

株価の動的相関の推定と変化点検出・ ネットワーク分析への応用

磯貝 孝

首都大学東京 大学院経営学専攻 特任教授、
日本銀行金融機構局

統数研共同研究集会「経済物理学とその周辺」
— H28 年度第一回研究会 —

概要

- ① はじめに
- ② 資産相関とボラティリティ
- ③ 動的相関の推定
- ④ 動的相関行列の固有値分解
- ⑤ 動的ネットワークとしての分析・観察
- ⑥ まとめ

— 本日の発表の詳細については、以下の文献をご参照ください

Isogai, T., "Building a dynamic correlation network for fat-tailed financial asset returns", Applied Network Science 1(7):1-24, 2016, <http://link.springer.com/article/10.1007/s41109-016-0008-x>

リサーチの背景と課題

背景: 資産収益率の動的な相関とは?

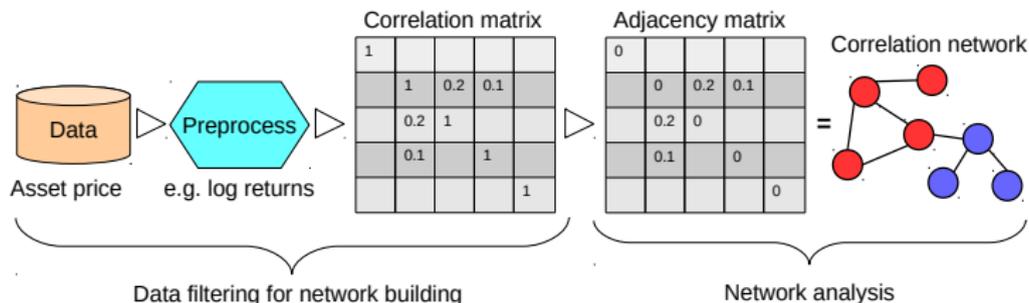
- 二つの資産間のペアワイズの相関からなる相関行列 (静的)
 - ▶ 例: 株価や為替レート
- 相関行列の動的な拡張
 - ▶ 相関はダイナミックに変化している?
- 動的な相関の分析は、**ポートフォリオのリスク管理**や**市場モニタリング**に有益

課題: 動的な相関行列の推定およびその観察に関する手法の提案

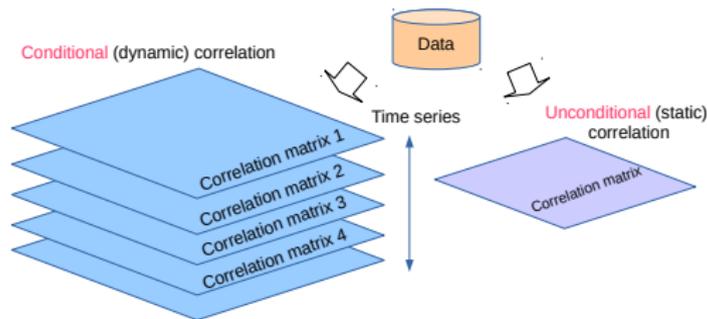
- ファットテール性に起因する**偽相関 (spurious correlation)** 問題への対処
- **モデルベース**の動的相関行列の推定
- 推定した相関行列の変化を観察し、相関の強さの変化や異例な動きを発見する (**変化点の検出**)
- **動的ネットワーク**としての特徴変化の観察

資産収益率データから相関行列を計算する

Static な分析の流れ-単一の相関（隣接）行列:



Dynamic なモデルでは、複数の相関（隣接）行列を推定する:

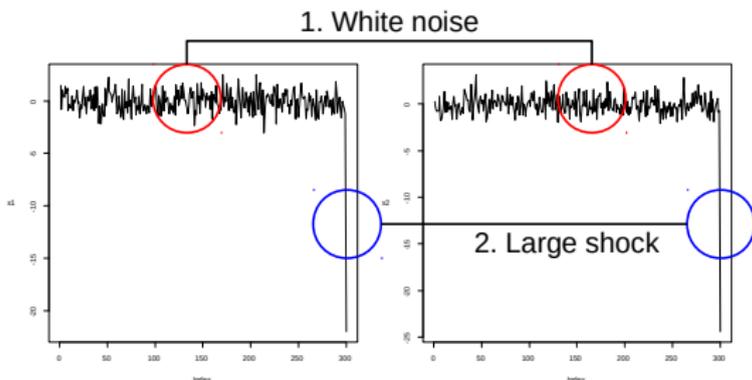


- ① はじめに
- ② **資産相関とボラティリティ**
- ③ 動的相関の推定
- ④ 動的相関行列の固有値分解
- ⑤ 動的ネットワークとしての分析・観察
- ⑥ まとめ

ボラティリティ変動から生じる線形相関の歪み

線形相関:

$$\rho_{X,Y} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \text{Var}(Y)}}$$

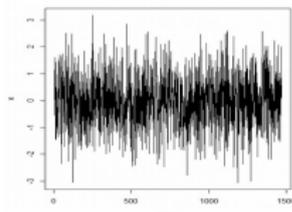
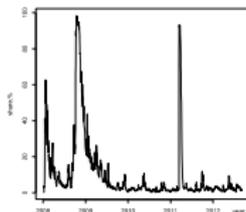
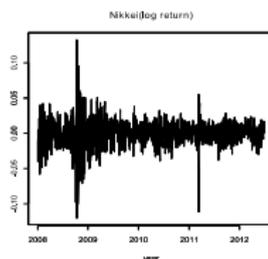
線形相関 $\rho_{X,Y}$ の高さ と 相関関係?

- 相関ゼロのランダムノイズの二つの時系列 $\sim N(0, 1) \Rightarrow \text{cor}(x_1, x_2) = 0$
- 新たにシンクロした大きなショックが加わると $\sim N(0, 10) \Rightarrow \text{cor}(x_1, x_2) = 0.8$

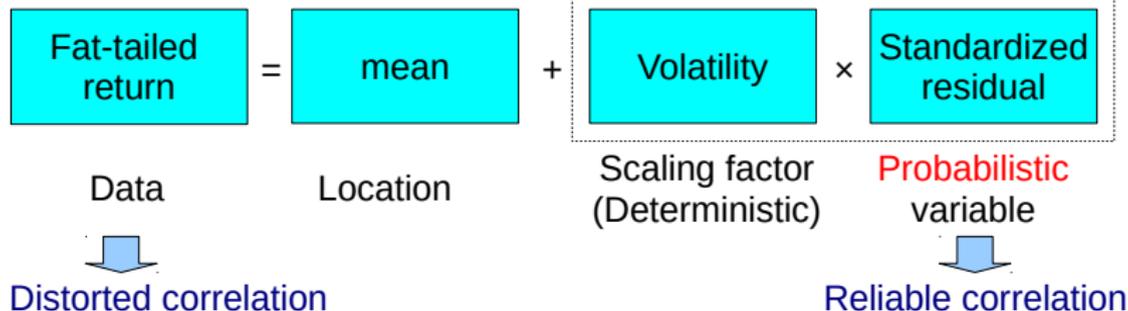
ボラティリティ変動要因の分離（残差の相関）

GARCH モデルでフィルタリング:

- 推定したボラティリティの変動要因をリターン系列から除去する



i.i.d.; variance=1



ARMA-GARCH モデル

(M)GARCH:

$$\begin{aligned} \mathbf{r}_t &= \boldsymbol{\mu}_t + \boldsymbol{\varepsilon}_t = \boldsymbol{\mu}_t + \mathbf{H}_t^{1/2} \mathbf{z}_t \\ \boldsymbol{\mu}_t &= E(\mathbf{r}_t | \mathcal{F}_{t-1}), \quad E(\mathbf{z}_t) = 0, \quad \text{Var}(\mathbf{z}_t) = \mathbf{I}_n \end{aligned} \quad (1)$$

where \mathbf{H}_t is a conditional variance covariance matrix (volatility), \mathbf{I}_n is an identity matrix of order n , and \mathcal{F}_{t-1} is the information set at time t .

Mean model: conditional means modeled by ARMA(P , Q), separately

$$\mu_t = \mu + \sum_{i=1}^P a_i r_{t-i} + \sum_{j=1}^Q b_j \varepsilon_{t-j} \quad (2)$$

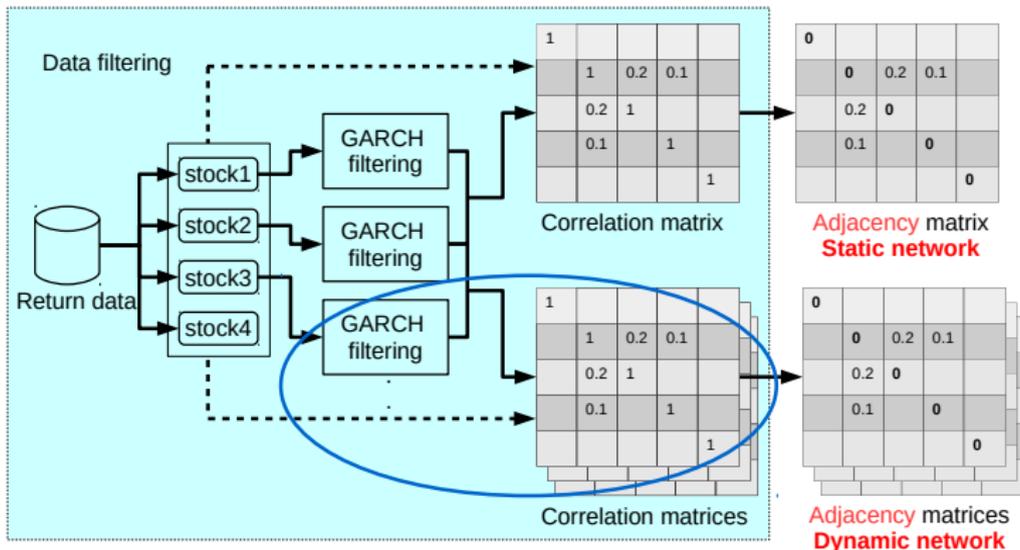
Volatility model: conditional volatilities as a vector form of GARCH(p, q)

$$\mathbf{h}_t = \boldsymbol{\omega} + \sum_{i=1}^q \mathbf{A}_i \boldsymbol{\varepsilon}_{t-i} \odot \boldsymbol{\varepsilon}_{t-i} + \sum_{j=1}^p \mathbf{B}_j \mathbf{h}_{t-j} \quad (3)$$

where \odot denotes the Hadamard operator (the entrywise product).

動的相関行列の推定のためのデータフィルタリング

データプロセッシングのイメージ:



● 動的な相関行列の推定方法

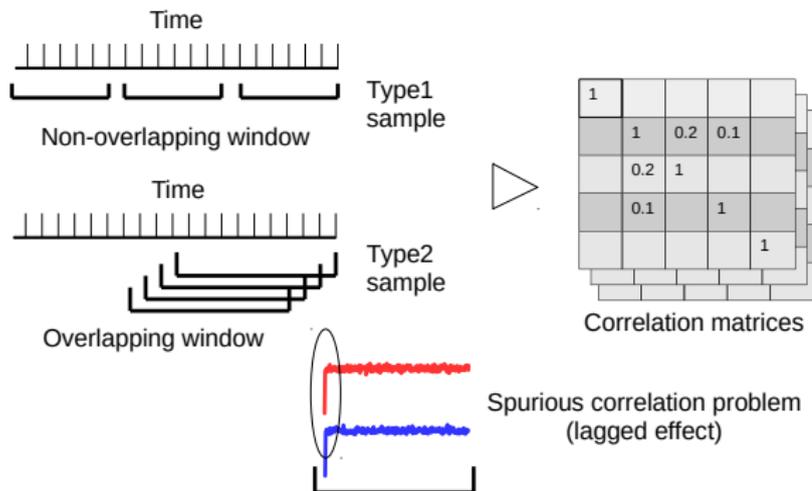
- ▶ 'Moving window' (--->) あるいは 'モデルベース' の推定方法 (→)

- ① はじめに
- ② 資産相関とボラティリティ
- ③ 動的相関の推定**
- ④ 動的相関行列の固有値分解
- ⑤ 動的ネットワークとしての分析・観察
- ⑥ まとめ

(1) Moving window 法による標本相関の計算 (一般的な手法)

Unfiltered log-returns の標本相関行列 R_t :

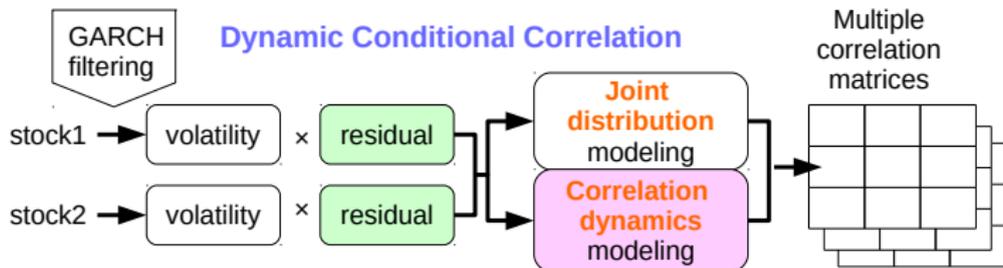
- R_t は時点 t ではなく、window (e.g., 30 取引日) に依存する。



- 市場横断的に大きなショックが発生すると、そのショックが Window に含まれる限り、相関の歪み・偽相関などの問題を引き起こし得る。⇒ Moving window 法の問題点

(2) モデルベースの動的相関行列の推定

DCC-GARCH: 動的な相関を持つ多変量 GARCH モデル



- 同時分布のモデル化には、 t -copula を用いる

$$f\left(\mathbf{r}_t | \boldsymbol{\mu}_t, \sqrt{\mathbf{h}_t}, \mathbf{R}_t, \eta\right) = c^{S_t}\left(u_{1 \cdot t}, \dots, u_{N \cdot t} | \mathbf{R}_t, \eta\right) \prod_{i=1}^N \frac{1}{\sqrt{h_{i \cdot t}}} f_{i \cdot t}\left(z_{i \cdot t} | \theta_i\right) \quad (4)$$

where $u_{i \cdot t} = F_i(r_{i \cdot t} | \mu_{i \cdot t}, \sqrt{h_{i \cdot t}}, \theta_i)$, θ_i is a parameter set including the ARMA-GARCH parameters and distributional parameters of i.i.d. residual z_i , $c^{S_t}(\cdot)$ is the Student t -copula density function, and η is the shape parameter of the Student t -copula.

(2) モデルベースの動的相関行列の推定

DCC-GARCH: 動的な相関のモデル化

- R_t の時間依存構造のモデル化: 代理変数 Q を用いて間接的に時間依存の構造を示す

$$R_t = \text{diag}(Q_t)^{-\frac{1}{2}} Q_t \text{diag}(Q_t)^{-\frac{1}{2}} \quad (5)$$

$$Q_t = \bar{Q} + \sum_{i=1}^m a_i (z_{t-i} z'_{t-i} - \bar{Q}) + \sum_{j=1}^n b_j (Q_{t-i} - \bar{Q}) \quad (6)$$

Q_t は R_t の proxy 変数、 z_t は standardized residuals (shocks).

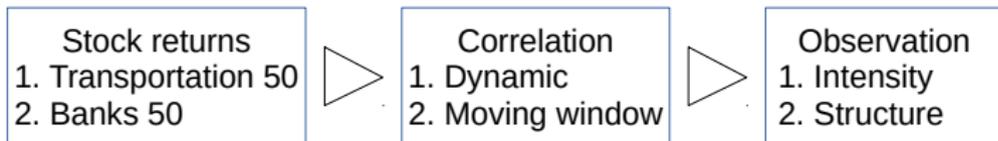
- ① はじめに
- ② 資産相関とボラティリティ
- ③ 動的相関の推定
- ④ 動的相関行列の固有値分解**
- ⑤ 動的ネットワークとしての分析・観察
- ⑥ まとめ

日本の株価を用いたケーススタディ

2 業種のサンプルデータセットを用意した:

- 時価総額順に 50 の銘柄を 2 業種からそれぞれ選んだ:
 - **輸送用機器** (自動車や関連部品等)
 - **銀行** (メガバンク、地銀等)
- 観測期間: 2008 年 1 月 ~ 2015 年 6 月; 日次終値ベースの対数収益率
 - ▶ リーマンショック (2008 年) と東日本大震災 (2011 年) を含む。
- 相関: **動的な相関 (DCC)** と **moving window** 法による動的な相関 (**窓の幅 200 取引日**)

Comparative analysis



モデルベースの動的相関の計算

DCC-GARCH モデルのパラメータ推定:

- 株価データから DCC-GARCH のパラメータを推定する
- 推定したパラメータに基づいて、各時点の動的相関 R_t を計算できる

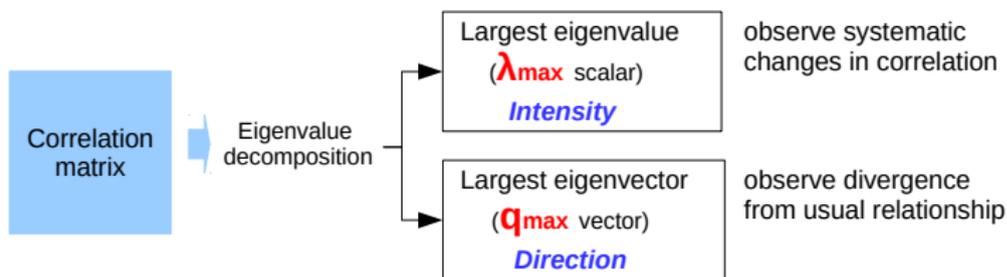
DCC Estimation Result

Sector	m, n	a_1	b_1	b_2	b_1+b_2	η
Transportation equipment	1, 2	0.0061 (0.0007)	0.4161 (0.0808)	0.4493 (0.0805)	0.8654	30.4160 (1.2898)
Banks	1, 2	0.0094 (0.0009)	0.3940 (0.0686)	0.4461 (0.0695)	0.8402	21.5952 (1.0033)

Note: DCC order (m, n) and parameters a_1 , b_1 , and b_2 are defined in (6). η is the shape parameter of the Student t -copula in (4).

相関行列の動的な変化の分析

動的相関行列の固有値分解 (固有値、固有ベクトル):



- 固有ベクトル \Rightarrow 最大固有値に対応する固有ベクトル $\mathbf{q}_{max \cdot t}$ について、**全期間平均 $\bar{\mathbf{q}}_{max}$** (benchmark) からの距離をもとに相関構造における重大な変化を検出する。
- 距離指標として、**標準コサイン距離** (normalized cosine distance) を用いる

$$\cos(\theta) = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|} = \frac{\sum x_i y_i}{\sqrt{\sum x_i^2 \sum y_i^2}}, \quad \gamma(\mathbf{x}, \mathbf{y}) = 1 - \cos(\theta) \quad (7)$$

$$\nu_{max \cdot t} = \frac{\gamma(\mathbf{q}_{max \cdot t}, \bar{\mathbf{q}}_{max})}{\text{std}(\gamma_{max})} \quad (8)$$

最大固有値だけみれば十分か？

同サイズのランダム行列の最大固有値の分布の上限と比べて判断する：

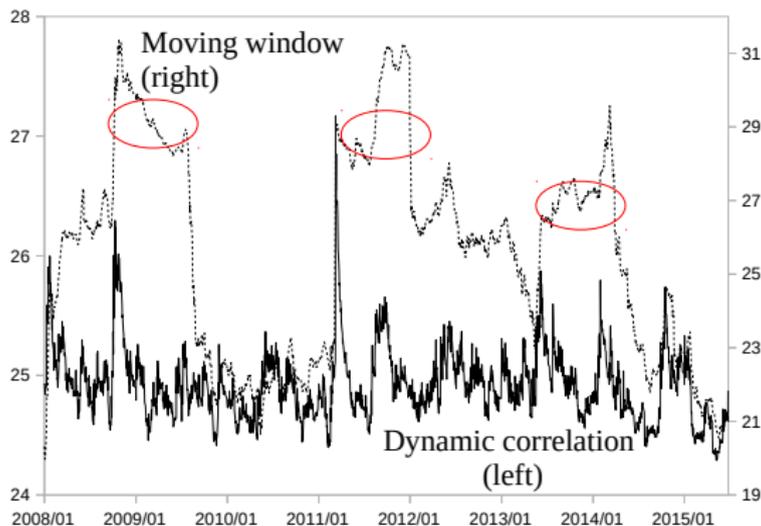
- 輸送用機器、銀行ともに最大固有値だけみればほぼ問題ない

Eigenvalues of Dynamic Correlation Matrix

	Eigenvalue		99 percentile
	Largest (min - max)	2nd largest (min - max)	Tracy-Widom
Transportation equipment			
Dynamic	24.28 - 27.17	0.97 - 1.40	1.38
Moving Average	19.97 - 31.39	1.32 - 2.67	
Banks			
Dynamic	31.56 - 35.20	1.25 - 1.73	1.38
Moving Average	25.89 - 39.60	0.84 - 2.70	

Note: Eigenvalues of R_t are calculated on every trading day during the observation period. The min and max represent the minimum and maximum of the vector of corresponding eigenvalues, respectively.

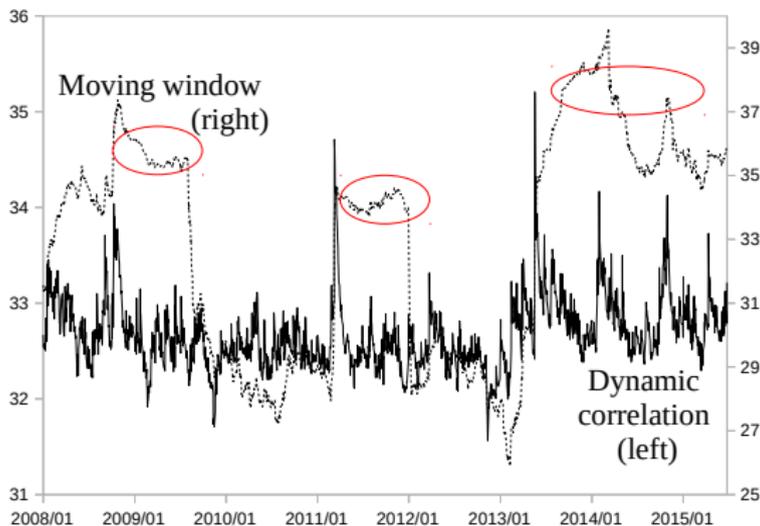
最大固有値の推移: 輸送用機器



DCC による動的相関の方が変化に対する反応がよい:

- **市場混乱期** (2008, 2011) には、相関の強さが増している。
- DCC と moving window では、固有値の水準、変化のパターンがかなり違う。
- 予想通り 'Moving window' には、**lagged effects** が見られた。
 - ▶ 多数のショックが続くとこの効果がより強く現れる。

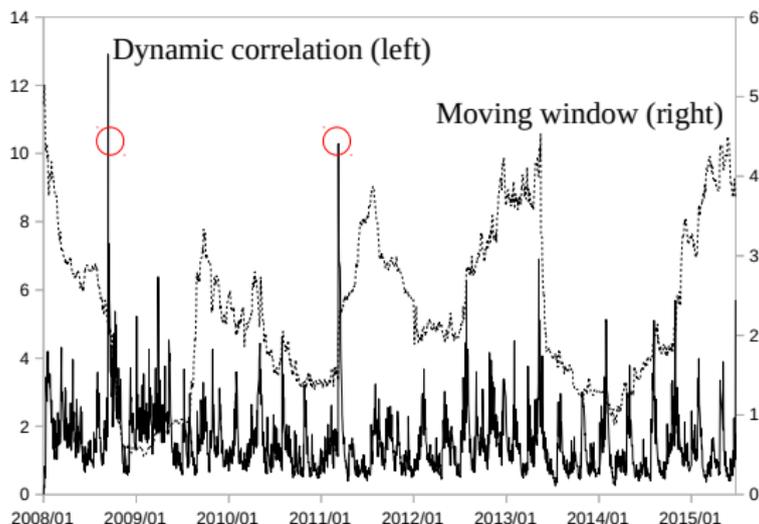
最大固有値の推移: 銀行



DCC による動的相関の方が変化に対する反応がよい:

- 市場混乱時に高い相関.
- 'Moving window' には明らかに lagged effect あり.
 - ▶ 相関の強さの変化の度合いとタイミングを把握し難い.

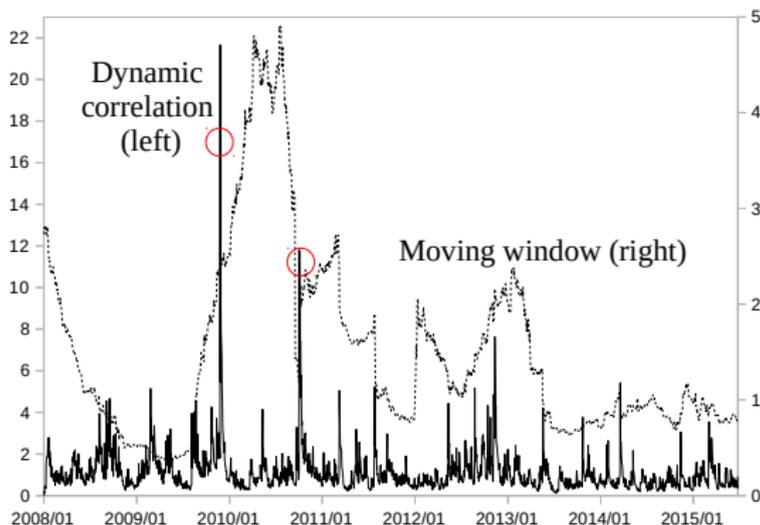
標準コサイン距離の推移: 輸送用機器



DCC による動的相関の方が変化に対する反応がよい:

- 期間平均からの乖離が大きい ⇒ 通常とは異なる相関パターン.
- ‘DCC による動的な相関’ は、相関の変化の発生時点を明確に示唆している.
- ‘Moving window’ は、相関の強さ同様、深刻な lagged effects が認められる.

標準コサイン距離の推移: 銀行



DCC による動的相関の方が変化に対する反応がよい:

- 2008年のリーマンショック時に目立ったベンチマークからの乖離はない
⇒ 平常時と同じ相関パターンであったが、より強い相関が認められた。
- 'Moving window' は、lagged effects のため使いづらい

- ① はじめに
- ② 資産相関とボラティリティ
- ③ 動的相関の推定
- ④ 動的相関行列の固有値分解
- ⑤ 動的ネットワークとしての分析・観察
- ⑥ まとめ

動的ネットワークのトポロジー変化

相関ネットワークの変化を観察する:

注目する指標

- ネットワーク密度 $D(\mathbf{A})$: 全ノード (銘柄) 間の全体的な連結強度 (相関の強さ) を表す指標. 1 に近いほど全てのノードの相関が強いことを意味する.

$$D(\mathbf{A}) = \frac{\sum_i \sum_{j>i} \mathbf{A}_{ij}}{n(n-1)/2} = \frac{\text{mean}(\mathbf{k})}{n-1} \approx \frac{\text{mean}(\mathbf{k})}{n} \quad (9)$$

$$k_i = \sum_{j \neq i} \mathbf{A}_{ij} \quad (10)$$

k_i はノード次数

動的ネットワークのトポロジー変化

相関ネットワークの変化を観察する:

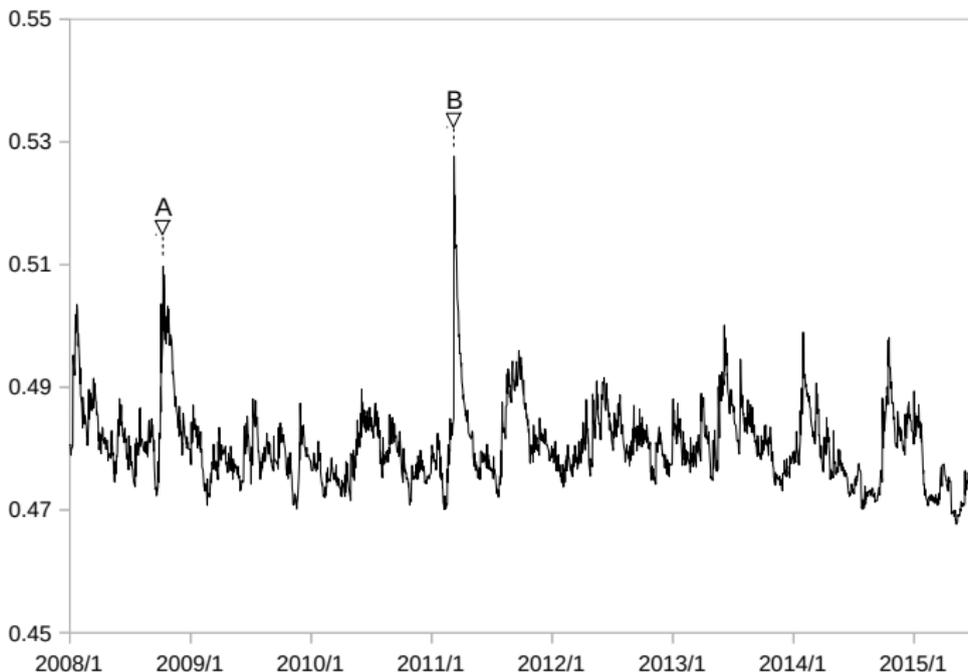
- ネットワーク中心性 $C(\mathbf{A})$: 一つのノードが全ノードと完全に連結 (相関 1) していて、それらのノードは互いにつながっていない (星型ネットワーク) 時に 1 となる. 全ノードが同じ程度の連結度を持っていれば 0.

$$\begin{aligned}
 C(\mathbf{A}) &= \frac{n}{n-2} \left(\frac{\max(\mathbf{k})}{n-1} - \frac{\text{mean}(\mathbf{k})}{n-1} \right) \\
 &= \frac{n}{n-2} \left(\frac{\max(\mathbf{k})}{n-1} - D(\mathbf{A}) \right) \approx \frac{\max(\mathbf{k})}{n} - D(\mathbf{A})
 \end{aligned} \tag{11}$$

- ネットワーク異質性 $H(\mathbf{A})$: ノード間の連結度の違いを示す (連結度の分布の相対的な散らばりとして定義). 高いほど相関の高低が存在する.

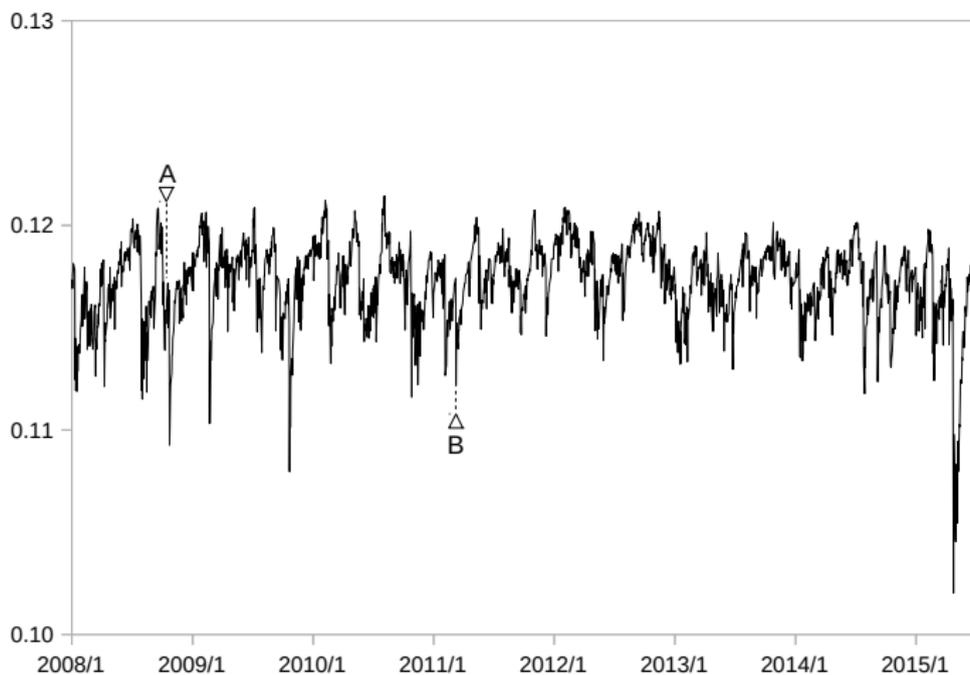
$$H(\mathbf{A}) = \frac{\sqrt{\text{var}(\mathbf{k})}}{\text{mean}(\mathbf{k})} = \sqrt{\frac{n \sum_i k_i^2}{(\sum_i k_i)^2} - 1} \tag{12}$$

ネットワーク密度: 輸送用機器



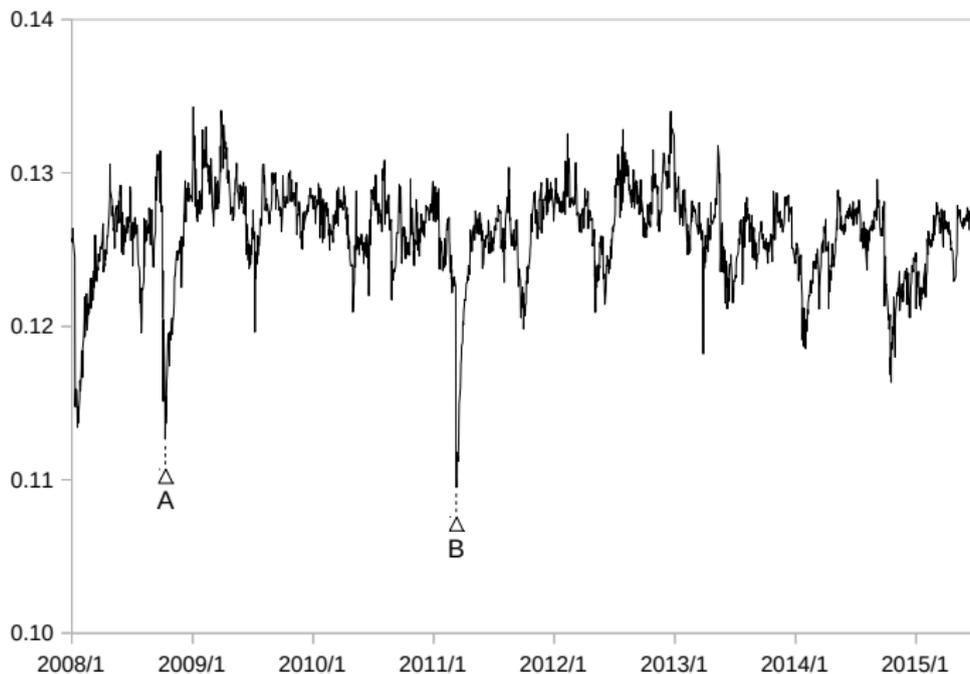
- リーマンショック (A)、東日本大震災 (B) では、ネットワーク密度が高まっている (市場混乱時の相関の強まり)。

ネットワーク中心性: 輸送用機器



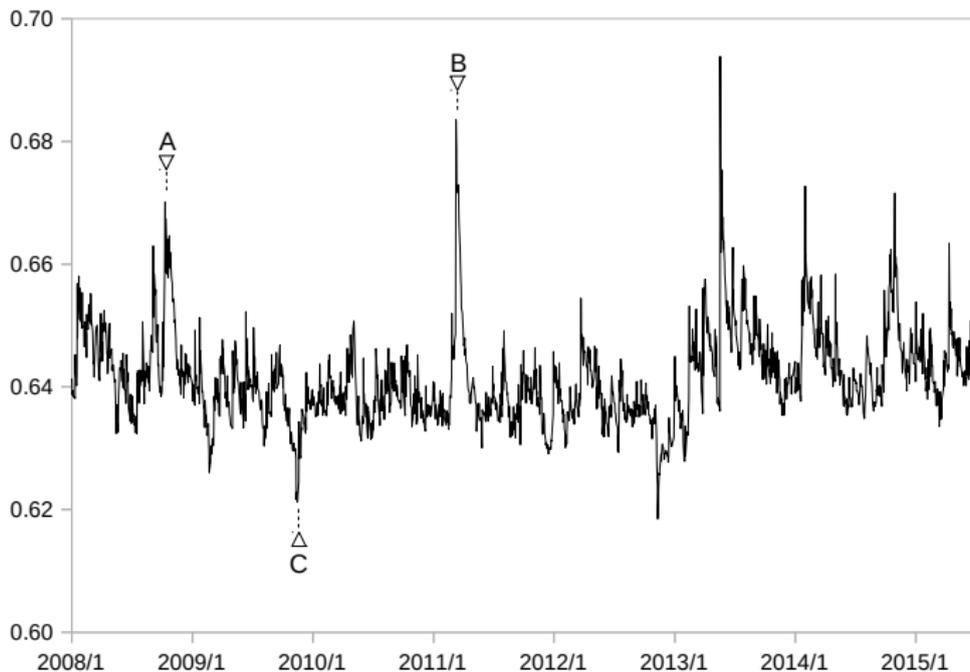
- 中心性には大きな変化はない（銀行よりも中心性は相対的に高い）。

ネットワーク異質性: 輸送用機器



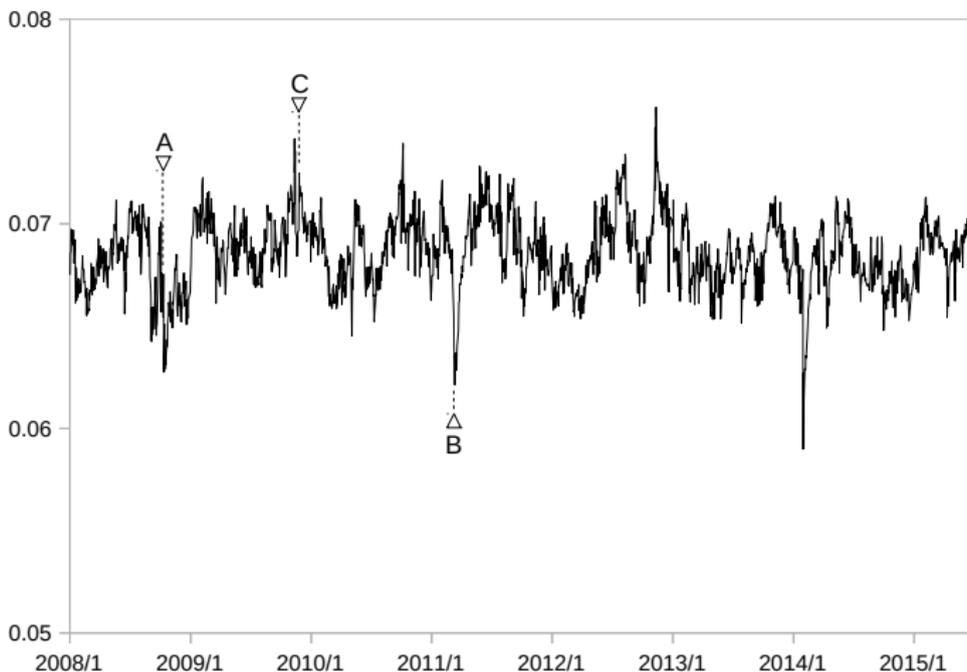
- リーマンショック (A)、東日本大震災 (B) では、ネットワーク異質性が低下している (市場混乱時に相関パターンがより一様化している)。

ネットワーク密度: 銀行



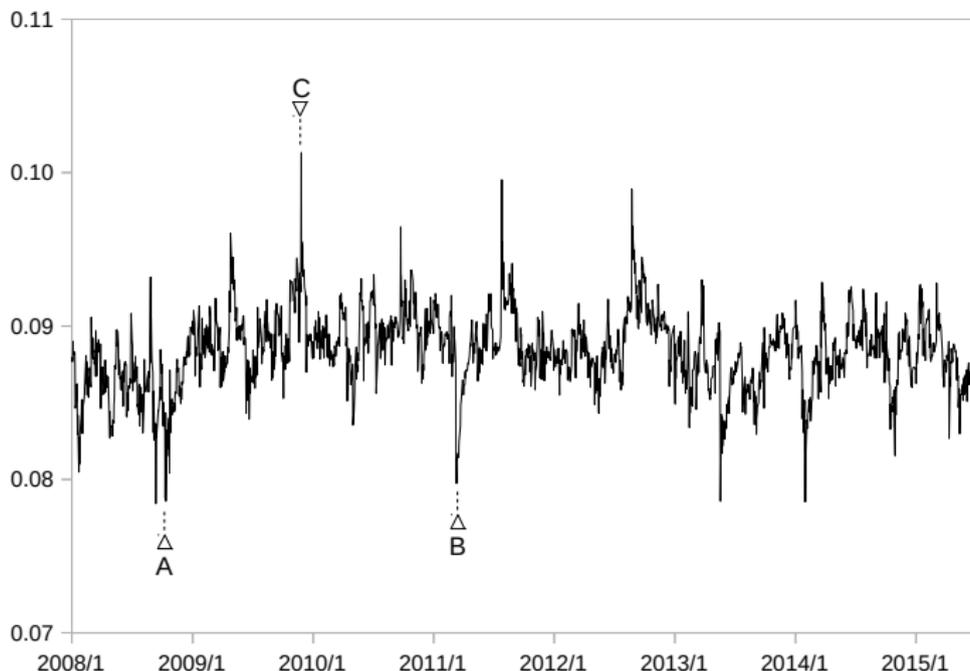
- リーマンショック (A)、東日本大震災 (B) では、ネットワーク密度が高まっている (市場混乱時の相関の強まり)。輸送用機器よりも銀行の方が密度が高め。

ネットワーク中心性: 銀行



- 中心性には大きな変化はない.

ネットワーク異質性: 銀行



- リーマンショック (A)、東日本大震災 (B) では、ネットワーク異質性が低下している (市場混乱時に相関パターンがより一様化している)。

- ① はじめに
- ② 資産相関とボラティリティ
- ③ 動的相関の推定
- ④ 動的相関行列の固有値分解
- ⑤ 動的ネットワークとしての分析・観察
- ⑥ **まとめ**

まとめ

主な分析結果:

- ファットテールな株価の収益率の動的な相関を DCC-GARCH モデルを用いて推定した.
- 日本株のデータによる実際の推定例では、異時点の比較や変化点の特定などの目的に対しては、動的相関行列は利便性が高いことが示された.
- 伝統的に用いられている標本データに基づく moving window 法による相関は、偽相関のリスクが高く、ダイナミックな分析でも問題が多い.
- 動的ネットワークとしての分析では、輸送用危機、銀行の 2 業種について、市場混乱時にネットワーク密度の上昇、異質性の低下が顕著にみられた。ポートフォリオのリスク管理や市場モニタリングに応用できる可能性あり.

今後の課題:

- より包括的かつ詳細な株価の動的な相関分析（異時点間比較、構造変化抽出など）.
- ダイナミックネットワーク分析へのさらなる応用、動的コミュニティ分析など.

ご清聴ありがとうございました.