są niezależne, lecz występują między nimi korelacje.

Możemy zdefiniować macierz korelacji:

$$C_{ij} = \langle x_i x_j \rangle_t$$

Korelacje a optymalny portfel

Notowania spółek nie są niezależne, lecz występują między nimi korelacje.

Możemy zdefiniować macierz korelacji:

$$C_{ij} = \langle x_i x_j \rangle_t$$

- Dobór optymalnego portfela polega na maksymalizacji zysku (średni zwrot) przy jednoczesnym minimalizowaniu ryzyka (odchylenie standardowe).

Korelacje a optymalny portfel

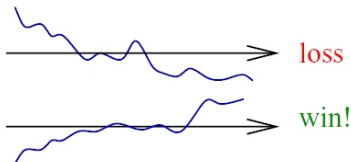
Notowania spółek nie są niezależne, lecz występują między nimi korelacje.

Możemy zdefiniować macierz korelacji:

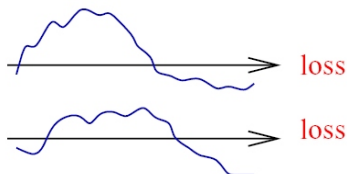
$$C_{ij} = \langle x_i x_j \rangle_t$$

- Dobór optymalnego portfela polega na maksymalizacji zysku (średni zwrot) przy jednoczesnym minimalizowaniu ryzyka (odchylenie standardowe).
- Markowitz: Inwestuj w niezależne spółki

uncorrelated



correlated



- Wektory własne macierzy korelacji odpowiadają nieskorelowanym superpozycjom spółek
- Znajomość korelacji między spółkami pozwala ocenić ryzyko i dokonać wyboru optymalnego portfela
- Interesuje nas widmo macierzy \mathbf{C}
- Jak znaleźć widmo macierzy korelacji?

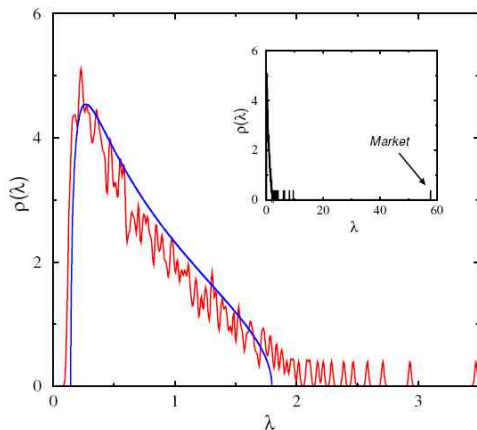
- Nie znamy prawdziwej macierzy korelacji \mathbf{C}
- Mając wyniki historyczne - macierz \mathbf{x} - możemy przybliżyć \mathbf{C} jej estymatorem
- Estymator macierzy korelacji nazywamy macierzą kowariancji \mathbf{c} :

$$\mathbf{c} = \frac{1}{T} \mathbf{x} \mathbf{x}^T = \frac{1}{T} \boxed{\mathbf{x}} \boxed{\mathbf{x}^T} \Leftrightarrow c_{ij} = \frac{1}{T} \sum_{t=1}^T x_{it} x_{jt}$$

- Dla krótkich czasów pomiaru T macierz \mathbf{c} jest zaszumiona, co utrudnia wydobycie prawdziwych korelacji.
- Dla $T \rightarrow \infty : \mathbf{c} \rightarrow \mathbf{C}$
- Gdy T maleje widmo bc rozptywa się

Czyszczenie widma macierzy kowariancji

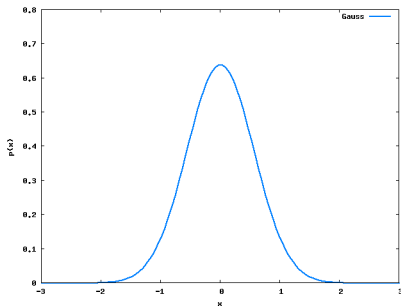
- Widmo macierzy kowariancji jest zaszumione
- Czyszczenie, polega na wyodrębnieniu z widma prawdziwych korelacji. Odpowiadają im wartości własne z górnej części widma



Losowy szum tworzy główną część widma. Niebieską linią oznaczono widmo dla czysto przypadkowych fluktuacji gaussowskich.

- Model gaussowski

$$P(\mathbf{x}) \propto e^{-\frac{1}{2} \text{Tr} \mathbf{x}^T \mathbf{C}^{-1} \mathbf{x}}$$



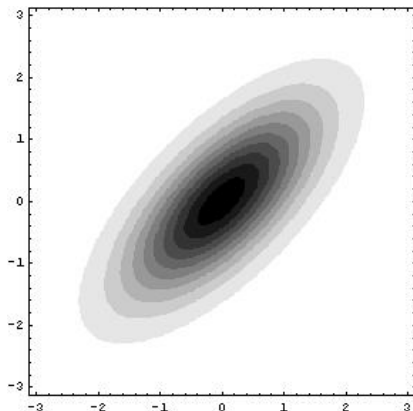
Fluktuacje gaussowskie

- Model gaussowski

$$P(\mathbf{x}) \propto e^{-\frac{1}{2} \text{Tr} \mathbf{x}^T \mathbf{C}^{-1} \mathbf{x}}$$

- Dopuszcza korelacje

$$C_{ij} = \langle x_i x_j \rangle_t$$



Fluktuacje gaussowskie

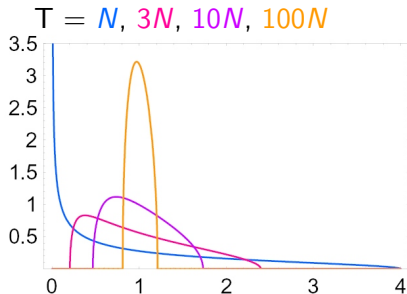
- Model gaussowski

$$P(\mathbf{x}) \propto e^{-\frac{1}{2}\text{Tr}\mathbf{x}^T\mathbf{C}^{-1}\mathbf{x}}$$

- Dopuszcza korelacje

$$C_{ij} = \langle x_i x_j \rangle_t$$

- Znamy widmo macierzy kowariancji,



Z. Burda, A. Görlich, A. Jarosz, J. Jurkiewicz, "Signal and noise in correlation matrix"

- Model gaussowski

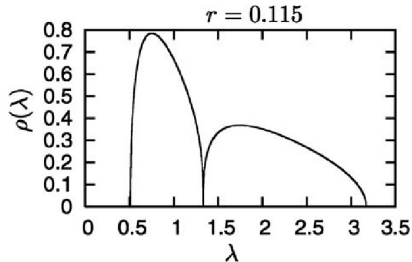
$$P(\mathbf{x}) \propto e^{-\frac{1}{2}\text{Tr}\mathbf{x}^T\mathbf{C}^{-1}\mathbf{x}}$$

- Dopuszcza korelacje

$$C_{ij} = \langle x_i x_j \rangle_t$$

- Znamy widmo macierzy kowariancji, również gdy występują korelacje

Wyprowadzenie korzysta z funkcji Greena i grafów Feynmana



Z. Burda, A. Görlich, A. Jaroš, J. Jurkiewicz, "Signal and noise in correlation matrix"

Fluktuacje gaussowskie

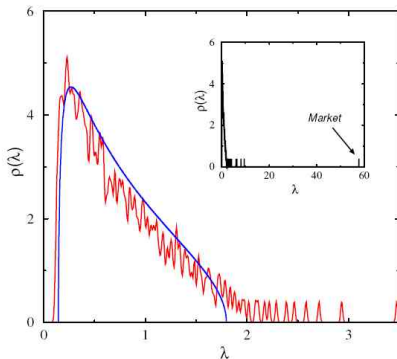
- Model gaussowski

$$P(\mathbf{x}) \propto e^{-\frac{1}{2}\text{Tr}\mathbf{x}^T\mathbf{C}^{-1}\mathbf{x}}$$

- Dopuszcza korelacje

$$C_{ij} = \langle x_i x_j \rangle_t$$

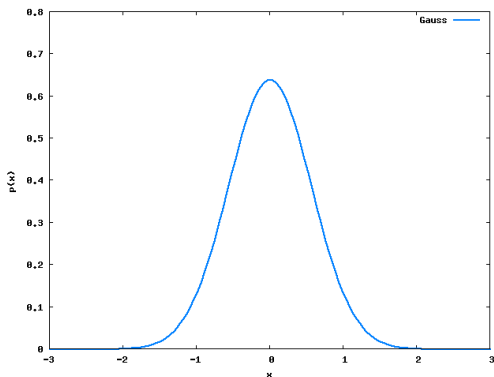
- Znamy widmo macierzy kowariancji, również gdy występują korelacje
- Możemy je zastosować do *oczyszczenia* widma z szumu



J.-P. Bouchaud, P. Cizeau, L. Laloux, M. Potters,
1999-2005

- Rozkład brzegowy jest rozkładem normalnym

$$P(x_i) \propto e^{-\frac{1}{2C_{ii}}x_i^2}$$

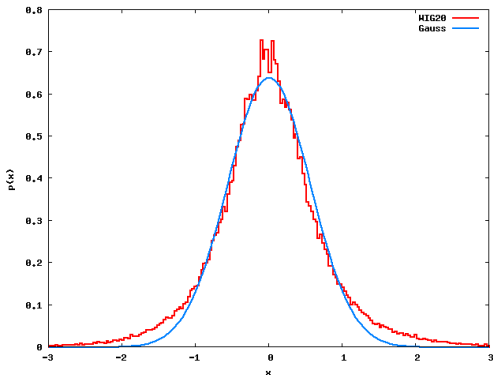


Fluktuacje gaussowskie

- Rozkład brzegowy jest rozkładem normalnym

$$P(x_i) \propto e^{-\frac{1}{2C_{ii}}x_i^2}$$

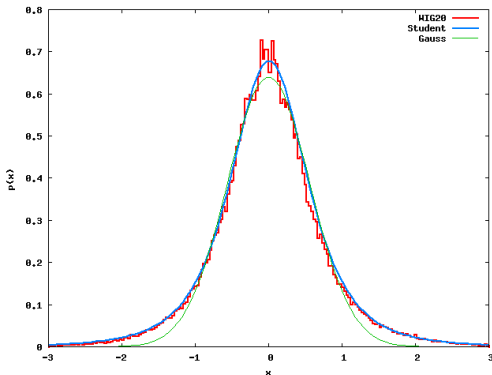
- Niestety rynek nie jest opisywany fluktuacjami gaussowskimi



Fluktuacje niegaussowskie

Rozwiązaniem są rozkłady o długich ogonach, np. rozkład Studenta.

$$P(x_i) \propto \left(1 + \frac{x_i^2}{\nu C_{ii}}\right)^{-\frac{\nu+1}{2}}$$



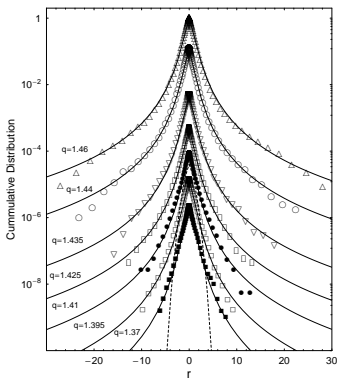
Rozkład zanika potęgowo dla dużych fluktuacji

$$P(x_i) \propto x_i^{-1-\nu}$$

Fluktuacje niegaussowskie

Rozwiązaniem są rozkłady o długich ogonach, np. rozkład Studenta.

$$P(x_i) \propto \left(1 + \frac{x_i^2}{\nu C_{ii}} \right)^{-\frac{\nu+1}{2}}$$



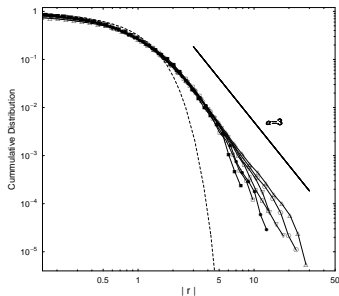
(R. Rak, S. Drozd, J. Kwapien, physics/0603071)

Wykres przedstawia $\ln P(x)$ dla WIG20 i różnych częstotliwości próbkowania.

Fluktuacje niegaussowskie

Rozwiązaniem są rozkłady o długich ogonach, np. rozkład Studenta.

$$P(x_i) \propto \left(1 + \frac{x_i^2}{\nu C_{ii}}\right)^{-\frac{\nu+1}{2}}$$



Wykres przedstawia $\ln P(x)$ od $\ln|x|$ dla WIG20.
Dopasowany wykładnik wynosi $\nu = 3$.

(R. Rak, S. Drozd, J. Kwapien, physics/0603071)

- Aby uwzględnić korelacje między spółkami musimy rozważyć wielowymiarowy skorelowany rozkład Studenta

$$P_\nu(\mathbf{x}) \propto \left(1 + \frac{1}{\nu^2} \text{Tr } \mathbf{x}^T \mathbf{C}^{-1} \mathbf{x} \right)^{-\frac{\nu+d}{2}}, d = N \cdot T$$

- Model ten wymusza korelacje

$$C_{ij} = \langle x_i x_j \rangle_t$$

- Rozkład brzegowy jest rozkładem Studenta

$$P(x_i) \propto \left(1 + \frac{x_i^2}{\nu C_{ii}} \right)^{-\frac{\nu+1}{2}}$$

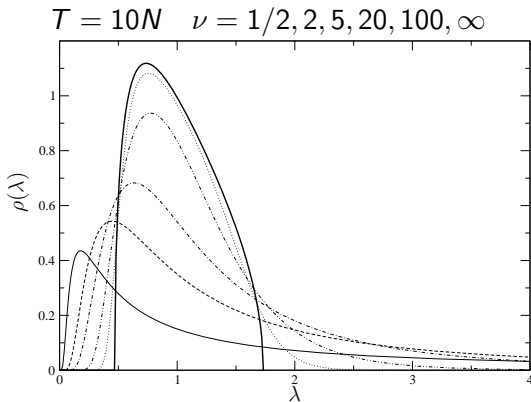
- Teraz znamy także widmo estymatora macierzy korelacji

$$\rho_\nu(\lambda) = \frac{1}{\Gamma(\frac{\nu}{2})} \left(\frac{\nu}{2}\right)^{\nu/2} \lambda^{-\frac{\nu}{2}-1} \int_0^\infty \rho_G(y) e^{-\frac{\nu y}{2\lambda}} y^{\frac{\nu}{2}} dy.$$

- $\rho_\nu(\lambda)$ wyraża się przez widmo dla rozkładu gaussowskiego $\rho_G(y)$.
- Widmo zanika potęgowo dla dużych λ , $\rho_\nu(\lambda) \propto \lambda^{-\frac{\nu}{2}-1}$.
- $\rho_\nu(\lambda)$ jest niezerowe dla $\lambda > 0$.
- W granicy $\nu \rightarrow \infty$ przechodzi w $\rho_G(\lambda)$.

Widmo macierzy kowariancji c dla $C_{ij} = \delta_{ij}$

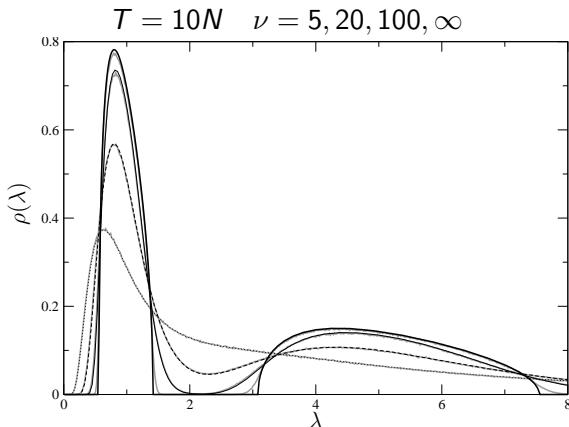
$$\rho_\nu(\lambda) = \frac{1}{\Gamma(\frac{\nu}{2})} \left(\frac{\nu}{2}\right)^{\nu/2} \lambda^{-\frac{\nu}{2}-1} \int_0^\infty \rho_G(y) e^{-\frac{\nu y}{2\lambda}} y^{\frac{\nu}{2}} dy$$



Z. Burda, A. Görlich, B. Waclaw, "Spectral properties of empirical covariance matrices for data with power-law tails", physics/0603186

Widmo macierzy kowariancji \mathbf{c} dla $\mathbf{C} = \text{Diag}(1, 5)$

$$\rho_\nu(\lambda) = \frac{1}{\Gamma(\frac{\nu}{2})} \left(\frac{\nu}{2}\right)^{\nu/2} \lambda^{-\frac{\nu}{2}-1} \int_0^\infty \rho_G(y) e^{-\frac{\nu y}{2\lambda}} y^{\frac{\nu}{2}} dy$$



Z. Burda, A. Görlich, B. Waclaw, "Spectral properties of empirical covariance matrices for data with power-law tails", physics/0603186

- Macierz korelacji pełni ważną rolę w ocenie ryzyka i tworzeniu portfela
- Eksperymentalna macierz kowariancji pozwala znaleźć oryginalną macierz korelacji
- Znamy widmo macierzy kowariancji dla fluktuacji gaussowskich
- Rzeczywisty rynek lepiej opisuje model z fluktuacjami o długich ogonach
- Teraz możemy wyznaczyć widmo macierzy kowariancji dla skorelowanych i długozasięgowych fluktuacji (rozkład Studenta)

Dziękuję za współpracę i udzieloną pomoc

Z. Burda, J. Jurkiewicz, M. A. Nowak i B. Wacław

- Z. Burda, A. Görlich, B. Waćław, "Spectral properties of empirical covariance matrices for data with power-law tails", physics/0603186
- Z. Burda, A. Görlich, A. Jarosz, J. Jurkiewicz, "Signal and noise in correlation matrix", Physica **A343** (2004) 295-310, [cond-mat/0305627]
- R. Rak, S. Drożdż, J. Kwapien, "Nonextensive statistical features of the Polish stock market fluctuations", physics/0603071
- L. Laloux, P. Cizeau, J.-P. Bouchaud and M. Potters, Phys. Rev. Lett. **83** (1999) 1467.
- H. Markowitz, *Portfolio Selection: Efficient Diversification of Investments*, J. Wiley and Sons, Nowy Jork 1959.