平成 23 年度科学研究費基盤研究(A)
研究課題番号 ： 23244011
研究課題名 ： 非対称・非線形統計理論と経済・生体科学への応用
研究代表者 ： 谷口 正信（早稲田大学）

生体数理・社会数理の統計科学

プログラム

2012 年 3 月 1 日～ 3 日

早稲田大学理工学部 55 号館 S 棟 第 3 会議室

開催責任者 ： 谷口 正信 （早大基幹理工）
村田 昇 （早大先進理工）
南 美穂子 （慶大理工）
谷合 弘行 （早大国際教養）

3 月 1 日（木）

(Chair H. Ogata)

13:30– 14:00 : Statistically Efficient Construction of α-risk Minimizing Portfolio
Hiroyuki TANIAI (Waseda Univ.)

14:00 – 14:30 : Semiparametric penalized spline regression
Takuma YOSHIDA* (Shimane Univ.) &
Kanta NAITO (Shimane Univ.)

14:30 – 15:00 : A Conjugate Property between Loss Functions and Uncertainty Sets in Classification Problems
Takafumi KANAMORI* (Nagoya Univ.),
Akiko TAKEDA (Keio Univ.) & Taiji SUZUKI (Tokyo Univ.)

(Chair A. Petkovic)

15:30 – 16:00 : Statistical Estimation for CAPM with Long-Memory Dependence
Tomoyuki AMANO* (Wakayama Univ.),
Tsuyoshi KATO & Masanobu TANIGUCHI (Waseda Univ.)

16:00 – 16:30 : Statistical Estimation of Dynamic Portfolios
Hiroshi SHIRAISHI (Jikei Medical Univ.)

16:30 – 17:00 : When does seasonality exist ? A synchrosqueezing approach to determining the seasonality
Ming-Yen CHENG (National Taiwan Univ.)

3 月 2 日（金）

(Chair H. Taniai)

9:30 – 10:00 : Learning a linear non-Gaussian structural equation model from multiple datasets
Shohei SHIMIZU (Osaka Univ.)

10:00 – 10:30 : A Simple Test of Long Memory Versus Structural Breaks
Katsumi SHIMOTSU (Hitotsubashi Univ.) &
Keiko YAMAGUCHI* (Waseda Univ.)

(Chair T. Amano)

11:00 – 11:30 : The optimal portfolio problem with the ideal balance and exogenous variable causality
Kiyoaki SASAKI (Niigata Univ.) &
Junichi HIRUKAWA* (Niigata Univ.)

11:30 – 12:00 : Smooth Transition Quantile Capital Asset Pricing Models with Heteroscedasticity
Cathy W. S. CHEN (Feng Chia Univ.)

(Chair N. Murata)

13:30 – 14:00 : Cyclic cubic regression spline smoothing and analysis of slowly changing cyclic variation
Mihoko MINAMI (Keio Univ.)

14:00 – 14:30 : Some generalization of universal portfolio
Jun'ichi TAKEUCHI* (Kyushu Univ.) &
Mariko TSURUSAKI (Kyushu Univ.)

14:30 – 15:00 : Machine learning methods for brain-machine interface
Shin ISHII (Kyoto Univ.)

(Chair　M. Taniguchi)

15:30 – 16:00 : Generalized Cordeiro & Ferrari
Bartlett-type adjustment

Yoshihide KAKIZAWA (Hokkaido Univ.)

16:00 – 16:30 : TailCoR: a new measure of tail association

David Veredas (Univ. Libre de Bruxelles)

16:30 – 17:00 : Group LASSO for Structual Break Time Series.

Ngai Hang CHAN (Chinese University of Hong Kong)

17:00 – 17:30 : Frequency domain techniques in the analysis of DNS sequences

David Stoffer (Univ. Pittsburgh)

3 月 3 日（土）

(Chair　H. Shiraishi)

9:30 – 10:00 : Robust Portfolio Estimation under Skew-Normal Return Processes.

Alexandre PETKOVIC* (Waseda Univ.),
Masanobu TANIGUCHI (Waseda Univ.) &
Thomas DiCiccio (Cornell Univ.)

10:00 – 10:30 : Joint estimation of copula and quantiles for time series
--- with estimating function approach ---

Hiroaki OGATA (Waseda Univ.)

(Chair　J. Hirukawa)

11:00 – 11:30 : Dynamic estimation of investment style for Equity managers

Takashi YAMASHITA (GPIF)

11:30 – 12:00 : Maximum entropy test for time series models

Sangyeol LEE (Seoul National Univ.)

Statistically Efficient Construction of α-risk Minimizing Portfolio

Hiroyuki TANIAI (*Waseda University*)

Symposium: *Statistics for Biomedical & Social Mathematical Sciences*
March 2012, Waseda

In this talk, we consider the problem that minimize a quantity known as α-risk by reallocate the weight of several financial assets. Particularly, we remark that we can improve the statistical efficincy of minimizing procedure by regarding it from an semiparametric vies point. The summary is as follows.

To begin with, the "α-risk" here we concern is those which defined by

$$\varrho_{\nu_\alpha}(X) := -\frac{1}{\alpha}\int_0^\alpha F_X^{\leftarrow}(u)du \qquad (F_X^{\leftarrow}(\alpha) := \inf\{x : F_X(x) \geq \alpha\}).$$

We consider a portfolio wich consists of p financial assets with allocation weight $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_p)^\top$, and seek to minimize its α-risk $\varrho_{\nu_\alpha}(\boldsymbol{X}^\top\boldsymbol{\pi})$. Bassett et al. (2004) already has showed that this optimization problem can be solved by means of Quantile Regression (QR) of Koenker and Bassett (1978). However, this QR estimator is not necessarily efficient: It can be point out that QR estimator can be regarded as a QMLE based on Asymmetric Laplace distribution: $\alpha(1-\alpha)\exp\{\rho_\alpha(\xi)\}$ (cf. Komunjer (2005)).

Here in this talk, we focus on the fact that QR is a prametrization which is essentially semi-parametric. That is, recalling that a QR model is such that $F^{\leftarrow}_{Z_i|\boldsymbol{W}_i=\boldsymbol{w}_i}(\alpha) = \boldsymbol{w}_i^\top\boldsymbol{\beta}(\alpha)$ with conditional quantile function $F^{\leftarrow}_{X|S}(\alpha) := \inf\{x : P(X \leq x|S) \geq \alpha\}$, such a model is a submodel of (e.g.,) the following class of quantile restricted models:

$$Z_i = \boldsymbol{W}_i^\top\boldsymbol{b} + \xi_i, \qquad \xi_i \overset{iid}{\sim} G, \tag{1}$$
$$\text{s.t. } g \in \mathcal{F}^\alpha := \Big\{f \Big| \int_{-\infty}^0 f(x)dx = \alpha = 1 - \int_0^\infty f(x)dx\Big\}.$$

Although these parematrization is not unique with respect to the given QR model, still if this parametrization, the model, is known to satisfy the Local Asymptotic Normality (LAN) then we may consider the semiparametrically efficient inference based on it. To this end, we apply the invariance approach of Hallin and Werker (2003). Namely, by referring Section 4.1 of Hallin et al. (2008), we construct the "one-step estimator" as

$$\tilde{\boldsymbol{b}}_f^{(n)} := \bar{\boldsymbol{b}}_n + \hat{\boldsymbol{\Sigma}}_{fg}^{-1}\frac{\underset{\sim}{\boldsymbol{\Delta}}^{(n)}_{\bar{\boldsymbol{b}}_n,f}}{\sqrt{n}}, \qquad \hat{\boldsymbol{\Sigma}}_{fg} := \hat{\mathcal{I}}_{fg}\hat{S}^{(n)}_{\boldsymbol{W}} + \frac{\alpha}{1-\alpha}\cdot\frac{f(0)}{-\alpha}\cdot\hat{\mu}_{\varphi_g,L}\bar{\boldsymbol{W}}^{(n)}\bar{\boldsymbol{W}}^{(n)\top},$$

where $\bar{\boldsymbol{b}}_n$ corresponds, in this case, to the discretization of QR estimator $\hat{\boldsymbol{\beta}}^{(n)}(\alpha) \equiv \operatorname{argmin}_{\boldsymbol{b}} \sum \rho_\alpha(Z_i - \boldsymbol{W}_i^\top \boldsymbol{b})$.

Consequently we provide several simulation results, which highlight the efficiency gain from $\hat{\boldsymbol{\beta}}^{(n)}(\alpha)$ to $\tilde{\boldsymbol{b}}_f^{(n)}$. Among these, here in this report we give Table 1 and Figure 1 below. The data are generated as the same situation as that of Bassett et al. (2004), so the true density g of (1) must be a mixture of Gaussian, χ^2 and Reversed χ^2 distribution. As shown in Table 1, we successfully improve the estimation errors. Still, you may think that Figure 1 is not as significant as Table 1 says: If we examine the result more deeply, we may find that those efficiency gain are obtained by trimming the behavior of the resulting mixture density at outliers.

$f \quad \diagdown \boldsymbol{\pi}(\alpha)$	$\pi_1(.1)$	$\pi_2(.1)$	$\pi_3(.1)$	$\pi_4(.1)$	$\pi_1(.5)$	$\pi_2(.5)$	$\pi_3(.5)$	$\pi_4(.5)$
$n = 1000$								
AL	0.8702	0.9084	0.5621	0.8193	0.9150	0.8985	0.2225	0.3614
N	0.9416	0.9485	0.8680	0.9112	0.7850	0.7714	0.4019	0.5214
LGT	0.9453	0.9532	0.8640	0.9085	0.8101	0.7981	0.4043	0.5199
$\mathrm{APD}_{1.5}$	0.9290	0.9296	0.9242	0.9645	0.8206	0.8077	0.3572	0.4806

Table. 1 $\mathrm{Var}[\tilde{\boldsymbol{b}}_f^{(n)}]/\mathrm{Var}[\hat{\boldsymbol{\beta}}^{(n)}(\alpha)]$ for situation of Bassett et al. (2004)

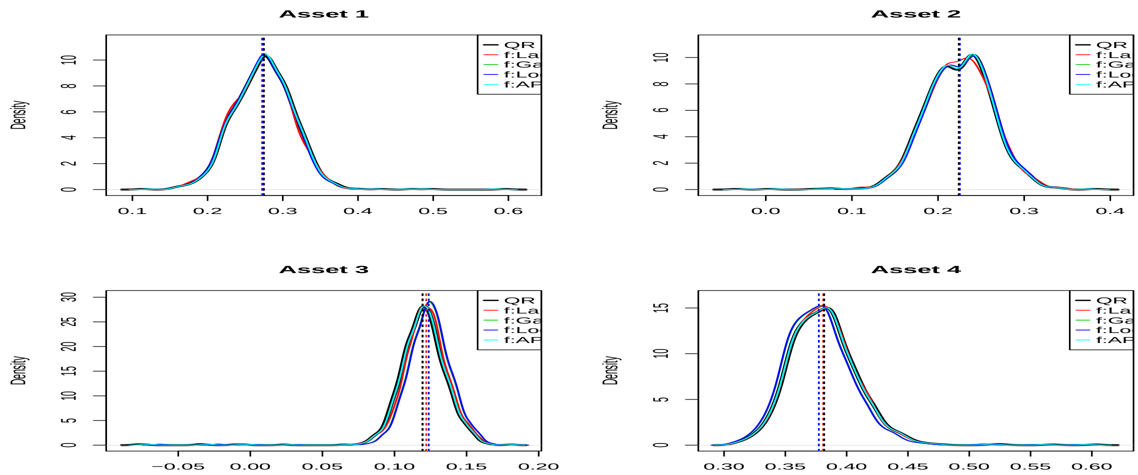

Fig. 1 Behavior of $\hat{\boldsymbol{\beta}}^{(n)}(\alpha)$ and $\tilde{\boldsymbol{b}}_f^{(n)}$, DGP = Bassett et al. (2004), $n = 1000$, $\alpha = 0.1$

References

Bassett, G. W. J., Koenker, R., and Kordas, G. (2004). Pessimistic portfolio allocation and choquet expected utility. *Journal of Financial Econometrics*, 2(4):477–492.

Hallin, M., Vermandele, C., and Werker, B. J. M. (2008). Semiparametrically efficient inference based on signs and ranks for median-restricted models. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 70(2):389–412.

Hallin, M. and Werker, B. J. M. (2003). Semi-parametric efficiency, distribution-freeness and invariance. *Bernoulli*, 9(1):137–165.

Koenker, R. and Bassett, Jr., G. (1978). Regression quantiles. *Econometrica*, 46(1):33–50.

Komunjer, I. (2005). Quasi-maximum likelihood estimation for conditional quantiles. *J. Econometrics*, 128(1):137–164.

Semiparametric penalized spline regression

吉田 拓真 [1]　　内藤 貫太 [2]

[1] 島根大学大学院総合理工学研究科 [2] 島根大学総合理工学部

1 はじめに

散布図平滑化問題を考えるとき, 2 つの代表的なアプローチとして, パラメトリック法とノンパラメトリック法がある. パラメトリック法は未知の回帰関数 $f(x)$ が有限個のパラメータによって規定されるモデル (例:多項式モデル) を想定し, 回帰関数を推定する代わりにパラメータを推定する手法である. この手法のメリットは推定量の解釈がしやすいことであるが, 一方で, 真の回帰関数がパラメトリックモデルであるかどうかはわからず, そこにはギャップ (Model bias) があるので, 想定したモデルが不適切である可能性がある. これはパラメトリック法のデメリットであると言える.

ノンパラメトリック法は回帰関数に大きい制約を置かず, 直接 $f(x)$ を推定する手法である. この手法のメリットは, 推定量は一般に一致性を有することである. しかし, 推定量の式は複雑で解釈がしにくい.

本講演ではパラメトリック法とノンパラメトリック法を組み合わせ, $f(x)$ の大きな挙動はパラメトリック推定量によって予測され, 挙動が複雑な細かいところはノンパラメトリック推定量によって修正されるようなセミパラメトリック法に基づく推定量を構築した. また, 推定量が一致性を有すること, 最適なパラメトリックモデルを選べばノンパラメトリック推定量よりも優れた推定量であること, パラメトリックモデルの選択法を示した. 本講演ではノンパラメトリック法として, スプライン法を用いている. 最後に, パラメトリックモデル選択の実装と構成した推定量の精度をシミュレーションによって検証した.

2 セミパラメトリック罰則付きスプライン推定量

与えられたデータ $\{(x_i, y_i) : i = 1, \cdots, n\}$ に対して, 回帰モデル

$$y_i = f(x_i) + \varepsilon_i, \quad i = 1, \cdots, n,$$

を考える. 目的は $f(x)$ をスプライン法に基づくセミパラメトリック法で推定することである. まず, パラメトリックモデル $f(x|\boldsymbol{\beta}), \boldsymbol{\beta} \in B \subseteq \mathbb{R}^m$ を用意し, 然るべき手法 (最小二乗法等) でパラメータ $\boldsymbol{\beta}$ の推定量 $\hat{\boldsymbol{\beta}}$ を得る. すると $f(x)$ は

$$f(x) = f(x|\hat{\boldsymbol{\beta}}) + f(x|\hat{\boldsymbol{\beta}})^\gamma r_\gamma(x, \hat{\boldsymbol{\beta}}),$$

と表せる. ここで, $r_\gamma(x, \boldsymbol{\beta}) = \{f(x) - f(x|\boldsymbol{\beta})\}/f(x|\boldsymbol{\beta})^\gamma$, $\gamma \in \{0, 1\}$ である. 次に, データ $\{(x_i, \{y_i - f(x_i|\hat{\boldsymbol{\beta}})\}/f(x_i|\hat{\boldsymbol{\beta}})^\gamma) : i = 1, \cdots, n\}$ に罰則付きスプライン法を適用し, $r_\gamma(x, \hat{\boldsymbol{\beta}})$ の推定量 $\hat{r}_\gamma(x, \hat{\boldsymbol{\beta}})$ を得る. セミパラメトリック罰則付きスプライン推定量 (SPSE) は

$$\hat{f}(x, \gamma) = f(x|\hat{\boldsymbol{\beta}}) + f(x|\hat{\boldsymbol{\beta}})^\gamma \hat{r}_\gamma(x, \hat{\boldsymbol{\beta}}),$$

と構成される.

3 SPSE の漸近的性質

SPSE $\hat{f}(x,\gamma)$ がいわゆる"良い推定量"であることをその漸近的性質によって保証する．適当な正則条件を仮定すると, 以下の Theorem が得られる．

Theorem 1. $f \in C^{p+1}, f(\cdot|\boldsymbol{\beta}_0) \in C^{p+1}$ とする．このとき, $x \in (0,1)$ に対して, $n \to \infty$ の元で,

$$\begin{aligned} E[\hat{f}(x,\gamma)] &= f(x) + b_a(x|\boldsymbol{\beta}_0,\gamma) + b_\lambda(x|\boldsymbol{\beta}_0,\gamma) + o_P(\xi_n) + o_P(C_n), \\ V[\hat{f}(x,\gamma)] &= O_P(K_n n^{-1}) \end{aligned}$$

が成り立つ．ただし, $\xi_n \to 0$, $C_n \to 0$, $\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) = O_P(1)$,

$$b_a(x|\boldsymbol{\beta}_0,\gamma) = \xi_n f(x|\boldsymbol{\beta}_0)^\gamma r_\gamma^{(p+1)}(x|\boldsymbol{\beta}_0), \quad b_\lambda(x|\boldsymbol{\beta}_0,\gamma) = C_n \psi(x|f, f(\cdot|\boldsymbol{\beta}_0)),$$

である．

Theorem 1 において, ξ_n, C_n, $\psi(x|f, f(\cdot|\boldsymbol{\beta}_0))$ の正確な表現は Yoshida and Naito (2012) を参照する．Theorem 1 より, SPSE は一致性を有することがわかる．

4 パラメトリックモデルの選択

SPSE を構成する上でパラメトリックモデルの選択は重要である．なぜなら, もし悪いモデルを選んでしまった場合, SPSE は通常のノンパラメトリック罰則付きスプライン (NPSE) より劣る推定量になりかねないからである．Claeskens et al. (2009) は NPSE の漸近バイアス $b_a(x), b_\lambda(x)$ を導出した．そこで, すべての $x \in (0,1)$ に対して

$$|b_a(x|\boldsymbol{\beta}_0,\gamma)| < |b_a(x)| \quad \text{and} \quad |b_\lambda(x|\boldsymbol{\beta}_0,\gamma)| < |b_\lambda(x)|, \text{ for all } x \in (0,1), \tag{1}$$

が成り立つように $f(x|\boldsymbol{\beta})$ を選べば, SPSE と NPSE の漸近分散は同等となるので, SPSE は NPSE よりも良い推定量であることが言える．本講演では (1) に基づいて $f(x|\boldsymbol{\beta})$ の決定アルゴリズムを示した:

1. 複数のパラメトリックモデルの候補を選ぶ: $\{f_k(x|\boldsymbol{\beta}_k)|k = 1, \cdots, K\}$.
2. 各 $f_k(\cdot|\boldsymbol{\beta}_k)$ に対して,
$$C(f_k(\cdot|\boldsymbol{\beta}_k)) = \#\{z_j|\ |b_a(z_j|\boldsymbol{\beta}_0,\gamma)| < |b_a(z_j)|, |b_\lambda(z_j|\boldsymbol{\beta}_0,\gamma)| < |b_\lambda(z_j)|, j = 1, \cdots, J\}$$
を計算する．(z_j は区間 $(0,1)$ の grid point である．)
3. 最適なパラメトリックモデルを $\underset{1 \le k \le K}{\operatorname{argmax}}\{C(f_k(\cdot|\boldsymbol{\beta}_k))\}$ とする．

5 シミュレーション

第 4 節で示した $f(x|\boldsymbol{\beta})$ の選択基準を用いて, 最適なモデル選択法を決定した．また SPSE と NPSE の精度比較を行った．$C(f_k(\cdot|\boldsymbol{\beta}_k))$ に基づいて選ばれたモデル $f(x|\boldsymbol{\beta})$ を用いて構築された SPSE$\hat{f}(x,\gamma)$ は NPSE よりも MISE 最小の意味で良い推定量であることが確認された.

References

[1] Claeskens,G., Krivobokova,T. and Opsomer,J.D. (2009). Asymptotic properties of penalized spline estimators. *Biometrika*. **96**, 529-544.

[2] Yoshida,T and Naito,K. (2012). Semiparametric penalized spline regression. submitted.

判別分析における損失関数と不確実性集合の共役性について

金森 敬文 (名古屋大)　武田朗子 (慶応大)　鈴木大慈 (東京大)

概 要

機械学習における２値判別問題に対して，これまで主に２つのアプローチが研究されてきた．ひとつは損失関数を用いる方法であり，もうひとつは不確実性集合を用いる方法である．

損失関数を用いる方法は，サポートベクターマシンやブースティングなど，主要な学習アルゴリズムに応用されている．データに対する適合の度合いを損失として定量化し，平均的な損失を最小化するという方針で学習アルゴリズムが組み立てられている．近年の数理最適化法の著しい発展もあり，損失関数に基づく学習法は広く応用されている．また，その統計的な性質も理論的に詳しく解析されている．

一方，不確実性集合を用いる方法として，ミニマックス確率マシンやマージン最大化ミニマックス確率マシンなどが提案されている．これは以下のように学習を行う．まず学習データの不確実性を表現するために，特徴空間において各出力ラベル対して不確実性集合を設定する．次に，それらの集合を最もよく分離する平面を判別境界とする．この方法は，ハードマージンサポートベクターマシンの幾何学的な側面を一般化した方法である．近年，ロバスト最適化との関連などが指摘され，研究が進展しつつある話題となっている．

本発表では,「共役性」をキーワードにして，これら２つのアプローチの関連について考察する．ここでは，ルジャンドル変換のもとでの共役性を考えている．まず，損失関数の共役関数を考え，そのレベルセットが不確実性集合を導出することを指摘する．その関係を軸として，不確実性集合を用いる学習法の統計的一致性など，理論的な性質について考察する．

Statistical Estimation for CAPM with Long-Memory Dependence.

Tomoyuki Amano (Wakayama Univ.)
Tsuyoshi Kato
Masanobu Taniguchi (Waseda Univ.)

Black version of CAPM expected return is given by $E[R_i] = \alpha_{im} + \beta_{im} E[R_m]$, where $\alpha_{im} = E[R_{om}(1-\beta_{im})]$.

However this model does not have a time dimension. For econometric analysis of the model, it is necessary to add an assumption concerning the time series behavior of returns and estimate the model over time.

In what follows, we discuss statistical estimation for Black version of CAPM.

Suppose that an n-dimensional financial return process $\{\boldsymbol{Y}_t = (Y_{1,t}, ..., Y_{n,t})'\}$ is generated by

$$\boldsymbol{Y}_t = \boldsymbol{\alpha} + \boldsymbol{B}'\boldsymbol{Z}_{mt} + \boldsymbol{\epsilon}_t, \tag{1}$$

where $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_n)'$ and $\boldsymbol{B} = \{\beta_{ij}; i = 1, ..., p, j = 1, ..., n\}$, $(n \leq p)$, are unknown vector and matrix, respectively, $\{\boldsymbol{Z}_{mt} = (Z_{1,t}, ..., Z_{p,t})'\}$ is an explanatory stochastic regressor process and $\{\boldsymbol{\epsilon}_t = (\epsilon_{1,t}, ..., \epsilon_{n,t})'\}$ is a sequence of disturbance process.

In view of empirical analysis for financial data, it is natural to assume that $\{\boldsymbol{Y}_t\}$ is short-memory, and $\{\boldsymbol{Z}_{mt}\}$ is long-memory.

For this we suppose that $\{\boldsymbol{Z}_{mt}\}$ and $\{\boldsymbol{\epsilon}_t\}$ are defined by

$$\boldsymbol{Z}_{mt} = \sum_{j=0}^{\infty} \boldsymbol{\gamma}_j \boldsymbol{a}_{t-j} + \sum_{j=0}^{\infty} \boldsymbol{\rho}_j \boldsymbol{b}_{t-j},$$
$$\boldsymbol{\epsilon}_t = \sum_{j=0}^{\infty} \boldsymbol{\eta}_j \boldsymbol{e}_{t-j} + \sum_{j=0}^{\infty} \boldsymbol{\xi}_j \boldsymbol{b}_{t-j},$$

where $\{\boldsymbol{a}_t\}$, $\{\boldsymbol{b}_t\}$ and $\{\boldsymbol{e}_t\}$ are p-dimensional zero-mean uncorrelated processes, and they are mutually independent. Here the coefficients $\{\boldsymbol{\gamma}_j\}$ nad $\{\boldsymbol{\rho}_j\}$ are $p \times p$-matrices, and all components of $\boldsymbol{\gamma}_j$ are ℓ^1-summable, (for short, $\boldsymbol{\gamma}_j \in \ell^1$), and those of $\boldsymbol{\rho}_j$ are ℓ^2-summable (for short, $\boldsymbol{\rho}_j \in \ell^2$). The coefficients $\{\boldsymbol{\eta}_j\}$ and $\{\boldsymbol{\xi}_j\}$ are $n \times p$-matrices, and $\boldsymbol{\eta}_j \in \ell^1$ and $\boldsymbol{\xi}_j \in \ell^2$.

Proposition 1.

If $\boldsymbol{B}'\boldsymbol{\rho}_j + \boldsymbol{\xi}_j = O(j^{-\alpha})$, $\alpha > 1$, then the process $\{\boldsymbol{Y}_t\}$ is short-memory dependent.

Proposition 1 provides an important view for the CAPM, i.e., if we assume natural conditions on (1) based on the empirical studies, then they impose a sort of "curved structure", i.e., $\boldsymbol{B}'\boldsymbol{\rho}_j + \boldsymbol{\xi}_j = O(j^{-\alpha})$.

Two-Stage Least Square Estimation

To estimate $\boldsymbol{B}$ in our setting, we introduce an instrument variable $\{\boldsymbol{X}_t\}$. $\{\boldsymbol{X}_t\}$ is $(r \times 1)$-vector $(p \leq r)$ and satisfies the condition $Cov(\boldsymbol{X}_t, \boldsymbol{Z}_{mt}) \neq 0$ and $Cov(\boldsymbol{X}_t, \boldsymbol{\epsilon}_t) = 0$. We consider the regression which is given by $\boldsymbol{Z}_{mt} = \boldsymbol{\delta}'\boldsymbol{X}_t + \boldsymbol{u}_t$.

Our 2SLS estimator $\hat{\boldsymbol{B}}_{2SLS}$ is defined by

$$\hat{\boldsymbol{B}}_{2SLS} = \left[\sum_{t=1}^{T} \hat{\boldsymbol{Z}}_{mt}\boldsymbol{Z}'_{mt}\right]\left[\sum_{t=1}^{T} \hat{\boldsymbol{Z}}_{mt}\boldsymbol{Y}'_{t}\right],$$

where $\hat{\boldsymbol{Z}}_{mt} = \hat{\boldsymbol{\delta}}'\boldsymbol{X}_t$ and $\hat{\boldsymbol{\delta}}$ is the OLS of $\boldsymbol{Z}_{mt}$ on $\boldsymbol{X}_t$.
To elucidate the asymptotics, we assume that the joint vector $(\boldsymbol{\epsilon}'_t, \boldsymbol{X}'_t)'$ is generated by

$$\begin{pmatrix} \boldsymbol{\epsilon}_t \\ \boldsymbol{X}_t \end{pmatrix} = \sum_{j=0}^{\infty} \boldsymbol{G}(j)\boldsymbol{\Gamma}(t-j) = \boldsymbol{A}_t \quad (say),$$

where $\{\boldsymbol{\Gamma}(t) = (\Gamma_1, ..., \Gamma_{n+r})'\}$ is an uncorrelated $(n+r)$-vector process with

$$E\left[\boldsymbol{\Gamma}(t)\right] = 0$$
$$E\left[\boldsymbol{\Gamma}(t)\boldsymbol{\Gamma}(s)^*\right] = \delta(t,s)\boldsymbol{K}$$

and $\boldsymbol{G}(j)$'s are $(n+r)\times(n+r)$ matrices which satisfy $\sum_{j=0}^{\infty} tr\left\{\boldsymbol{G}(j)\boldsymbol{K}\boldsymbol{G}(j)^*\right\} < \infty$. Denote the spectral density matrix of $\{\boldsymbol{A}_t\}$ by $\boldsymbol{f}(\omega)$.
Here the process $(\boldsymbol{\epsilon}'_t, \boldsymbol{X}'_t)'$ is possibly long-memory. For this type of linear processes, Hosoya(1997) developed the asymptotic theory for the Whittle estimator. Using his result, we can derive the asymptotic distribution of $\hat{\boldsymbol{B}}_{2SLS}$ as follows.

Theorem

Under appropriate regularity conditions, it holds that
(i)

$$p\text{-lim } \hat{\boldsymbol{B}}_{2SLS} = \boldsymbol{B}.$$

(ii)

$$\sqrt{T}\left(\hat{\boldsymbol{B}}_{2SLS} - \boldsymbol{B}\right) \xrightarrow{d} \boldsymbol{Q}^{-1}E\left[\boldsymbol{Z}_{mt}\boldsymbol{X}'_t\right]E\left[\boldsymbol{X}_t\boldsymbol{X}'_t\right]^{-1}\boldsymbol{U}$$

where

$$\boldsymbol{Q} = \left[E(\boldsymbol{Z}_{mt}\boldsymbol{X}'_t)\right]\left[E(\boldsymbol{X}_t\boldsymbol{X}'_t)\right]^{-1}\left[E(\boldsymbol{X_t}\boldsymbol{Z}'_{mt})\right]$$

and $\boldsymbol{U} = \{U_{i,j}; 1 \le i \le r, 1 \le j \le n\}$ is a random matrix whose elements follow the maltivariate normal distribution with mean $\boldsymbol{0}$ and

$$\begin{aligned} Cov[U_{i,j}, U_{k,l}] = 2\pi \int_{-\pi}^{\pi} &[f_{n+i,n+k}(\omega)\bar{f}_{j,l}(\omega) + f_{n+i,l}(\omega)\bar{f}_{j,n+k}(\omega)]d\omega \\ + 2\pi \sum_{\beta_1,\cdots,\beta_4=1}^{n+r} &\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \kappa_{n+i,\beta_1}(\omega_1)\kappa_{j,\beta_2}(-\omega_1) \\ &\times \kappa_{n+k,\beta_3}(\omega_2)\kappa_{l,\beta_4}(-\omega_2)Q^{\Gamma}_{\beta_1,\cdots,\beta_4}(\omega_1,-\omega_2,\omega_2)d\omega_1 d\omega_2. \end{aligned}$$

Statistical Estimation of Multiperiod Optimal Portfolios

Hiroshi Shiraishi *Jikei Medical University*

The original literature on dynamic portfolio choice, pioneered by Merton (1969) in continuous time and by Samuelson (1969) and Fama (1970) in discrete time, produced many important insights into the properties of optimal portfolio policies. Unfortunately, since it is known that the closed-form solutions are obtained only for a few special cases, the recent literature uses a variety of numerical and approximate solution methods to incorporate realistic features into the dynamic portfolio problem such as Ait-Sahalia and Brandt (2001) and Brandt et al. (2005).

We introduce an procedure to construct the dynamic portfolio weights based on AR bootstrap. Bose (1988) showed that the distribution of least-squares estimators in autoregressions can be resampled with accuracy $o(n^{-1/2})$ a.s..

Suppose the existence of a finite number of risky assets indexed by $i, (i = 1, \ldots, m)$. Let $\boldsymbol{X}_t = (X_1(t), \ldots, X_m(t))'$ denote the random excess returns on m assets from time t to $t+1$. Suppose too that there exists a risk-free asset with the excess return X_f. Based on the process $\{\boldsymbol{X}_t\}_{t=1}^T$ and X_f, we consider an investment strategy from time 0 to time T where $T(\in \mathbb{Z})$ denotes the end of the investment time. Let $\boldsymbol{w}_t = (w_1(t), \ldots, w_m(t))'$ be vectors of portfolio weight for the risky assets at the beginning of time $t+1$. Here we assume that the portfolio weights $\boldsymbol{w}_t$ can be rebalanced at the beginning of time $t+1$ and measurable (predictable) with respect to the past information $\mathcal{F}_t$. Then the return of the portfolio from time t to $t+1$ is written as $1 + X_f + \boldsymbol{w}_t'(\boldsymbol{X}_{t+1} - X_f \boldsymbol{e})$ and the return from time 0 to time T (called terminal wealth) is written as

$$W_T := \prod_{t=0}^{T-1} (1 + X_f + \boldsymbol{w}_t'(\boldsymbol{X}_{t+1} - X_f \boldsymbol{e})). \tag{0.1}$$

Suppose that a utility function $U : x \mapsto U(x)$ is differentiable, concave and strictly increasing for each $x \in \mathbb{R}$. Consider an investor's problem

$$\max_{\{\boldsymbol{w}_t\}_{t=0}^{T-1}} E[U(W_T)].$$

Following a formulation by the dynamic programming (e.g.,Bellman (2010)), it is convenient to express the expected terminal wealth in terms of a value function V_t which varies according to $\mathcal{F}_t = \sigma(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_t, \boldsymbol{w}_0, \ldots, \boldsymbol{w}_{t-1})$:

$$V_t = \max_{\{\boldsymbol{w}_s\}_{s=t}^{T-1}} E[U(W_T)|\mathcal{F}_t] = \max_{\boldsymbol{w}_t} E\left[\max_{\{\boldsymbol{w}_s\}_{s=t+1}^{T-1}} E[U(W_T)|\mathcal{F}_{t+1}] \middle| \mathcal{F}_t \right] = \max_{\boldsymbol{w}_t} E\left[V_{t+1}|\mathcal{F}_t\right] \tag{0.2}$$

subject to the terminal condition $V_T = U(W_T)$. The recursive equation (0.2) is the so-called Bellman equation and is the basis for any recursive solution of the dynamic portfolio choice problem. According to the literature (e.g,Brandt et al. (2005)), we can simplify this problem in case of a constant relative risk aversion (CRRA) utility function, that is,

$$U(W) = \frac{W^{1-\gamma}}{1-\gamma}, \quad \gamma \neq 1 \tag{0.3}$$

where γ denotes the coefficient of relative risk aversion. In this case, the Bellman equation simplifies to

$$V_t = U(W_t)\Psi_t, \quad \Psi_t = \max_{\boldsymbol{w}_t} E\left[(1 + X_f + \boldsymbol{w}_t'(\boldsymbol{X}_{t+1} - X_f \boldsymbol{e}))^{1-\gamma} \Psi_{t+1}|\mathcal{F}_t\right]$$

subject to the terminal condition $\Psi_T = 1$.

Suppose that $\{\boldsymbol{X}_t = (X_1(t), \ldots, X_m(t))'; t \in \mathbb{Z}\}$ is an m-vector AR(1) process defined by

$$\boldsymbol{X}_t = \boldsymbol{\mu} + A(\boldsymbol{X}_{t-1} - \boldsymbol{\mu}) + \boldsymbol{\epsilon}_t \tag{0.4}$$

where $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_m)'$ is a constant m dimensional vector, $\boldsymbol{\epsilon}_t = (\epsilon_1(t), \ldots, \epsilon_m(t))'$ are independent and identically distributed (i.i.d.) random m dimensional vectors with $E\boldsymbol{\epsilon}_t = \boldsymbol{0}$ and $E\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t' = \Gamma$ (Γ is a nonsingular m by m matrix), and A is a nonsingular m by m matrix. We construct an estimator of the optimal portfolio weight as follows;

Step1: First, we fix the current time t which implies that the observed stretch $n+t$ is fixed. Then, we can generate $\{\boldsymbol{X}_s^{(b_1,b_2,t)*}\}_{s=t+1}^T$ which is a bootstrapped sample based on $\{\boldsymbol{X}_s\}_{s=-n+1}^t$.

Step2: Next, for each $b_0 = 1,\ldots,B$, we obtain $\hat{\boldsymbol{w}}_{T-1}^{(b_0,t)}$ as the maximizer of

$$E^*_{T-1}\left[(1+X_f+\boldsymbol{w}'(\boldsymbol{X}_T^{(b_0,b,t)*}-X_f\boldsymbol{e}))^{1-\gamma}\right] = \frac{1}{B}\sum_{b=1}^{B}(1+X_f+\boldsymbol{w}'(\boldsymbol{X}_T^{(b_0,b,t)*}-X_f\boldsymbol{e}))^{1-\gamma}. \tag{0.5}$$

This $\hat{\boldsymbol{w}}_{T-1}^{(b_0,t)}$ corresponds to the estimator of myopic (single period) optimal portfolio weight.

Step3: Next, we construct estimators of Ψ_{T-1}. Since it is difficult to express the explicit form of Ψ_{T-1}, we parameterize it as linear functions of $\boldsymbol{X}_{T-1}$ as follows;

$$\Psi^{(1)}(\boldsymbol{X}_{T-1},\boldsymbol{\theta}_{T-1}) := [1,\boldsymbol{X}'_{T-1}]\boldsymbol{\theta}_{T-1} \tag{0.6}$$

$$\Psi^{(2)}(\boldsymbol{X}_{T-1},\boldsymbol{\theta}_{T-1}) := [1,\boldsymbol{X}'_{T-1}, vech(\boldsymbol{X}_{T-1}\boldsymbol{X}'_{T-1})']\boldsymbol{\theta}_{T-1}. \tag{0.7}$$

Note that the dimensions of $\boldsymbol{\theta}_{T-1}$ in $\Psi^{(1)}$ and $\Psi^{(2)}$ are $m+1$ and $m(m+1)/2+m+1$, respectively. The idea of $\Psi^{(1)}$ and $\Psi^{(2)}$ is inspired by the parameterization of the conditional expectations in Brandt et al. (2005).

In order to construct the estimators of $\Psi^{(i)}(i=1,2)$, we introduce the conditional least squares estimators of the parameter $\boldsymbol{\theta}_{T-1}^{(i)}$, that is,

$$\hat{\boldsymbol{\theta}}_{T-1}^{(i)} = \arg\min_{\boldsymbol{\theta}} Q_{T-1}^{(i)}(\boldsymbol{\theta}), \quad Q_{T-1}^{(i)}(\boldsymbol{\theta}) = E^*_{T-1}\left[(\Psi_{T-1}-\Psi^{(i)})^2\right].$$

Then, by using $\hat{\boldsymbol{\theta}}_{T-1}^{(i)}$, we can compute $\Psi^{(i)}(\boldsymbol{X}_{T-1}^{(b_0,b,t)*},\hat{\boldsymbol{\theta}}_{T-1}^{(i)})$.

Step4: Based on the above $\Psi^{(i)}$, we obtain $\hat{\boldsymbol{w}}_{T-2}^{(b_0,t)}$ as the maxmizer of

$$E^*_{T-2}\left[(1+X_f+\boldsymbol{w}'(\boldsymbol{X}_{T-1}^{(b_0,b,t)*}-X_f\boldsymbol{e}))^{1-\gamma}\Psi^{(i)}(\boldsymbol{X}_{T-1}^{(b_0,b,t)*},\hat{\boldsymbol{\theta}}_{T-1}^{(i)})\right]. \tag{0.8}$$

This $\hat{\boldsymbol{w}}_{T-2}^{(b_0,t)}$ does not correspond to the estimator of myopic (single period) optimal portfolio weight due to the effect of $\Psi^{(i)}$.

Step5: In the same manner of Step3-4, we can obtain $\hat{\boldsymbol{\theta}}_s^{(i)}$ and $\hat{\boldsymbol{w}}_s^{(b_0,t)}$, recursively, for $s=T-2,T-1,\ldots,t+1$.

Step6: Then, we define an optimal portfolio weight estimator at time t as $\hat{\boldsymbol{w}}_t^{(t)} := \hat{\boldsymbol{w}}_t^{(b_0,t)}$ by Step4. Note that $\hat{\boldsymbol{w}}_t^{(t)}$ is obtained as only one solution because $\boldsymbol{X}_{t+1}^{(b_0,b,t)*}(=\hat{\boldsymbol{\mu}}^{(t)}+\hat{A}^{(t)}(\boldsymbol{X}_t-\hat{\boldsymbol{\mu}}^{(t)})+\boldsymbol{\epsilon}_{t+1}^{(b,t)*})$ is independent of b_0.

Step7: For each time $t=0,1,\ldots,T-1$, we obtain $\hat{\boldsymbol{w}}_t^{(t)}$ by Step1-6. Finally, we can construct an optimal investment strategy as $\{\hat{\boldsymbol{w}}_t^{(t)}\}_{t=0}^{T-1}$.

References

Ait-Sahalia, Y. and Brandt, M. (2001). Variable selection for portfolio choice. *Journal of Finance*, 56(4):1297–1351.

Bellman, R. (2010). *Dynamic Programming (Princeton Landmarks in Mathematics)*. Princeton Univ Pr.

Bose, A. (1988). Edgeworth correction by bootstrap in autoregressions. *Ann. Statist.*, 16(4):1709–1722.

Brandt, M. W., Goyal, A., Santa-Clara, P., and Stroud, J. R. (2005). A simulation approach to dynamic portfolio choice with an application to learning about return predictability. *Review of Financial Studies*, 18(3):831–873.

Fama, E. F. (1970). Multiperiod consumption-investment decisions. *American Economic Review*, 60(1):163–74.

Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *The Review of Economics and Statistics*, 51(3):247–57.

Samuelson, P. A. (1969). Lifetime portfolio selection by dynamic stochastic programming. *Review of Economics and Statistics*, 51(3):239–46.

When does seasonality exist? A Synchrosqueezing approach to determining the seasonality.

Ming-Yen Cheng (National Taiwan University)

Abstract:

Seasonality exists in a given observed signal if the signal expresses repeated patterns regularly as time proceeds. Identifying seasonality is an important issue in many scientific fields. In general, if seasonality is observed or suspected, ARIMA model, frequency analysis (FFT) and seasonal index are applied to analyze, confirm and predict its effect. Besides their success in understanding many important information, such methods are limited to either model simplification or a priori model selection. In particular, the simplified assumptions might mask the nonlinearity or nonstationarity nature of the underlying system. Consequently, these approaches are not always feasible to analyze the real world data. For example, when we want to understand finer local information in the data, such as the effect of an unexpected disease breakout or the effect of a policy change, it seems difficult for the traditional methods to achieve this aim. It turns out that establishing an adaptive method, that is, one without any assumption on the model, is beneficial in this field. Upper respiratory tract infection (URI) is an illness caused by an acute infection involving the upper respiratory tract. From the viewpoint of public health, it has the feature referred to as the seasonality, that is, the prevalence oscillates according to the seasonal change ‡ the prevalence is higher in the winter and lower in the summer. This phenomena has been studied and well understood. The purpose of this study is to introduce an adaptive method, referred to as the Synchrosqueezing transform, to extract information from a given time varying signal so that the extracted information may be used as the initial guess of the model selection problem. Since the existence of the seasonality phenomena in the URI disease is commonly accepted, and policy changes and ages play roles in this disease, we focus on analyzing this disease and demonstrate how this method extracts the hidden information. In the end, we apply the extracted information to build up a new ARIMA model to understand statistically the URI dataset. This talk is based on a joint work with Jzeng-Ji Chen, Yu-Chun Chen, and Hau-Tieng Wu.

複数データセットによる非ガウス構造方程式モデルの推定

清水　昌平
大阪大学　産業科学研究所

1　はじめに

構造方程式モデル [1] は, データ生成過程のモデルとして使うことができる．重要な応用には, 因果推論がある [2]．従来は, 線形性とガウス分布の仮定が基本であったが, データ生成過程の構造に関する背景知識がない場合に, 識別できるモデルが少ないという問題があった．そのため最近は, ガウス分布の代わりに非ガウス分布を仮定するモデルが盛んに研究されるようになってきている [3,4]．データの非ガウス性を利用することで, これまで識別できなかったモデルの多くが識別可能になる．データの非ガウス性の利用という点で, 独立成分分析 [5] と強く関連している．多くの分野においてガウス分布では上手く近似できないようなデータがあり, 経済・生体科学などにおける実際の適用例も増えてきている [6–8]．本発表では, そのような非ガウス構造方程式モデルについて議論する．特に, 複数のデータセットを融合し推定精度を向上させることを考える．

2　モデル

複数グループのデータ生成過程を次のようにモデリングする：

$$x_i^{(g)} \;=\; \sum_{k(j)<k(i)} b_{ij}^{(g)} x_j^{(g)} + e_i^{(g)} \quad (g = 1, \cdots, c), \tag{1}$$

ここで g はグループの添え字, c はグループ数, $x_i^{(g)}$, $e_i^{(g)}$ と $b_{ij}^{(g)}$ はグループ g の観測変数, (未観測の) 外的影響変数, 変数間の結合の強さを表す．外的影響変数 $e_i^{(g)}$ は連続変数で, 平均ゼロで分散が非ゼロの非ガウス分布に従い, 互いに独立とする．そして、$k(i)$ は各グループ g において観測変数が有向非巡回グラフを成すような変数順序を表す．変数順序はグループ間で共通と仮定する．このモデルは識別可能であり [3,9], 推定には直接法が利用可能である [9,10]．グループ数 $c = 1$ の場合は, LiNGAM モデル [3] になる．共通の変数順序というグループ間の類似性を利用して推定精度を上げたい.

3　応用

このモデルを用いて, [11] のタンパク質発現量データを分析した [12]．これらは 9 つの異なる外的刺激条件で測定されたデータで、サンプルサイズはそれぞれ, 853, 902, 911, 723, 810, 799, 848, 913, 707 であった．測定されたのは Raf パスウェイのタンパク質で, Raf, Mek1/2, Plc-γ, PIP2, PIP3, Erk (p44/42), Akt, PKA, PKC, P38 と Jnk の 11 種類であった．共通の変数順序という仮定は, この例では, わるくはなかった [11].

90%以上の外的刺激条件で共通して推定された有向辺が図 1 である．7 つの有向辺のうち, 5 つが実質科学の知見に合っていた．残りの 2 つは, 逆向きだった．また, 90 %以上の条

件で, 94 の有向辺がないものとして推定された．そのうち 83 は背景知識に合っていた．次に, 複数データセット用のベイジアンネットワークを用いる最新手法である iMAGES [13] を適用し, 推定された非巡回グラフが図 2 である．従来法の iMAGES では, いくつかの無向辺が推定されたが, 向きまでは推定できなかった．

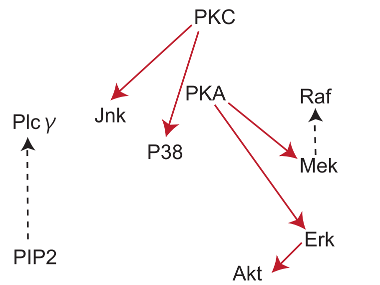

図 1: 非ガウス構造方程式モデルを用いて推定された有向辺．実線は背景知識に合う有向辺を表す．

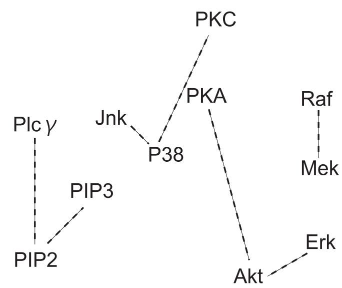

図 2: iMAGES によって推定された非巡回グラフ．

参考文献

[1] K. Bollen. *Structural Equations with Latent Variables.* John Wiley & Sons, 1989.

[2] J. Pearl. *Causality: Models, Reasoning, and Inference.* Cambridge University Press, 2000. (2nd ed. 2009).

[3] S. Shimizu, P. O. Hoyer, A. Hyvärinen, and A. Kerminen. A linear non-gaussian acyclic model for causal discovery. *Journal of Machine Learning Research*, 7:2003–2030, 2006.

[4] P. Spirtes, C. Glymour, R. Scheines, and R. Tillman. Automated search for causal relations: Theory and practice. In R. Dechter, H. Geffner, and J. Halpern, editors, *Heuristics, Probability, and Causality: A Tribute to Judea Pearl*, pages 467–506. College Publications, 2010.

[5] A. Hyvärinen, J. Karhunen, and E. Oja. *Independent component analysis.* Wiley, New York, 2001.

[6] L. Faes, S. Erla, E. Tranquillini, D. Orrico, and G. Nollo. An identifiable model to assess frequency-domain Granger causality in the presence of significant instantaneous interactions. In *Proceedings of the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS2010)*, pages 1699–1702, 2010.

[7] S. M. Smith, K. L. Miller, G. Salimi-Khorshidi, M. Webster, C. F. Beckmann, T. E. Nichols, J. D. Ramsey, and M. W. Woolrich. Network modelling methods for FMRI. *NeuroImage*, 54(2):875–891, 2011.

[8] E. Ferkingsta, A. Lølanda, and M. Wilhelmsen. Causal modeling and inference for electricity markets. *Energy Economics*, 33(3):404–412, 2011.

[9] S. Shimizu. Joint estimation of linear non-Gaussian acyclic models. *Neurocomputing*, xx:xx–xx, 2012. In press.

[10] S. Shimizu, T. Inazumi, Y. Sogawa, A. Hyvärinen, Y. Kawahara, T. Washio, P. O. Hoyer, and K. Bollen. DirectLiNGAM: A direct method for learning a linear non-Gaussian structural equation model. *Journal of Machine Learning Research*, 12:1225–1248, April 2011.

[11] K. Sachs, O. Perez, D. Pe'er, D. A. Lauffenburger, and G. P. Nolan. Causal protein-signaling networks derived from multiparameter single-cell data. *Science*, 308(5721):523–529, April 2005.

[12] S. Shimizu, T. Shimamura, S. Hara, T. Washio, and S. Imoto. Finding commonly estimated causal relations across different experimental conditions for better reliability. 2012. In preparation.

[13] J. D. Ramsey, S. J. Hanson, C. Hanson, Y. O. Halchenko, R. A. Poldrack, and C. Glymour. Six problems for causal inference from fMRI. *NeuroImage*, 49(2):1545–1558, 2010.

A Simple Test of Long Memory Versus Structural Breaks*

Katsumi Shimotsu
Department of Economics
Hitotsubashi University
shimotsu@econ.hit-u.ac.jp

Keiko Yamaguchi
Graduate school of Commerce
Waseda University
kyamaguchi@aoni.waseda.jp

January 29, 2012

Abstract

This paper proposes a simple test that is based on certain time domain properties of $I(d)$ processes. If a time series follows an $I(d)$ process, then its dth differenced series follows an $I(0)$ process. Simple as it may sound, its properties provides useful tools to distinguish the true and spurious $I(d)$ processes. We estimate d, use the estimate to take the d th difference of the sample, and apply the KPSS test to the differenced data and its partial sum. The KPSS test is applicable to both stationary and nonstationary $I(d)$ processes. The spurious long memory processes are essentially $I(0)$ or $I(1)$ in their nature, and taking their dth difference magnifies their non-$I(d)$ properties. We derive its limiting distribution and show that the test is consistent. The limiting distribution of these test statistics depends on d, and its simulated critical values are provided. Simulations show that the proposed tests have good power against the spurious long memory models considered in the literature.

JEL Classification Number: C12, C13, C14, C22
Keywords: long memory; fractional integration; structural breaks.

*The author thanks Torben Andersen, Rohit Deo, Christian Gourieroux, Alex Maynard, Nour Meddahi, Peter Phillips, and the participants at the CIRANO-CIREQ Conference on Time Series, 2006 International Symposium on Econometric Theory and Applications (SETA), and the seminar at Yokohama National University for helpful comments on an earlier version. The authors thank Qian Liu for research assistance. Financial support from SSHRC is gratefully acknowledged. This paper is a part of the outcome of research performed under a Waseda University Grant for Special Research Projects (Project number: 2011A-868).

The optimal portfolio problem with the ideal balance and exogenous variable causality

Kiyoaki Sasaki and Junichi Hirukawa*
Niigata University

abstract

The pension fund collect the premiums from the people and invest them. The premium reserve is a minimum asset which should be reserved in the investment. On the other hand the portfolio is a linear combination of the prices of assets, which are multiplied by allocations. In this paper, we assume that the premium reserve and portfolio are described by Fackler's recurrence formula. The cash flow of total of premium and sum insured can be decomposed into two part, namely, that of the premium reserve and that of portfolio. Our object is the ideal balance of premium reserve, which is given by the pension investment policy and expressed by the ideal curve function. We give the optimal portfolio which solve this recursive utility optimization problem. Finally, we consider the additional restriction of the causality from exogenous variable.

Let us assume that a total of premium at year t is C_t and sum insured is R_t. The cash flow constraint at year t is

$$CF_t = R_t - C_t,$$

which is cash flow total of premium and sum insured. Portfolio have price p_t and the allocation vector is given by α_t at year t. We assume that i is interest rate, r_t is return rate of portfolio, V_t is premium reserve and A_t is total of portfolio price. From Fackler's recurrence formula we have

$$V_{t+1} = V_t\left(1+i\right) + CF_t^1\left(1+\frac{i}{2}\right),$$

$$A_{t+1} = A_t\left(1+r_t\right) + CF_t^2\left(1+\frac{r_t}{2}\right).$$

The cash flow constraint is

$$CF_t = CF_t^1 + CF_t^2.$$

Solving Fackler's recurrence formula with respect to V_t:

$$V_t = V_0\left(1+i\right)^t + \sum_{k=0}^{t-1}\left(1+i\right)^{t-1-k} CF_k^1\left(1+\frac{i}{2}\right).$$

Since A_t is total of portfolio price, CF_t^2 is represented by p_t and α_t:

$$CF_t^2 = \frac{2\alpha_t' p_t \left(\alpha_{t+1}' - \alpha_t'\right) p_{t+1}}{\alpha_t' \left(p_t + p_{t+1}\right)}.$$

We consider model of p_t whether it contains exogenous variable or not.

- If p_t does not contain the exogenous variable, p_t is expressed by simple VAR model:

  $$p_t = Bp_{t-1} + w_t,$$

  where the error term w_t is weak white noise and A is an $n \times n$ matrix.

- If p_t contains the exogenous variable, we assume S_t is exogenous and

  $$\begin{aligned} Cp_t &= Dp_{t-1} + F_0 S_t + F_1 S_{t-1} + u_t, \\ S_t &= HS_{t-1} + v_t. \end{aligned}$$

Therefore,

$$\begin{pmatrix} p_t \\ S_t \end{pmatrix} = \begin{pmatrix} C^{-1}D & C^{-1}F_1 + C^{-1}F_0 H \\ 0 & H \end{pmatrix} \begin{pmatrix} p_{t-1} \\ S_{t-1} \end{pmatrix} + \begin{pmatrix} C^{-1}\left(u_t + F_0 v_t\right) \\ v_t \end{pmatrix}$$

This optimization problems are solved with minimizing the distance between the premium reserve and the objective function in terms of the loss function. Let us assume objective function l_t, is the ideal balance function. This function is the ideal balance curve of overall future pension given by pension investment policy. The minimization problem is

$$\min_{\alpha} E_t \left[\sum_{j=0}^{\infty} \delta^j U\left(l_{t+j} - V_{t+j}\right) \right],$$

$$\text{subject to} : CF_t = CF_t^1 + CF_t^2,$$

where δ , $0 \leq \delta \leq 1$ is subjective discount factor and the loss function U is increasing and convex. Let us denote

$$q_{t,t} = \frac{\partial CF_t^2}{\partial \alpha_t},\ q_{t,t-1} = \frac{\partial CF_{t-1}^2}{\partial \alpha_t},\ l_{t,t+j} = \frac{\partial V_{t+j}}{\partial \alpha_t}.$$

The solution of this problem can be obtained by differentiating with respect to the portfolio allocations. We can find the first-order condition called Euler condition:

$$U'\left(l_t - V_t\right)\left(-q_{t,t-1}\right) + E_t \left[\sum_{j=1}^{\infty} \delta^j U'\left(l_{t+j} - V_{t+j}\right) l_{t,t+j} \right] = 0.$$

where **1** is the vector whose components are all 1.

Smooth Transition Quantile Capital Asset Pricing Models with Heteroscedasticity

Cathy W. S. Chen(1)*, Simon Lin(1), and Philip L.H. Yu(2)
(1) Feng Chia University, Taiwan
(2) The University of Hong Kong

Abstract

Capital asset pricing model (CAPM) has become a fundamental tool in finance for assessing the cost of capital, risk management, portfolio diversification and other financial assets. It is generally believed that the market risks of the assets, often denoted by a beta coefficient, should change over time. In this paper, we model time-varying market betas in CAPM by a smooth transition regime switching CAPM with heteroscedasticity, which provides flexible nonlinear representation of market betas as well as flexible asymmetry and clustering in volatility. We also employ the quantile regression to investigate the nonlinear behavior in the market betas and volatility under various market conditions represented by different quantile levels. Parameter estimation is done by a Bayesian approach. Finally, we analyze some Dow Jones Industrial stocks to demonstrate our proposed models. The model selection method shows that the proposed smooth transition quantile CAPM-GARCH model is strongly preferred over a sharp threshold transition and a symmetric CAPM-GARCH model.

周期的スプライン平滑法と変化する周期的変動の解析

Cyclic cubic regression spline smoothing and analysis of slowly changing cyclic variation

慶應義塾大学 理工学部　　南　美穂子

周期的３次回帰スプライン平滑法（Cyclic cubic regression splines smoothing）は両端点の２次微分までの値が等しいという制約を加えた３次回帰スプライン平滑法である。本研究では、周期的３次回帰スプライン平滑法の数理的特徴を調べ、この特徴を利用して周期成分を柔軟に、かつ、より少ない基底関数で表現してテンソルプロダクト法を用いることにより、変化する周期変動を解析する方法を提案する。大気中の PM2.5 濃度データを例に特徴を示し、CO2 濃度データの季節によって変化する日次変動を解析した。

周期的３次回帰スプライン平滑法

周期的３次スプライン関数は、区分的３次多項式であり、節点を $t^{(0)}, t^{(1)} \dots,\ t^{(k)}$ とし、区間 $[t^{(j-1)}, t^{(j)}]$, $j = 1, 2, \dots, k,$ 上の３次多項式を $S_j(t)$ としたとき、節点 $t^{(j)}$ $(j = 1, \dots, k-1)$ において２次微分まで連続という通常の３次スプライン関数の満たすべき制約に加えて、両端点 $t^{(0)}, t^{(k)}$ での値が２次微分まで等しい、つまり、$S_1(t^{(0)}) = S_k(t^{(k)}),\ S_1'(t^{(0)}) = S_k'(t^{(k)}),\ S_1''(t^{(0)}) = S_k''(t^{(k)})$ という制約をも満たすものである。このとき周期は k である。周期的３次回帰スプライン平滑法 (cf. Wood, 2006) は、観測点 $t_1, \cdots, t_n$ における観測値 $y_1, \cdots, y_n$ が得られたとき、平滑パラメータを $\lambda(\lambda \geq 0)$ とする罰則付きの残差２乗和

$$\sum_{i=1}^{n} \{y_i - g(t_i)\}^2 + \lambda \int_{t^{(0)}}^{t^{(k)}} g''(t)^2 dx \tag{1}$$

を最小にする周期的３次スプライン関数 $g(t) = \sum_{j=1}^{k} I_{[t^{(j-1)},\ t^{(j)}]}(t) S_j(t)$ を平滑曲線とする平滑手法である。

節点での関数値をパラメータとする表現

周期的３次スプライン関数 $g(t)$ の節点での関数値と２次微分値を $\beta_j = g(t^{(j)}),\ \delta_j = g''(t^{(j)})$ とおくと

$$g(t) = \begin{cases} a_j^-(t)\beta_{j-1} + a_j^+(t)\beta_j + c_j^-(t)\delta_{j-1} + c_j^+(t)\delta_j & \text{if } t^{(j-1)} \leq t \leq t^{(j)},\ j = 2, ..., k \\ a_1^-(t)\beta_k + a_1^+(t)\beta_1 + c_1^-(t)\delta_k + c_1^+(t)\delta_1 & \text{if } t^{(0)} \leq t \leq t^{(1)} \end{cases} \tag{2}$$

と表現できる。ここで、

$$\begin{aligned} a_j^-(t) &= \frac{t^{(j)} - t}{h_j}, \quad a_j^+(t) = \frac{t - t^{(j-1)}}{h_j}, \quad h_j = t^{(j)} - t^{(j-1)} \\ c_j^-(t) &= \frac{1}{6}\left[\frac{(t^{(j)} - t)^3}{h_j} - h_j(t^{(j)} - t)\right], \quad c_j^+(t) = \frac{1}{6}\left[\frac{(t - t^{(j-1)})^3}{h_j} - h_j(t - t^{(j-1)})\right]. \end{aligned}$$

節点で２次微分まで連続で、両端点での２次微分までの値が等しいという条件は、関数値ベクトルを $\boldsymbol{\beta} = (\beta_1,\ \cdots,\ \beta_k)^T$, ２次微分値ベクトルを $\boldsymbol{\delta} = (\delta_1,\ \cdots,\ \delta_k)^T$ としたとき、$B\boldsymbol{\delta} = D\boldsymbol{\beta}$ と同値であることが示せる。ここで行列 B, D は大きさ k の正方行列であり、

$$B_{i,i} = \frac{1}{3}(h_i + h_{i+1}),\ B_{i,\ i+1} = B_{i+1,\ i} = \frac{1}{6}h_{i+1}\ (i = 1, \cdots, k-1),\ B_{k,k} = \frac{1}{3}(h_k + h_1),\ B_{1,k} = B_{k,1} = \frac{1}{6}h_1$$

$$D_{i,i} = -\frac{1}{h_i} - \frac{1}{h_{i+1}},\ D_{i,\ i+1} = D_{i+1,\ i} = \frac{1}{h_{i+1}}\ (i = 1, \cdots, k-1),\ D_{k,k} = -\frac{1}{h_k} - \frac{1}{h_1},\ D_{1,\ k} = D_{k,\ 1} = \frac{1}{h_1}$$

で、その他はすべて０の行列である。$\boldsymbol{\delta} = B^{-1}D\boldsymbol{\beta}$ として (2) に代入することにより節点での関数値をパラメータとする基底表現 $g(t) = \sum_{i=1}^{k} b_i(t)\beta_i$ を得る。また、罰則項は $\int_{t^{(1)}}^{t^{(k+1)}} g''(t)^2 dx = \boldsymbol{\beta}^T D^T B^{-1} D \boldsymbol{\beta}$ と表わすことができる。

節点が等間隔の場合の行列

節点間の距離がすべて等しく h であり、観測はすべて節点上で、各節点での観測数が m の場合を考える。観測数は $n = m \cdot k$ である。各節点での標本平均ベクトルを $\overline{\boldsymbol{y}}$ で表す。(3) 式の最小化は $S(\boldsymbol{\beta}) = m\|\overline{\boldsymbol{y}} - \boldsymbol{\beta}\|^2 + \lambda\boldsymbol{\beta}^T D^T B^{-1} D\boldsymbol{\beta}$ の最小化と同値になり、$\hat{\boldsymbol{\beta}} = \left(I_k + \frac{\lambda}{m} D^T B^{-1} D\right)^{-1} \overline{\boldsymbol{y}} \, (\equiv H\overline{\boldsymbol{y}})$ を得る。さて、$G(a, b)$ を対角成分が a、その上下の非対角成分と $(1, k)$ 成分、$(k, 1)$ 成分が b の周期的 3 重対角行列（巡回行列）とすると、行列 B と C は $B = \dfrac{h}{6}\, G(4, 1),\ D = \dfrac{1}{h}\, G(-2, 1)$ と表せる。 $G(a, b)$ の固有値は k が偶数 $(= 2q)$ の場合, 降順に並べると $l_1 = a + 2b,\ l_{2j} = l_{2j+1} = a + 2b\cos\frac{2\pi j}{k}(j = 1, \cdots, q-1),\ l_{2q} = a - 2b$ であり、対応する基準化固有ベクトルは、$\boldsymbol{u}_1$ は定数ベクトル、$\boldsymbol{u}_{2j}, \boldsymbol{u}_{2j+1}$ は周期 k/j の $\sin, \cos$ 関数の節点での値からなるベクトル、$\boldsymbol{u}_{2q}$ は 1 と -1 が交互に値を取るベクトルを、それぞれ基準化したものとなる。

行列 B, D および、I_k は固有ベクトルが共通であることから、影響行列 H の固有値は、降順に

$$\gamma_1 = 1, \quad \gamma_{2j} = \gamma_{2j+1} = \left(1 + \frac{\lambda}{m} \cdot \frac{12}{h^3} \cdot \frac{\left(1 - \cos\dfrac{2\pi j}{k}\right)^2}{\left(2 + \cos\dfrac{2\pi j}{k}\right)}\right)^{-1} \quad \gamma_{2q} = \left(1 + \frac{\lambda}{m} \cdot \frac{48}{h^3}\right)^{-1}$$

であり、固有ベクトルは周期的 3 重対角行列と同じである。推定値ベクトルは $\hat{\boldsymbol{\beta}} = \sum_{j=1}^{k} \gamma_j (\boldsymbol{u}_j, \overline{\boldsymbol{y}})\, \boldsymbol{u}_j$ と表せるが、上記より、定数成分は縮小されないが、周期が短い成分ほどより縮小されるということがわかる。Wahba(1990) は、三角関数を基底関数とした場合の周期的 3 次平滑法による係数を示しているが、罰則項の評価が異なるため、固有ベクトルは同じであるが、固有値は異なっている。

例として、福岡市で 2001 年 7 月 10 日から 8 月 6 日までの 28 日間に毎時計測された大気中の PM2.5(直径 2.5 μ m 以下の粒子状物質) 濃度データの解析結果、影響行列の固有値、固有ベクトルと成分の縮小の様子、また、固有値の小さい成分を除いて成分数を減らしても推定結果があまり変わらないことを示した。

混合効果モデルとしての解釈と平滑パラメータの推定、平滑曲線の推定と信用曲線

関数空間を罰則が課される関数の空間と罰則が課されない滑らかな関数の空間に分けることにより、周期的スプライン平滑法を混合効果モデルとみなすことができる。平滑パラメータは分散パラメータで表すことができ、分散パラメータの推定に制限付き最尤推定法 (REML) を用いると、直観的に理解しやすい推定方程式が得られる。また、関数値ベクトルの事後分布を考えることにより、平滑曲線の推定、および、信用曲線を求めることができる。

テンソルプロダクト法による変化する周期的変動の推定

テンソルプロダクト法は、単変量平滑法の基底関数の積を基底関数とする多次元平滑法である。季節で変化する日次変動の推定法として、日次変動、および、年次変動を表す周期的スプライン平滑法の基底関数の積 $b_i(t)c_j(s)$ を基底関数とするテンソルプロダクト法を考える。これはトーラス上の関数の平滑法ということになる。基底数は 2 つの単変量平滑法の基底数の積になるから、それぞれの単変量平滑法の基底数を低く抑える必要がある。そこで、影響行列の大きい固有値に対応する成分のみを基底とすることにする。

例として、南極昭和基地で計測された CO2 濃度のデータを解析した。これは毎時計測された約 26 年分のデータであるが欠測もある。日次変動の季節的な変化が推定できたが、日次変動は傾向変動や年次変動に比べるとかなり小さいものであり、信用曲線の幅も相対的に広いものとなった。

参考文献

Wahba, G. (1990), Spline Models for Observational Data, SIAM. CBMS-NSF Regional Conference Series in Applied Mathematics, v. 59.

Wood, Simon N.(2006) Generalized Additive Models, An Introduction with R, Chapman & Hall.

Wood, S.N. (2006). Low-rank scale-invariant tensor product smooths for generalized additive mixed models. Biometrics 62, 1025-1036.

Some Generalization of Universal Portfolio

Jun'ichi Takeuchi and Mariko Tsurusaki
Graduate School of Information Science and Electrical Engineering,
Kyushu University, Fukuoka, Japan.

金融におけるポートフォリオの問題は，様々な分野において研究されてきたが，Cover らは，情報理論を援用してユニバーサルポートフォリオの枠組みを提案した ([1] 等)．特に，ポートフォリオの性能を示す wealth ratio の minimax 値を導出している [2]．すなわち，stock 数 m の市場において，Constantly Rebalanced Portfolio (CRP) と呼ばれる戦略のクラスをターゲットとしたとき，minimax log wealth ratio は

$$\frac{m-1}{2}\log\frac{n}{2\pi}+\log\int\sqrt{J_B(\theta)}d\theta+o(1)$$

となる．ただし，θ は，アルファベット数 m の多項 Bernoulli モデルのパラメータ，J_B は θ の Fisher 情報量，n は取引日数を表す．また，$o(1)$ は，n が大きくなるにつれ 0 に収束する量である．その後，この値がユニバーサルデータ圧縮における多項 Bernoulli モデルに関する minimax regret[3] と全く同じであることが判明した [5]．regret は冗長度の一種であり，ユニバーサルデータ圧縮の評価基準として導入された [3]．上記の関係は，ユニバーサルポートフォリオの枠組みと，データ圧縮のそれとの強い関連性の表れである．

本発表では，この関係を見通しよく理解するために，ユニバーサルポートフォリオをユニバーサルデータ圧縮の枠組み [3] で論じる．すなわち，CRP に対応する確率過程のクラス定義し，それをターゲットとした場合の minimax regret について考察する．また，CRP に対応する確率過程は隠れ Markov モデル (HMM) とみなせることに着目し，CRP を一般化したターゲットクラスとして，Constant Markov Portfolio (CMP) を定義する．さらに，CRP と CMP が HMM であることを利用して，効率的なユニバーサルアルゴリズムを提案する．

本発表は，主に [4] に基づいている．

参考文献

[1] T. M. Cover & J. A. Thomas, *Elements of Information Theory, second edition* Wiley-Interscience, 2006.

[2] T. M. Cover & E. Ordentlich, "Universal Portfolios with Side Information." *IEEE Trans. on Inform. Theory,* 42(2):348-363, March 1996.

[3] Y. M. Shtar'kov, "Universal sequential coding of single messages," *Problems of Information Transmission*, vol. 23, pp. 3-17, July 1987.

[4] M. Tsurusaki & J. Takeuchi, "Stochastic Interpretation of Universal Portfolio and Generalized Target Classes," *Proc. of 2011 IEEE International Symposium on Information Theory,* 2011.

[5] Q. Xie & A. R. Barron, "Asymptotic minimax regret for data compression, gambling and prediction," *IEEE trans. IT* vol. 46, no. 2, pp. 431-445, 2000.

Machine learning methods for brain machine interface

Shin Ishii (Kyoto University)

Bain machine interface (BMI) is a technology to directly connect brains and computers. It recently emerges due to the progress in neuroscience, signal processing and machine learning. In this talk, I put my focus on the non-invasive BMI; that is, there is no need for surgical operation. Magneto-encephalography (MEG) becomes recently available to measure in millisecond order magnetic fields produced by neural activities in the brain. Since we try to detect brain activities with fine temporal resolution, MEG is an attractive brain measurement modality. In order to extract brain's information processing from MEG, we may need to solve a statistical inverse problem, because the forward process is disturbed by various probabilistic factors and distortions, so it has been a target of statistical inference. However, this statistical inference problem usually suffers from a serious ill-conditioned situation; the possible activity source location in the brain may have a wide variety whose possibility number may become for example 25000, while the number of measurement sensors is typically up to 400. That is, the inverse regression is completely ill-conditioned. So, some constraints like sparseness are definitely necessary. A previous study [1] presented a hierarchical Bayesian approach, in which the hierarchical prior called automatic relevant determination encourages many source components to be zero in effect. Another recent study has presented its dynamical version [2]; although the previous method [1] did not consider any dynamics in the brain activities, its dynamical version assumes the component-wise auto-regression (AR) process for the latent brain activities along time. If the AR parameters are estimated to be all zeros, the extended version reduces to the previous (no-dynamical) one. Then, this dynamical version has a temporally smoothing factor as well as the sparseness due to the Bayesian setting.

In the talk, I next introduce our network estimation method [3]. In the previous study, an individual AR process was assumed, but there was no consideration of special or functional structures of cortical activities. If we are interested in such coordination structure in the cortex, it is the issue of network estimation. Although there are many methods for network estimation in a particular interest in its static structure, there are only few for their dynamic counterparts. When we attempt to extract dynamically changing network structure from observation time-series, the ill-conditioned-ness becomes more serious, because we have to basically estimate the network parameter for every time point along the time-series. There should be more strong constraints other than the sparseness. According to our method, we introduce, in addition to the L1 norm penalty term which encourages sparseness, the nuclear norm penalty term is attached to the optimization problem. Interesting is, if the objective function is convex, the total optimization problem is also convex even with the L1 norm and nuclear norm terms, so techniques for convex optimization can be used after some newly introduced modifications. As a benchmark test, we evaluated the prediction performance of the estimated network, when applied to the dataset of US senate roll-call votes. Our estimation method for dynamically changing network structure successfully achieved a higher prediction performance, in terms of cross-validation pseudo-likelihood, than its static version or the method without the low-rank regularization. We are currently working on the combination of such a network estimation model and a forward physical model, in order to realize a more sophisticated BMI analysis tool.

References:

[1] Sato, M, Yoshioka, T., Kajihara, S., Toyama, K., Goda, N., Doya, K., & Kawato, M. Hierarchical Bayesian estimation for MEG inverse problem. *NeuroImage*, **23**, 806-826, 2004.

[2] Fukushima, M., Yamashita, O., Kanemura, A., Ishii, S., Kawato, M., & Sato, M. A state-space modeling approach for localization of focal current sources from MEG. *IEEE Transactions on Biomedical Engineering*, to appear.

[3] Hirayama, J., Hyvarinen, A., & Ishii, S. Sparse and low-rank matrix regularization for learning time-varying Markov networks.

Generalized Cordeiro–Ferrari Bartlett-type adjustment

柿沢 佳秀 (北大経済)

1. はじめに バートレット調整として知られている LR 統計量の帰無分布のカイ 2 乗近似を改良する話題は, Bartlett (1937;Proc.R.Soc.Lond.) による分散均一性の具体例が始まりであり, 一般論は, Lawley (1956;Biometrika) が $(\boldsymbol{\theta}'_{(1)}, \boldsymbol{\theta}'_{(2)})' \in \mathbf{R}^p$ に関する帰無仮説 $\boldsymbol{\theta}_{(1)} = \boldsymbol{\theta}_{(1)0}$ の LR 統計量の平均 $E^{(N)}_{\theta_{(1)0},\theta^\dagger_{(2)}}[\mathrm{LR}^{(N)}] = p_1 + (\mathcal{E}_p - \mathcal{E}_{p-p_1})/N + \mathrm{o}(N^{-1})$ を計算し, さらに $p_1 \mathrm{LR}^{(N)}/E^{(N)}_{\theta_{(1)0},\theta^\dagger_{(2)}}[\mathrm{LR}^{(N)}]$ の任意次のキュムラントは $\mathrm{o}(N^{-1})$ を無視すれば $\chi^2_{p_1}$ のキュムラントに一致することを示したことによる. 実際, これらの結果は LR 統計量の漸近展開 (Hayakawa (1977,1987 訂正;AISM))

$$P^{(N)}_{\theta_{(1)0},\theta^\dagger_{(2)}}[\mathrm{LR}^{(N)} \le x] = G_{p_1}(x) + \frac{A_1}{24N}\{G_{p_1+2}(x) - G_{p_1}(x)\} + \mathrm{o}(N^{-1})$$

に整合する.

$\mathrm{LR}^{(N)}$ がバートレット調整可能である, すなわち, $\mathrm{BLR}^{(N)} = \mathrm{LR}^{(N)}/\{1+A_1/(12p_1N)\}$, $\{1-A_1/(12p_1N)\}\mathrm{LR}^{(N)}$ の漸近展開の N^{-1} 項が消失する; $P^{(N)}_{\theta_{(1)0},\theta^\dagger_{(2)}}[\mathrm{BLR}^{(N)} \le x] = G_{p_1}(x) + \mathrm{o}(N^{-1})$ という性質は, Rao/Wald 統計量などで一般に成立せず, それら統計量に対する異なる調整法が 1991 年の 3 論文 [Chandra and Mukerjee;JMVA, Cordeiro and Ferrari;Biometrika, Taniguchi;JMVA] で論じられた [今日にバートレット型調整として知られるようになった原型である]. CF 法は複合帰無仮説でも適用でき, 過去 20 年間に数値実験を含めて先行研究が多数あるものの, CM/T 法は報告者の知る限り, Mukerjee (1992;SPL) 以外には複合帰無仮説を扱っていない [文献の多くはスカラー母数の単純帰無仮説であった] から, 報告者は 11 月 30 日の京都シンポジウムで CM/T 法を統一的に再検討してそれらを含む一般化バートレット型調整 [Kakizawa(2012;JMVA)] を複数導入し, 局所対立仮説の下での検出力関数の漸近展開を比較した.

本報告では Cordeiro–Ferrari バートレット型調整の一般化を提示して, 局所対立仮説の下での検出力関数の漸近展開も導出した.

2. 設定と記号 d_X-次元の密度関数モデル $\mathcal{P} = \{f(\mathbf{x}, \boldsymbol{\theta}) : \boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbf{R}^p\}$ で未知の母数ベクトルを $\boldsymbol{\theta} = (\boldsymbol{\theta}'_{(1)}, \boldsymbol{\theta}'_{(2)})'$, $\boldsymbol{\theta}_{(1)} = (\theta_1, \ldots, \theta_{p_1})'$, $\boldsymbol{\theta}_{(2)} = (\theta_{p_1+1}, \ldots, \theta_p)'$ のように分割し, 帰無仮説 $\boldsymbol{\theta}_{(1)} = \boldsymbol{\theta}_{(1)0}$ を検定する. $\ell_{j_{i1}\cdots j_{iR_i}}(\mathbf{x}, \boldsymbol{\theta}) = \frac{\partial}{\partial\theta_{j_{i1}}} \cdots \frac{\partial}{\partial\theta_{j_{iR_i}}} \log f(\mathbf{x}, \boldsymbol{\theta})$, $i = 1, \ldots, v$ の $f(\mathbf{x}, \boldsymbol{\theta})$ に関する v 次キュムラントを $\nu_{j_{11}\cdots j_{1R_1},\ldots,j_{v1}\cdots j_{vR_v}}(\boldsymbol{\theta})$ と書き, $Z^{(N)}_j(\boldsymbol{\theta}) = N^{-1/2}\sum_{i=1}^N \ell_j(\mathbf{X}_i, \boldsymbol{\theta})$, $Z^{(N)}_{j_1j_2}(\boldsymbol{\theta}) = N^{-1/2}\sum_{i=1}^N \{\ell_{j_1j_2}(\mathbf{X}_i, \boldsymbol{\theta}) - \nu_{j_1j_2}(\boldsymbol{\theta})\}$ などを定義する. このとき, 情報量行列 $[\nu_{j,k}(\boldsymbol{\theta})] = -[\nu_{jk}(\boldsymbol{\theta})]$ とスコアベクトルを分割し, $[\nu_{j,k}(\boldsymbol{\theta})] = \begin{pmatrix} \boldsymbol{\nu}_{(1,1)}(\boldsymbol{\theta}) & \boldsymbol{\nu}_{(1,2)}(\boldsymbol{\theta}) \\ \boldsymbol{\nu}_{(2,1)}(\boldsymbol{\theta}) & \boldsymbol{\nu}_{(2,2)}(\boldsymbol{\theta}) \end{pmatrix}$, $[Z^{(N)}_j(\boldsymbol{\theta})] = \begin{pmatrix} \mathbf{Z}^{(N)}_{(1)}(\boldsymbol{\theta}) \\ \mathbf{Z}^{(N)}_{(2)}(\boldsymbol{\theta}) \end{pmatrix}$ と書く. $p \times p_1$ 行列 $[\mathcal{G}_{j,a}(\boldsymbol{\theta})] = \begin{pmatrix} \mathbf{I}_{p_1} \\ -\boldsymbol{\nu}^{-1}_{(2,2)}(\boldsymbol{\theta})\boldsymbol{\nu}_{(2,1)}(\boldsymbol{\theta}) \end{pmatrix}$ を定義する. 以下, 任意の関数 $Q(\cdot)$ に対し i)–iii)

の表記に従う;

i) Unrestricted MLE $\widehat{\boldsymbol{\theta}}_{\mathrm{ML}}^{(N)}$ での評価は, $\widehat{\boldsymbol{\theta}}_{\mathrm{ML}}^{(N)}$ を省略してハット記号 ^ を付ける.

ii) Restricted MLE $\widetilde{\boldsymbol{\theta}}_{\mathrm{ML}}^{(N)}$ での評価は, $\widetilde{\boldsymbol{\theta}}_{\mathrm{ML}}^{(N)}$ を省略してチルド記号 ~ を付ける.

iii) $\boldsymbol{\theta}^{\dagger}=(\boldsymbol{\theta}'_{(1)0},(\boldsymbol{\theta}^{\dagger}_{(2)})')'$ での評価は, $\boldsymbol{\theta}^{\dagger}$ を省略する.

なお, 添え字については, $j_1,\ldots,j_R\in\{1,\ldots,p\}$ (j,k も同様), $a_1,\ldots,a_R\in\{1,\ldots,p_1\}$ (a,b も同様), $r_1,\ldots,r_R\in\{p_1+1,\ldots,p\}$ (r,s も同様) と約束し, それらの和 $\sum_{j_1,\ldots,j_R=1}^{p}$, $\sum_{a_1,\ldots,a_R=1}^{p_1}$, $\sum_{r_1,\ldots,r_R=p_1+1}^{p}$ に対しては Einstein 表記法を採用する.

3. 一般化された Cordeiro–Ferrari バートレット型調整　統計量のクラス $\mathcal{T}_N$

$$
\begin{aligned}
T^{(N)}(C,D)\approx{}&\widetilde{Z}_a^{(N)}\widetilde{\nu}_{(11\cdot2)}^{a,b}\widetilde{Z}_b^{(N)}+\frac{2}{N^{1/2}}\Big(\widetilde{C}^{\mathcal{G}\,\mathcal{G}\,\mathcal{G}}_{b_1b_2b_3}\prod_{i=1}^{3}[\widetilde{\boldsymbol{\nu}}^{-1}_{(11\cdot2)}\widetilde{\mathbf{Z}}^{(N)}_{(1)}]_{b_i}+\widetilde{C}^{\mathcal{G}\,\mathcal{G}}_{b_1b_2,k_1k_2}\widetilde{Z}^{(N)}_{k_1k_2}\prod_{i=1}^{2}[\widetilde{\boldsymbol{\nu}}^{-1}_{(11\cdot2)}\widetilde{\mathbf{Z}}^{(N)}_{(1)}]_{b_i}\Big)\\
&+\frac{2}{N}\Big(\widetilde{D}^{\mathcal{G}\,\mathcal{G}\,\mathcal{G}\,\mathcal{G}}_{b_1b_2b_3b_4}\prod_{i=1}^{4}[\widetilde{\boldsymbol{\nu}}^{-1}_{(11\cdot2)}\widetilde{\mathbf{Z}}^{(N)}_{(1)}]_{b_i}+\widetilde{D}^{\mathcal{G}\,\mathcal{G}\,\mathcal{G}}_{b_1b_2b_3,k_1k_2}\widetilde{Z}^{(N)}_{k_1k_2}\prod_{i=1}^{3}[\widetilde{\boldsymbol{\nu}}^{-1}_{(11\cdot2)}\widetilde{\mathbf{Z}}^{(N)}_{(1)}]_{b_i}\\
&\quad+\widetilde{D}^{\mathcal{G}\,\mathcal{G}\,\mathcal{G}}_{b_1b_2b_3,k_1k_2k_3}\widetilde{Z}^{(N)}_{k_1k_2k_3}\prod_{i=1}^{3}[\widetilde{\boldsymbol{\nu}}^{-1}_{(11\cdot2)}\widetilde{\mathbf{Z}}^{(N)}_{(1)}]_{b_i}+\widetilde{D}^{\mathcal{G}\,\mathcal{G}}_{b_1b_2,k_1k_2,k_3k_4}\widetilde{Z}^{(N)}_{k_1k_2}\widetilde{Z}^{(N)}_{k_3k_4}\prod_{i=1}^{2}[\widetilde{\boldsymbol{\nu}}^{-1}_{(11\cdot2)}\widetilde{\mathbf{Z}}^{(N)}_{(1)}]_{b_i}\Big)
\end{aligned}
$$

に対して, 以下のような形式の調整を考えた:

$$
T^{\mathrm{GCF}(N)}=T^{(N)}(C,D)+\frac{2}{N}\sum_{R=2,4,6}\widetilde{\Gamma}_{b_1\ldots b_R}\prod_{i=1}^{R}[\widetilde{\boldsymbol{\nu}}^{-1}_{(11\cdot2)}\widetilde{\mathbf{Z}}^{(N)}_{(1)}]_{b_i}\,.
$$

このとき, $P^{(N)}_{\theta_{(1)0},\theta^{\dagger}_{(2)}}[T^{\mathrm{GCF}(N)}\le x]=\Pr[\chi^2_{p_1}\le x]+\mathrm{o}(N^{-1})$ なる対称 R 次元配列関数 $\Gamma_R(\cdot)=[\Gamma_{a_1\ldots a_R}(\cdot)]$, $R=2,4,6$ は無数に存在する.

例 1　対称行列関数 $\mathbf{A}(\cdot)=[A_{a_1a_2}(\cdot)]$ 毎に

$$
\Gamma^{ACD}_{a_1a_2}(\cdot)=-\frac{\beta_1^{CD}(\cdot)}{\phi_1^A(\cdot)}\,A_{a_1a_2}(\cdot)\,,\quad \Gamma^{ACD}_{a_1a_2a_3a_4}(\cdot)=-\frac{\beta_2^{CD}(\cdot)}{\phi_2^A(\cdot)}\,\frac{\langle 3\rangle}{3}\,A_{a_1a_2}(\cdot)A_{a_3a_4}(\cdot)\,,
$$
$$
\Gamma^{AC}_{a_1a_2a_3a_4a_5a_6}(\cdot)=-\frac{\beta_3^{C}(\cdot)}{\phi_3^A(\cdot)}\,\frac{\langle 15\rangle}{15}\,A_{a_1a_2}(\cdot)A_{a_3a_4}(\cdot)A_{a_5a_6}(\cdot)
$$

を採用できる. ただし

$$
\phi_1^A(\boldsymbol{\theta})=\mathrm{tr}\,\{\mathbf{A}(\boldsymbol{\theta})\boldsymbol{\nu}^{-1}_{(11\cdot2)}(\boldsymbol{\theta})\}\neq0\,,\ \phi_2^A(\boldsymbol{\theta})=[\mathrm{tr}\,\{\mathbf{A}(\boldsymbol{\theta})\boldsymbol{\nu}^{-1}_{(11\cdot2)}(\boldsymbol{\theta})\}]^2+2\mathrm{tr}\,[\{\mathbf{A}(\boldsymbol{\theta})\boldsymbol{\nu}^{-1}_{(11\cdot2)}(\boldsymbol{\theta})\}^2]\neq0\,,
$$
$$
\phi_3^A(\boldsymbol{\theta})=[\mathrm{tr}\,\{\mathbf{A}(\boldsymbol{\theta})\boldsymbol{\nu}^{-1}_{(11\cdot2)}(\boldsymbol{\theta})\}]^3+6[\mathrm{tr}\,\{\mathbf{A}(\boldsymbol{\theta})\boldsymbol{\nu}^{-1}_{(11\cdot2)}(\boldsymbol{\theta})\}]\mathrm{tr}\,[\{\mathbf{A}(\boldsymbol{\theta})\boldsymbol{\nu}^{-1}_{(11\cdot2)}(\boldsymbol{\theta})\}^2]+8\mathrm{tr}\,[\{\mathbf{A}(\boldsymbol{\theta})\boldsymbol{\nu}^{-1}_{(11\cdot2)}(\boldsymbol{\theta})\}^3]\neq0
$$

を満たす. $\mathbf{A}(\cdot)=\boldsymbol{\nu}_{(11\cdot2)}(\cdot)$ のときオリジナルの Cordeiro–Ferrari バートレット型調整とみなせるが, その他にも (i) 任意に固定された $\boldsymbol{\gamma}\in\mathbf{R}^{p_1}\backslash\{\mathbf{0}_{p_1}\}$ について $\mathbf{A}(\cdot)\equiv\boldsymbol{\gamma}\boldsymbol{\gamma}'$, (ii) 任意の整数 $n=0,\pm1,\pm2,\ldots$ について $\mathbf{A}(\cdot)=\boldsymbol{\nu}^{n}_{(11\cdot2)}(\cdot)$ など存在する.

例 2　例 1 はスカラー関数 $\beta_R(\cdot)$ に比例する $\Gamma_R(\cdot)$ を選び, 帰無分布の漸近展開 $\int_0^x g_{p_1}(t)\,dt-\frac{2}{N}\{\pi_1^{CD\Gamma_2}g_{p_1+2}(x)+\pi_2^{CD\Gamma_4}g_{p_1+4}(x)+\pi_3^{C\Gamma_6}g_{p_1+6}(x)\}+\mathrm{o}(N^{-1})$ の N^{-1} 項を消失させるというアイデアに基づくが, $\pi_1^{CD\Gamma_2}=-\Gamma^{CD}_{b_1b_2}\nu^{b_1,b_2}_{(11\cdot2)}+\Gamma_{b_1b_2}\nu^{b_1,b_2}_{(11\cdot2)}$ の構造に着目すれば, 明らかに $\Gamma_{a_1a_2}(\cdot)=\Gamma^{CD}_{a_1a_2}(\cdot)$ も候補になる. さらに, 同様の考察から $\Gamma_6(\cdot)=\Gamma_6^{\star C}(\cdot)$, $\Gamma_4(\cdot)=\Gamma_4^{\star CD}(\cdot)$ も提示した [例 1 と例 2 の混合も許される].

TailCor: a new measure for tail association

David Veredas (Univ. Libre de Brussels)
(joint with Lorenzo Ricci)

Abstract:

A quantile-based method to measure tail correlations within the elliptical class distributions (TailCor) is proposed. It differs from tail dependence in that TailCor is not based on tail asymptotic arguments, and hence can be applied to any probability level. The use of TailCor is straightforward: it is a simple function and it disentangles the contribution of linear and non-linear correlation, the latter depending on the tail index. A Monte Carlo study reveals the goodness of the measure, both in terms of computational time and for finite samples. An empirical illustration to a large panel of securities (the constituents of S&P500) over the financial crisis illustrates the usefulness of TailCor.

Group LASSO for Structural Break Time Series [1]

Ngai Hang Chan
Department of Statistics
Chinese University of Hong Kong
Shatin, NT, Hong Kong
nhchan@sta.cuhk.edu.hk

Abstract

Consider a structural break autoregressive (SBAR) process

$$Y_t = \sum_{j=1}^{m+1} \boldsymbol{\beta}_j^{0\mathrm{T}} \boldsymbol{Y}_{t-1} I(t_{j-1} \leq t < t_j) + \varepsilon_t,$$

where $\boldsymbol{Y}_{t-1} = (1, Y_{t-1}, \ldots, Y_{t-p})^{\mathrm{T}}$, $\boldsymbol{\beta}_j^0 = (\beta_{j0}^0, \beta_{j1}^0, \ldots, \beta_{jp}^0)^{\mathrm{T}} \in \boldsymbol{R}^{p+1}$, $j = 1, \ldots, m+1$, $1 = t_0 < t_1 < \ldots < t_{m+1} = n+1$, $\{t_1, \ldots, t_m\}$ are change points, $\{\varepsilon_t\}$ are independent and identically distributed (i.i.d.) innovations with zero mean and unit variance. In practice, it is usually assumed that m is known and small, because a large m would involve a huge amount of computational burden in parameters estimation. By reformulating the problem in a regression variable selection context, the group least absolute shrinkage and selection operator (LASSO) is proposed to estimate an SBAR model when the number of change points m is unknown. It is shown that the number of change points and the locations of the changes can be consistently estimated from the data and the computation can be efficiently performed. Furthermore, the convergence rate of the breaks is shown to be nearly optimal. An improved practical version that incorporates group LASSO and stepwise regression variable selection technique is discussed. Simulation studies are conducted to assess the finite sample performance.

[1]Waseda Time Series Symposium, Waseda University, Tokyo. Feb 27 – March 4, 2012. Joint work with C.Y. Yau and R.M. Zhang. Research supported in part by grants from HKSAR-RGC-GRF.

Frequency Domain Techniques in the Analysis of DNA Sequences

David Stoffer, (University of Pittuburgh)

Abstract:

The concept of the spectral envelope for analyzing periodicities in categorical-valued time series was introduced in the statistics literature in Stoffer et al. (1993a) as a computationally simple and general statistical methodology for the harmonic analysis and scaling of non-numeric sequences. In the process of developing the technology, many possible interesting adaptations became apparent; for example, Stoffer & Tyler (1998) consider the maximal squared coherency between two categorical-valued time series. One of the most interesting directions was the use of the technology in the analysis of long DNA sequences. A benefit of the techniques was that it combined rigorous statistical analysis with modern computer power to quickly search for diagnostic patterns within long DNA sequences. The methodology is closely related to frequency domain principal component analysis and canonical correlation analysis of time series. I will present some of the theory and methods of the spectral envelope and related techniques, various analyses of DNA sequences are included. The investigations focus primarily, but not exclusively, on the analysis of viruses. The problems addressed concern period lengths in nucleosome positioning signals, optimal alphabets in codon usage, and sequence alignment.

Robust portfolio estimation under skew-normal return processes

Alexandre Petkovic (Waseda University)

Abstract:

In this talk, we study issues related to the optimal portfolio estimators and the local asymptotic normality (LAN) of the return process under the assumption that the return process has an infinite moving average (MA) (∞) representation with skew-normal innovations. The paper consists of two parts. In the first part, we discuss the influence of the skewness parameter δ of the skew-normal distribution on the optimal portfolio estimators. Based on the asymptotic distribution of the portfolio estimator $\hat{g}$ for a non-Gaussian dependent return process, we evaluate the influence of δ on the asymptotic variance $V(\delta)$ of $\hat{g}$. We also investigate the robustness of the estimators of a standard optimal portfolio via numerical computations. In the second part of the paper, we assume that the MA coefficients and the mean vector of the return process depend on a lower-dimensional set of parameters. Based on this assumption, we discuss the LAN property of the return's distribution when the innovations follow a skew-normal law. The influence of δ on the central sequence of LAN is evaluated both theoretically and numerically.

Joint estimation of copula and quantiles for time series

Hiroaki Ogata

Waseda University
hiroakiogata@aoni.waseda.jp

We consider a bivariate strictly stationary process $\{\boldsymbol{Y}_t = (Y_{1t}, Y_{2t})', t \in \mathbb{Z}\}$ and assume a realization $\{\boldsymbol{Y}_t, t = 1, \ldots, T\}$. Denote the joint distribution function of $(Y_{1t}, Y_{2t})'$ and the marginal distribution functions of Y_{1t} and Y_{2t} by $F(\cdot,\cdot)$, $F_1(\cdot)$ and $F_2(\cdot)$, respectively. For a given pair of numbers $0 \le u_1, u_2 \le 1$, we are interested in estimating their quntiles

$$q_1 = F_1^{-1}(u_1),\ q_2 = F_2^{-1}(u_2)$$

and copula

$$c = C(u_1, u_2) = F(F_1^{-1}(u_1), F_2^{-1}(u_2))$$

where F_1^{-1} and F_2^{-1} are quasi inverse functions of F_1 and F_2, that is,

$$F_j^{-1}(u_j) = \inf\{y | F_j(y) \ge u_j\}, \quad (j = 1, 2).$$

We denote the interested quantities by $\boldsymbol{\theta} = (q_1, q_2, c)'$ and their true values by $\boldsymbol{\theta}_0 = (q_{10}, q_{20}, c_0)'$.

To estimate $\boldsymbol{\theta}$, we make use of the function

$$K(x) = \int_{-\infty}^{x} k(z)dz$$

where $k(z)$ is a kernel function satisfying Assumption 1, which is mentioned later. We construct the estimating functions:

$$\begin{aligned}
w_{1t}(\boldsymbol{\theta}) &= K\left(\frac{q_1 - Y_{1t}}{h}\right) - u_1, \\
w_{2t}(\boldsymbol{\theta}) &= K\left(\frac{q_2 - Y_{2t}}{h}\right) - u_2, \\
w_t(\boldsymbol{\theta}) &= K\left(\frac{q_1 - Y_{1t}}{h}\right) K\left(\frac{q_2 - Y_{2t}}{h}\right) - c,
\end{aligned}$$

where $h = h_T$ is a bandwidth. Now, we formulate the penalty function

$$\begin{aligned} Q_T(\boldsymbol{\theta}) &= Q_T(\boldsymbol{\theta})(\mathbf{Y}_1, \dots, \mathbf{Y}_T; \boldsymbol{\theta}) \\ &= \{\bar{w}_{1T}(\boldsymbol{\theta})\}^2 + \{\bar{w}_{2T}(\boldsymbol{\theta})\}^2 + \{\bar{w}_T(\boldsymbol{\theta})\}^2 \end{aligned}$$

where

$$\bar{w}_{jT}(\boldsymbol{\theta}) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} w_{jt}(\boldsymbol{\theta}), \quad (j = 1, 2)$$

$$\bar{w}_T(\boldsymbol{\theta}) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} w_t(\boldsymbol{\theta})$$

and consider to find $\boldsymbol{\theta}$ minimizing the penalty function. That is, the estimator is defined as

$$\hat{\boldsymbol{\theta}}_T = \arg\min_{\boldsymbol{\theta}} Q_T(\boldsymbol{\theta}).$$

We give the assumptions on the kernel function, the bandwidth and the process.

Assumption 1 (Kernel function and bandwidth) (i) Bandwidths satisfy $Th^4 \to 0$.

(ii) The kernel $k(z)$ has a support $(-1, 1)$.

(iii) The kernel $k(z)$ is symmetric.

(iv) The kernel $k(z)$ has a bounded derivative $k'(z)$.

Assumption 2 (Process) (i) The process $(\boldsymbol{Y}_t)$ is strong mixing with coefficients α_t such that $\alpha_T = o(T^{-d})$ for some $d > 1$, as $T \to \infty$.

(ii) The marginal distribution functions $F_j, j = 1, 2$, are continuously differentiable on the intervals $[F_j^{-1}(a) - \varepsilon, F_j^{-1}(b) + \varepsilon]$ for every $0 < a < b < 1$ and some $\varepsilon > 0$, with positive derivatives f_j. Moreover, the first partial derivatives of F exist and are Lipschitz continuous on the product of these intervals.

Then, we obtain the following theorem.

Theorem 1 *Under Assumptions 1 and 2,*

$$\sqrt{T}(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0) \xrightarrow{d} \mathcal{N}(\mathbf{0}, V^{-1} W V^{-1}).$$

where V and W are certain constant matrices.

Dynamic estimation of investment style for Equity managers

Takashi Yamashita (GPIF)

Abstract:

In fund management, manager allocation is one of the most important issues for investors. Traditionally, many investors use the returns-base style analysis to deal with this issue. This research indicates the statistical difficulty in practical use. I propose a dynamic model of returns-based style analysis. We can recognize manager's style drift and improve manager allocation with this model.

Maximum entropy test for time series models

Sangyeol Lee (Seoul National Univ.)

Abstract:

In this talk, we discuss an application of the maximum entropy test, developed for a goodness of fit in iid samples by Lee et al. (2011), to time series models including non-stationary unstable models. Its asymptotic distribution is derived under the null hypothesis and its performance is investigated through Monte Carlo simulations. Vasicek's test will be also discussed.