観測誤差と線形制約を伴う真のデータの推定に関する新たな近接法（未定稿）

千木良 弘朗　　山本 拓

東北大学大学院経済学研究科　日本大学経済学部

研究集会「経済統計・政府統計の数理的基礎と応用」

2015 年 1 月 30 日 東京大学

1 はじめに

“観測誤差と線形制約を伴う真のデータ” とは?

例. 国民経済計算における供給表と使用表

供給表:

商品 \ 産業	製造業	サービス業	総計
製造品	900	100	1000
サービス	300	500	800
総計	1200	600	

使用表:

商品 \ 産業	製造業	サービス業	消費	総計
製造品	150	50	750	950
サービス	250	100	250	600
賃金	400	250		
営業余剰	350	100		
総計	1150	500		

$\Rightarrow$ 　　　、　　　、　　　、　　　 は等しくなければならない

$\Rightarrow$ **真のデータでは等しいが、観測誤差のために等しくなくなっている**

真のデータの推定

供給表:

商品＼産業	製造業	サービス業	総計
製造品	880	90	970
サービス	260	460	720
総計	1140	550	

使用表:

商品＼産業	製造業	サービス業	消費	総計
製造品	160	40	770	970
サービス	270	150	300	720
賃金	410	280		
営業余剰	300	80		
総計	1140	550		

⇒ [ ]、[ ]、[ ]、[ ]を等しくするよう観測誤差を修正する

⇒ **修正法は国によって違う。アメリカとオランダは統計学の理論に基づいたストーン法（Stone，Champernowne and Meade (1942)）を採用**

⇒ **本稿では、ストーン法も部分的に使うが、全く新しいアプローチを導入する**

発表の構成

2. ストーン法と先行研究の概観

3. 新しいアプローチ

4. モンテカルロ実験

5. まとめ

3 ストーン法と先行研究の概観

モデル

$$\underset{(N\times 1)}{\boldsymbol{X}} = \boldsymbol{X}^* + \boldsymbol{u},\ \boldsymbol{u} \sim (\boldsymbol{0}, \boldsymbol{\Omega}_{\boldsymbol{u}})$$

$$\underset{(R\times N)}{\boldsymbol{A}'} \boldsymbol{X}^* = \boldsymbol{h}$$

- $\boldsymbol{X}$: 観測されたデータ
- $\boldsymbol{X}^*$: 真のデータ
- $\boldsymbol{u}$: 観測誤差
- $\boldsymbol{A}$、$\boldsymbol{h}$: $\boldsymbol{X}^*$ が満たすべき線形制約

例. 国民経済計算における供給表と使用表

供給表:

商品＼産業	製造業	サービス業	総計
製造品	x_1	x_2	
サービス	x_3	x_4	
総計			

使用表:

商品＼産業	製造業	サービス業	消費	総計
製造品	x_5	x_6	x_7	
サービス	x_8	x_9	x_{10}	
賃金	x_{11}	x_{12}		
営業余剰	x_{13}	x_{14}		
総計				

$\rightarrow A'X^* = h$ は、

$$\begin{bmatrix} 1 & 1 & 0 & 0 & -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & -1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1^* \\ \vdots \\ x_{14}^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

ストーン法

$$\min(X - X^*)'\Omega_u^{-1}(X - X^*), \text{ s.t.} A'X^* = h$$

$$\rightarrow \boxed{\widehat{X}^* = X - \Omega_u A(A'\Omega_u A)^{-1}(A'X - h)}$$

別の解釈:
簡単化のため $h = 0$ とすると制約条件は $A'X^* = 0$ で

$$\begin{aligned}\widehat{X}^* &= X - \Omega_u A(A'\Omega_u A)^{-1}A'X \\ &= (I - \tilde{A}(A'\tilde{A})^{-1}A')X \qquad (\tilde{A} = \Omega_u A) \\ &= A_\perp(\tilde{A}'_\perp A_\perp)^{-1}\tilde{A}'_\perp X \qquad (\text{“}\perp\text{” は直行補空間})\end{aligned}$$

→ X^* が $A_\perp$ に属している（フルランク N ではなくランク落ちして $N - R$）ため、X も $A_\perp$ に属するよう変換しているのがストーン法

Ω_uの設定について

ストーン法はΩ_uが解らないと実行不能

→ **統計理論に基づいた解決法はそう多くなく、ある程度恣意的に決められることが多い**

		業種						
		農林水産業	鉱業	製造業	建 設 業	電気・ガス・水道業	卸売・小売業	金融・保険業
商品	農林水産業	1,981.2	1.5	8,412.4	250.1	1.7	1,605.1	0.0
	鉱業	0.3	8.4	8,506.8	1,037.0	2,336.4	3.5	0.0
	製造業	3,510.3	281.2	138,525.2	28,727.9	1,669.4	5,702.9	1,517.6
	建 設 業	85.5	9.7	1,410.0	208.1	1,131.2	554.8	160.1
	電気・ガス・水道業	106.8	43.3	6,522.2	500.4	1,335.1	1,128.4	220.5
	卸売・小売業	4.0	0.0	0.0	0.0	0.0	672.0	0.0
	金融・保険業	137.3	42.1	1,419.1	480.4	224.5	1,723.5	1,150.0

野木森 稔, 「加重最小二乗法を利用したバランシング・モデル-SUT バランスシステム開発に向けた一考察」, 季刊国民経済計算, 147, 69–89, 2012 年
より抜粋

○ 統計理論に基づいた解決法

Weale (1992) 等のアイデア:

$$X_t = X_t^* + u_t,\ u_t \sim (0, \Omega_u),\ t = 1, \cdots, T$$

$$\begin{cases} X_t、X_t^*、u_t \text{は定常} \\ Cov(X_t^*, u_t) = 0 \end{cases}$$

とすると、簡単化のため $E(X_t^*) = 0$ として

$$\begin{aligned} \boxed{\widehat{A'\Omega_u}} &= \frac{1}{T}\sum_{t=1}^{T} A' X_t X_t' \\ &= \frac{1}{T}\sum_{t=1}^{T} A' u_t (X_t^* + u_t)' \xrightarrow{p} \boxed{A'\Omega_u} \end{aligned}$$

⇒ Ω_u ではなく $A'\Omega_u$ の一致推定でストーン法を実行可能にしている

先行研究の問題点と本稿での新しいアプローチ

問題点:

$$\begin{cases} \widehat{A'\Omega_u} \text{は} N << T \text{でないとパフォーマンスが悪い} \\ X_t^* \text{に定常性を仮定するのは非現実的} \\ \text{ストーン法では} u_t \text{を一致推定できない} \\ \Omega_u \text{は一致推定できない} \end{cases}$$

$\Rightarrow$

新しいアプローチ:

$$\begin{cases} \text{大きな} N \text{に対して良好なパフォーマンス} \\ X_t^* \text{に定常性を仮定しない} \\ u_t \text{(の一部)を一致推定する} \\ \Omega_u \text{を一致推定する} \end{cases}$$

3 新しいアプローチ

モデル

観測誤差にファクターモデルを適用する

$$X_t = X_t^* + u_t = X_t^* + \underset{(N\times q)}{C} \underset{(q\times 1)}{F_t} + \varepsilon_t,\ t = 1, \cdots, T$$

$$\begin{cases} C \text{はファクター負荷} \\ F_t \text{はファクター} \\ q \text{はファクター数} \end{cases}$$

$\Rightarrow$ **CF_tを一致推定して観測誤差を修正する。ε_tについてはストーン法で修正**

（簡単化のため、以下では$h = 0$とする）

モデルの推定の直感的な概観

$$X_t = X_t^* + CF_t + \varepsilon_t,\ \varepsilon_t \sim (0, \Omega),\ \Omega = \begin{bmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_N^2 \end{bmatrix}$$

⇓

$$A_N' X_t = A_N' C F_t + A_N' \varepsilon_t,\ A_N = \underset{(R\times R)}{A\ \Upsilon_N^{-1}},\ \Upsilon_N = \begin{bmatrix} n_1 & & 0 \\ & \ddots & \\ 0 & & n_R \end{bmatrix},$$

n_rはAの第r列のnon-zero要素の数

⇓

この式にBai (2003)の方法を適用して$A_N'\widehat{CH'^{-1}}$と$\widehat{H'F_t}$を得る

$$\Downarrow$$

$Cov(X_t^*, \varepsilon_t) = 0$、$Cov(F_t, \varepsilon_t) = 0$を仮定すると

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T}\Upsilon_N(A_N'X_t - A_N'\widehat{CH'}^{-1}\widehat{H'}F_t)X_t' &\approx \frac{1}{T}\sum_{t=1}^{T}\widehat{A'}\varepsilon_t X_t' \\ &\approx \frac{1}{T}\sum_{t=1}^{T}\widehat{A'}\varepsilon_t\varepsilon_t' \\ &\approx A'\Omega\end{aligned}$$

$$\Downarrow$$

この式を積率条件としてGMMで$\hat{\sigma}_1^2, \cdots, \hat{\sigma}_N^2$を得る

$$\Downarrow$$

$$\frac{1}{T}\sum_{t=1}^{T} A'_N X_t X'_t \approx A'_N C \Omega_F C' + A'_N \Omega \qquad (Cov(X_t^*, F_t) = 0 \text{を追加})$$

$$= A'_N C H'^{-1} H' \Omega_F H H^{-1} C' + A'_N \Omega$$

において、

$$\underbrace{\frac{1}{T}\sum_{t=1}^{T} A'_N X_t X'_t}_{\text{計算可能}} = \underbrace{A'_N C H'^{-1} H' \Omega_F H}_{\widehat{A'_N C H'^{-1}} \text{ と } \widehat{H' F_t} \text{ で推定可能}} \underbrace{H^{-1} C'}_{\text{未知パラメーター}} + \underbrace{A'_N \Omega}_{\hat{\sigma}_1^2, \cdots, \hat{\sigma}_N^2 \text{ で推定可能}}$$

$$\Downarrow$$

この式から $H^{-1}C'$ を逆算して $\widehat{H^{-1}C'}$ を得る

仮定:

1. $E(F_t) = 0$、$F_t, \varepsilon_t \sim I(0)$

2. Aの第r列と第s列で共通する項目の数c_{rs}が十分小さい
（全てのrについて$\sum_{s=1}^{R} c_{rs}$が$R \to \infty$としても発散しない）

3. $q < R$

4. Aの全行には少なくとも1つnon-zero要素がある

5. $X_t^* \sim I(0)$

6. $N, R, T \to \infty$ with $\frac{N^K}{\sqrt{T} R^{1-W}} \to 0$
($\sum_{r=1}^{R} n_R = O(N^K)$, $1 \le K \le 2$、
$n_1, \cdots, n_R$の内で$O(1)$なものの個数を$O(R^W)$, $0 \le W \le 1$)

$$\widehat{CH'^{-1}}\widehat{H'F_t} \xrightarrow{p} CF_t$$ （観測誤差（の一部）の一致推定）

$$\hat{\sigma}_1^2, \cdots, \hat{\sigma}_N^2 \xrightarrow{p} \sigma_1^2, \cdots, \sigma_N^2$$ （ストーン法の改善）

実際のprocedure

ファクター-ストーン法:

$\widetilde{X_t^*} = X_t - \widehat{CH'^{-1}}\widehat{H'F_t}$に、$\hat{\Omega} = \begin{bmatrix} \hat{\sigma}_1^2 & & 0 \\ & \ddots & \\ 0 & & \hat{\sigma}_N^2 \end{bmatrix}$を使ってストーン法を適用

$\rightarrow \widehat{X_t^*} = \widetilde{X_t^*} - \hat{\Omega}A(A'\hat{\Omega}A)^{-1}A'\widetilde{X_t^*}$

4 モンテカルロ実験

DGPと実験の設定

$$X_t = X_t^* + u_t = X_t^* + CF_t + \varepsilon_t$$

$$\begin{cases} X_t^* \sim \text{トレンド定常VAR(1)} \\ F_t \sim \text{期待値0の定常VAR(1)} \quad (C\Omega_F C'/Var(X_t^*) \approx 5\%) \\ \varepsilon_t \sim \text{期待値0の定常VAR(1)} \quad (\Omega/Var(X_t^*) \approx 5\%) \\ X_t^*\text{、}F_t\text{、}\varepsilon_t\text{は全て正規乱数で互いに独立} \end{cases}$$

$\Rightarrow$ $Var(X_t)$の内、$Var(X_t^*)$が約90%でΩ_uが約10%

$Var(X_t^*)$、$C\Omega_F C'$、Ω全てにおいて最小の分散と最大の分散には約500～1000倍の差をつけ、不均一性を表現

- X_t^* のVAR(1)パラメーターと定数・トレンド項は乱数で生成して固定
- F_t と ε_t のVAR(1)パラメーターは乱数で生成して固定
- C は乱数で生成して固定
- A は $-1, 0, 1, 2$ しかとらないsparse行列として乱数で生成して固定
- $q = 2$ で既知
- $N = 25, 100, 400$、$R = N/4, \sqrt{N}$、$T = 25, 50, 800$
- 繰り返し回数1000回

各修正法の具体的な計算

○ 原始的なストーン法（Stone1）

$$\Omega_u = I_N \text{として} \widehat{X_t^*} = X_t - A(A'A)^{-1}A'X_t$$

○ Weale (1992) の方法（Weale）

demean、detrend した X_t^d より $\widehat{A'\Omega_u} = \frac{1}{T}\sum_{t=1}^{T} A'X_tX_t^{d\prime}$ として $\widehat{X}_t^* = X_t - \widehat{\Omega_u A}(\widehat{A'\Omega_u}A)^{-1}A'X_t$

○ ファクター-ストーン法（F-Stone）

先述の通りだが、

$$\begin{cases} \frac{1}{T}\sum_{t=1}^{T} \Upsilon_N(A_N'X_t - A_N'\widehat{CH'}^{-1}\widehat{H'F}_t)X_t^{d\prime} \approx A'\Omega \\ \frac{1}{T}\sum_{t=1}^{T} A_N'X_tX_t^{d\prime} \approx A_N'CH'^{-1}H'\Omega_F HH^{-1}C' + A_N'\Omega \end{cases}$$

のように X_t^d で計算する。また、$\hat{\sigma}_i^2 \leq 0$ だった時は $\hat{\sigma}_i^2 = 0.01 \times \frac{1}{T}\sum_{t=1}^{T}(x_{it}^d)^2$ とする

推定精度の尺度

MSE$= E\{(X_t^* - \widehat{X_t^*})'(X_t^* - \widehat{X_t^*})\}$だと、

- 変動の大きい$x_{it}^* - \hat{x}_{it}^*$だけにdominateされてしまい不公平
- NやRを変えるとDGPのパラメーターの値の変化を反映してしまい、純粋なsample sizeの効果を捉えられない

$\Rightarrow$ **基準化MSE（sMSE）：**

$$\text{sMSE}= \frac{1}{N}E\left\{(X_t^* - \widehat{X_t^*})' \begin{bmatrix} Var(x_1^*) & & 0 \\ & \ddots & \\ 0 & & Var(x_N^*) \end{bmatrix}^{-1} (X_t^* - \widehat{X_t^*})\right\}$$

$$= \frac{1}{N}E\left\{\sum_{i=1}^{N} \frac{(x_{it}^* - \hat{x}_{it}^*)^2}{Var(x_i^*)}\right\}$$

$$\Rightarrow \quad \begin{cases} x^*_{it} = \mu_i + \delta_i t + z_{it},\ E(z_{it}) = 0,\ z_{it} \sim I(0) \\ \hat{x}^*_{it} = \mu_i + \delta_i t \end{cases}$$

（確定項だけ捉えている“最低限”の推定量）

とすると、$E\left\{\frac{(x^*_{it}-\hat{x}^*_{it})^2}{Var(x^*_i)}\right\} = E\left\{\frac{z^2_{it}}{Var(x^*_i)}\right\} = 1$

$$\Rightarrow \quad \begin{cases} \text{sMSE} < 1 \text{ ···最低限の規準はクリア} \\ \text{sMSE} \geq 1 \text{ ···推定精度に問題あり} \end{cases}$$

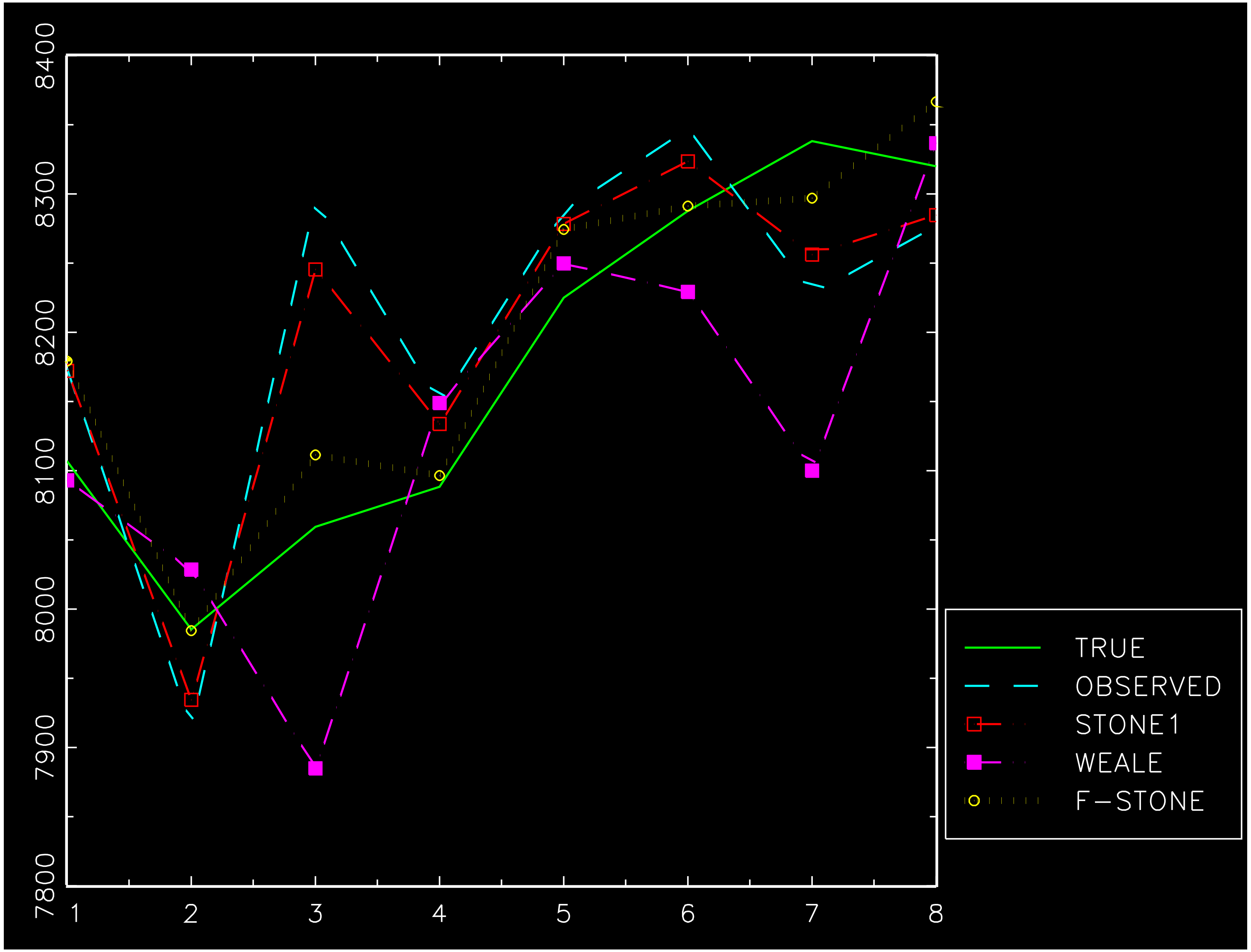
（実際はsMSEが$t = 1, \cdots, T$について得られるので$\frac{1}{T}\sum_{t=1}^{T} \text{sMSE}_t$を計算）

sMSE

		Stone1			Weale			F-Stone		
N	R	T			T			T		
		25	50	800	25	50	800	25	50	800
25	5	0.137	0.136	0.136	0.294	0.178	0.062	0.144	0.109	0.064
	7	0.152	0.152	0.151	0.398	0.222	0.060	0.143	0.106	0.057
100	10	0.138	0.139	0.139	0.599	0.307	0.068	0.144	0.110	0.059
	25	0.192	0.192	0.193	.	0.632	0.081	0.147	0.106	0.052
400	20	0.116	0.115	0.115	2.005	0.524	0.081	0.135	0.105	0.053
	100	0.137	0.138	0.138	.	.	0.195	0.139	0.104	0.045

実験データの時系列プロット

$N = 100$、$R = 25$、$T = 50$



5 まとめ

本稿の貢献

観測誤差を伴う真のデータの推定について、

- 観測誤差（の一部）の一致推定による直接的な推定
- （それ以外の観測誤差に関する）修正精度の改善

を提案した

推定精度

sMSE という尺度では、概ね F-Stone>Stone1>Weale

トレンド・季節性とマクロ時系列：(非定常) 変数誤差問題 (Trend, seasonality and macro time series : Nonstationary errors-in-variables)

国友直人・佐藤整尚

2015 年 1 月 30 日

1. A General Problem

Let y_{ij} be the $i-$th observation of the $j-$th time series at t_i^n for $i = 1, \cdots, n; j = 1, \cdots, p; 0 = t_0^n \leq t_1^n \leq \cdots \leq t_n^n = T$. We set $n = T$, $t_i^n - t_{i-1}^n = 1$, $\mathbf{y}_i = (y_{1i}, \cdots, y_{pi})'$ be a $p \times 1$ vector and $\mathbf{Y}_n = (\mathbf{y}_i') \; (= (y_{ij}))$ be an $n \times p$ matrix of observations. ($\mathbf{y}_0$ is the initial observation vector.) We consider the situation when the underlying non-stationary trends $\mathbf{x}_i \; (= (x_{ji}))$ at t_i^n $(i = 1, \cdots, n)$ are not necessarily the same as the observed time series and let $\mathbf{c}_i' = (c_{1i}, \cdots, c_{pi})$, $\mathbf{s}_i' = (s_{1i}, \cdots, s_{pi})$ and $\mathbf{v}_i' = (v_{1i}, \cdots, v_{pi})$ be the vectors of the seasonal components, the cycle components and the noise components at t_i^n, respectively, which are independent of $\mathbf{x}_i$. Then we use the additive decomposition (see Kitagawa (2010))

$$\mathbf{y}_i = \mathbf{x}_i + \mathbf{s}_i + \mathbf{c}_i + \mathbf{v}_i \; , \tag{1.1}$$

where $\mathbf{x}_i$ are a sequence of non-stationary trend components satisfying $\Delta \mathbf{x}_i = (1 - \mathcal{L})\mathbf{x}_i = \mathbf{w}_i^{(x)}$ (with $\mathcal{L}\mathbf{x}_i = \mathbf{x}_{i-1}$, $\mathcal{E}(\mathbf{w}_i^{(x)}) = \mathbf{0}$, $\mathcal{E}(\mathbf{w}_i^{(x)}\mathbf{w}_i^{(x)'}) = \mathbf{\Sigma}_x$), $\mathbf{s}_i$ are a sequence of seasonal components satisfying $(1 + \mathcal{L} + \cdots + \mathcal{L}^{s-1})\mathbf{s}_i = \mathbf{w}_i^{(s)}$ (with $\mathcal{E}(\mathbf{w}_i^{(s)}) = \mathbf{0}$, $\mathcal{E}(\mathbf{w}_i^{(s)}\mathbf{w}_i^{(s)'}) = \mathbf{\Sigma}_s$), $\mathbf{c}_i$ are a sequence of stationary cycle components (with $\mathcal{E}(\mathbf{c}_i) = \mathbf{0}$, $\mathcal{E}(\mathbf{c}_i\mathbf{c}_i') = \mathbf{\Sigma}_c$) and $\mathbf{v}_i$ are a sequence of independent noise components (with $\mathcal{E}(\mathbf{v}_i) = \mathbf{0}$, $\mathcal{E}(\mathbf{v}_i\mathbf{v}_i') = \mathbf{\Sigma}_v$).

We assume that $\mathbf{w}_i^{(x)}, \mathbf{w}_i^{(s)}$ and $\mathbf{v}_i$ are the sequence of i.i.d. random variables with $\mathbf{\Sigma}_v$ being positive definite and finite, and the random variables $\mathbf{w}_i^{(x)}, \mathbf{w}_i^{(s)}$, $\mathbf{c}_i$ and $\mathbf{v}_i$ are mutually independent. We also set $\mathbf{c}_i = \mathbf{0}$ in this occasion, but we can generalize the following arguments into more general directions with $\mathbf{c}_i \neq \mathbf{0}$ $(i = 1, \cdots, n)$.

The main purpose of this study is to estimate the structural relationships among the hidden random variables; the trend components and seaonal components in particular when we have stationary and non-stationary errors-in-variables models. Let $\boldsymbol{\beta}$ be a $p \times 1$ vector and we want to estimate

$$\boldsymbol{\beta}'\mathbf{x}_i = o_p(1) \tag{1.2}$$

when we have the observations of $p \times 1$ vectors $\mathbf{y}_i$ $(i = 1, \cdots, n)$. More generally, let $\mathbf{B}$ be a $q \times p$ $(q \leq p)$ matrix and we want to estimate

$$\mathbf{B}\mathbf{x}_i = o_p(1); \tag{1.3}$$

when we have the observations of $p \times 1$ vectors $\mathbf{y}_i$ $(i = 1, \cdots, n)$. Similarly, some structural relations among seasonal components can be written as

$$\mathbf{B}_s\mathbf{s}_i = o_p(1) , \tag{1.4}$$

and they imply that the observed multivariate time series have common seasonality.

2. Simple Cases

Consider some examples when $p = 2$ and $\mathbf{c}_i = \mathbf{0}$ (i.e. no-cycle components).

Example 1 : Assume that the random variables $x_{1i} = \nu_i = \beta_2\mu_i$ and $x_{2i} = \mu_i$ satisfy $\mu_i = \mu_{i-1} + w_i^{(x)}$ $(i = 1, \cdots, n)$ and $w_i^{(x)}$ are i.i.d. random variables with $\mathcal{E}(w_i^{(x)}) = 0$ and $\mathcal{E}(w_i^{(x)2}) = \sigma_x^2$. Then we can write

$$\mathbf{y}_i = \begin{pmatrix} \beta_2 \\ 1 \end{pmatrix} \mu_i + \mathbf{v}_i \ . \tag{2.1}$$

Since μ_i follows the random walk model,

$$\frac{1}{n^2}\sum_{i=1}^{n} \mu_i^2 \xrightarrow{p} \sigma_x^2 \int_0^1 B_s^2 ds \ , \tag{2.2}$$

where B_s is the standard Brownian Motion on $[0, 1]$.

If we multiply the vector $\boldsymbol{\beta}' = (1, -\beta_2)$ to (2.1) from the left, we have the relation

$$\boldsymbol{\beta}'\mathbf{y}_i = u_i \ (= \boldsymbol{\beta}'\mathbf{v}_i) \ , \tag{2.3}$$

which is a structural equation.

Example 2 : We take the case when $\mathbf{x}_i = \boldsymbol{\mu}_i$, and $\boldsymbol{\mu}_i = \boldsymbol{\mu}_{i-1} + \mathbf{w}_i^{(x)}$, which has been often called **spurious regression**. It can be written as

$$\mathbf{y}_i = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \boldsymbol{\mu}_i + \mathbf{v}_i \tag{2.4}$$

and the dimension of random walk is 2 and $\boldsymbol{\beta}'\mathbf{y}_i = \boldsymbol{\beta}'\boldsymbol{\mu}_i + u_i$, $u_i = \boldsymbol{\beta}_x'\mathbf{v}_i$ for any $\boldsymbol{\beta} \neq \mathbf{0}$ (the term of $\boldsymbol{\beta}'\boldsymbol{\mu}_i$ cannot be disappeared).

Example 3 : Assume that the random variables $s_{1i} = \nu_i^{(s)} = \beta_2^{(s)}\mu_i^{(s)}$ and $s_{2i} = \mu_i^{(s)}$

satisfy $\mu_i^{(s)} = \mu_{i-s}^{(s)} + w_i^{(s)}$ $(s \geq 1\ ;\ i = 1, \cdots, n)$ and $w_i^{(s)}$ are i.i.d. random variables with $\mathcal{E}(w_i^{(s)}) = 0$ and $\mathcal{E}(w_i^{(s)2}) = \sigma_s^2$. Then we can write

$$\mathbf{y}_i = \begin{pmatrix} \beta_2^{(s)} \\ 1 \end{pmatrix} \mu_i + \mathbf{v}_i \ . \tag{2.5}$$

If we multiply the vector $\boldsymbol{\beta}_s^{'} = (1, -\beta_2^{(s)})$ to (2.5) from the left, we have the relation

$$\boldsymbol{\beta}_s^{'}\mathbf{y}_i = u_i \ (= \boldsymbol{\beta}_s^{'}\mathbf{v}_i) \tag{2.6}$$

and $\mathbf{y}_i$ has the common seasonal components.

3. The Case without Seasonality

Let $p \geq 2$ and $\mathbf{s}_i = \mathbf{c}_i = \mathbf{0}$, we have the representation

$$\mathbf{y}_i = \mathbf{x}_i + \mathbf{v}_i = \boldsymbol{\Pi}\boldsymbol{\mu}_i + \mathbf{v}_i \ , \tag{3.1}$$

where $\mathcal{E}(\mathbf{w}_i^{(x)}) = \mathbf{0}$, $\mathcal{E}(\mathbf{w}_i^{(x)}\mathbf{w}_i^{(x)'}) = \boldsymbol{\Sigma}_x$ and $\mathbf{w}_i^{(x)} = \Delta\mathbf{x}_i$. We assume that the rank of $p \times q$ matrix $\boldsymbol{\Pi}$ is q $(\leq p)$, $\boldsymbol{\mu}_i$ are $q \times 1$ vectors, and there exists a $q \times p$ matrix $\mathbf{B}$ such that $\mathbf{B}\mathbf{y}_i = u_i$ $(= \mathbf{B}\mathbf{v}_i)$, which are the set of q structural equations.

We consider the situation when $\Delta\mathbf{x}_i$ and $\mathbf{v}_i$ $(i = 1, \cdots, n)$ are independent and they are independently, identically and normally distributed as $N_p(\mathbf{0}, \boldsymbol{\Sigma}_x)$ and $N_p(\mathbf{0}, \boldsymbol{\Sigma}_v)$, respectively. We use an $n \times p$ matrix $\mathbf{Y}_n = (\mathbf{y}_i^{'})$ and consider the distribution of $np \times 1$ random vector $(\mathbf{y}_1^{'}, \cdots, \mathbf{y}_n^{'})^{'}$. Given the initial condition $\mathbf{y}_0$, we have

$$\mathbf{Y}_n \sim N_{n \times p}\left(\mathbf{1}_n \cdot \mathbf{y}_0^{'}, \mathbf{I}_n \otimes \boldsymbol{\Sigma}_v + \mathbf{C}_n\mathbf{C}_n^{'} \otimes \boldsymbol{\Sigma}_x\right) \ , \tag{3.2}$$

where $\mathbf{1}_n^{'} = (1, \cdots, 1)$ and

$$\mathbf{C}_n = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 1 & \cdots & 0 \\ 1 & \cdots & 1 & 1 & 0 \\ 1 & \cdots & 1 & 1 & 1 \end{pmatrix}_{n \times n} \ . \tag{3.3}$$

Then given the initial condition $\mathbf{y}_0$ the maximum likelihood (ML) estimator can be defined as the solution of maximizing the log-likelihood function except a constant as

$$\begin{aligned} L_n^* &= \log|\mathbf{I}_n \otimes \boldsymbol{\Sigma}_v + \mathbf{C}_n\mathbf{C}_n^{'} \otimes \boldsymbol{\Sigma}_x|^{-1/2} - \frac{1}{2}[vec(\mathbf{Y}_n - \bar{\mathbf{Y}}_0)^{'}]^{'}[\mathbf{I}_n \otimes \boldsymbol{\Sigma}_v + \mathbf{C}_n\mathbf{C}_n^{'} \otimes \boldsymbol{\Sigma}_x]^{-1} \\ &\qquad \times[vec(\mathbf{Y}_n - \bar{\mathbf{Y}}_0)^{'}] \end{aligned}$$

and

(3.4) $$\bar{\mathbf{Y}}_0 = \mathbf{1}_n \cdot \mathbf{y}_0^{'} \ .$$

We transform $\mathbf{Y}_n$ to $\mathbf{Z}_n \ (= (\mathbf{z}_k^{'}))$ by

(3.5) $$\mathbf{Z}_n = \mathbf{P}_n \mathbf{C}_n^{-1} \left(\mathbf{Y}_n - \bar{\mathbf{Y}}_0 \right)$$

where

(3.6) $$\mathbf{C}_n^{-1} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & 0 & \cdots \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}_{n \times n} ,$$

(3.7) $$\mathbf{P}_n = (p_{jk}^{(n)}) \ , \ p_{jk}^{(n)} = \sqrt{\frac{2}{n+\frac{1}{2}}} \cos \left[\frac{2\pi}{2n+1} (k - \frac{1}{2})(j - \frac{1}{2}) \right] \ .$$

By using the spectral decomposition $\mathbf{C}_n^{-1} \mathbf{C}_n^{'-1} = \mathbf{P}_n \mathbf{D}_n \mathbf{P}_n^{'}$ and $\mathbf{D}_n$ is a diagonal matrix with the k-th element

$$d_k = 2[1 - \cos(\pi(\frac{2k-1}{2n+1}))] \ (k = 1, \cdots, n) \ .$$

Then the log-likelihood function is proportional to

(3.8) $$L_n = \sum_{k=1}^{n} \log |a_{kn} \mathbf{\Sigma}_v + \mathbf{\Sigma}_x|^{-1/2} - \frac{1}{2} \sum_{k=1}^{n} \mathbf{z}_k^{'} [a_{kn} \mathbf{\Sigma}_v + \mathbf{\Sigma}_x]^{-1} \mathbf{z}_k \ ,$$

where

(3.9) $$a_{kn} \ (= d_k) \ = 4 \sin^2 \left[\frac{\pi}{2} \left(\frac{2k-1}{2n+1} \right) \right] \ (k = 1, \cdots, n) \ .$$

Since we are dealing with the errors-in-variables model, there is an issue if we can identify the structural equation of our interest. When $\mathbf{x}_i$ are i.i.d. random variables, for instance, the coefficient parameters are not identified without some further restrictions. In the classical case we need to impose some conditions on the covariance matrix (such as the homoscedasticity and zero covariance) when we have the functional relationship model in Example 1 except $x_i \ (= \mu_i \ ; i = 1, \cdots, n)$ with

(3.10) $$\frac{1}{n} \sum_{i=1}^{n} \mu_i^2 \xrightarrow{p} \sigma_x^2 \ .$$

(See Fuller (1987) for instance.) Here we say that the parameter vector $\boldsymbol{\theta} \ (= (\theta_j))$ is identified if $\boldsymbol{\theta} \neq \boldsymbol{\theta}^{'}$ implies that $L_n(\theta) \neq L_n(\theta^{'})$.

We illustrate our arguments on the likelihood function when $q = 1$. We take $\boldsymbol{\theta} \ (= \boldsymbol{b})$ and apply the matrix formulae that for a positive definite $\mathbf{A}$ we have

$$|\mathbf{A} + \mathbf{b}\mathbf{b}^{'}| = |\mathbf{A}|[1 + \mathbf{b}^{'} \mathbf{A}^{-1} \mathbf{b}]$$

and

$$[\mathbf{A}+\mathbf{b}\mathbf{b}^{'}]^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{b}[1+\mathbf{b}^{'}\mathbf{A}^{-1}\mathbf{b}]^{-1}\mathbf{b}^{'}\mathbf{A}^{-1}$$

for $\mathbf{A} = a_{kn}\mathbf{\Sigma}_v$ $(k = 1, \cdots, n)$, $\mathbf{b} = \sigma_\mu \mathbf{\Pi}$ and $\mathbf{b}_* = \mathbf{\Sigma}_v^{-1}\sigma_\mu\mathbf{\Pi}$.
Then L_n is proportional to (-1/2) times

$$(3.11) \quad L_{1n} = \sum_{k=1}^{n}\left[\log|\mathbf{\Sigma}_v| + \log(a_{kn}+\mathbf{b}^{'}\mathbf{\Sigma}_v^{-1}\mathbf{b}) + a_{kn}^{-1}\mathbf{z}_k^{'}\mathbf{\Sigma}_v^{-1}\mathbf{z}_k - \frac{a_{kn}^{-1}(\mathbf{z}_k^{'}\mathbf{\Sigma}_v^{-1}\mathbf{b})^2}{a_{kn}+\mathbf{b}^{'}\mathbf{\Sigma}_v^{-1}\mathbf{b}}\right] .$$

We notice that L_{1n} is a concave function of $\mathbf{\Sigma}_v^{-1}$ and

$$\sum_{k=1}^{n}\frac{(\mathbf{z}_k^{'}\mathbf{\Sigma}_v^{-1}\mathbf{b})^2}{a_{kn}^2 + a_{kn}\mathbf{b}^{'}\mathbf{\Sigma}_v^{-1}\mathbf{b}} = \mathbf{b}_*^{'}\sum_{k=1}^{n}[\frac{1}{a_{kn}(a_{kn}+c)}\mathbf{z}_k\mathbf{z}_k^{'}]\mathbf{b}_* ,$$

where we set $\mathbf{b}^{'}\mathbf{\Sigma}_v^{-1}\mathbf{b} = c$. Then given $\mathbf{b}_*^{'}\mathbf{\Sigma}_v\mathbf{b}_* = c$, it is a quadratic form and its minimum is the smallest characteristic root. Thus we have the next results.

Proposition 1 : Under the above assumptions, the structural parameter $\boldsymbol{\theta}$ (i.e. $\mathbf{B}$) is identified (up to a normalization).

Given the initial condition $\mathbf{y}_0$, the unconditional maximum likelihood (ML) estimator can be defined as the solution of maximizing the log-likelihood function.

Proposition 2 : Under the above assumptions, there exists a unique ML estimator for $\mathbf{B}$.

Next, given $\mathbf{b}^{'}\mathbf{\Sigma}_v^{-1}\mathbf{b} = \mathbf{b}_*^{'}\mathbf{\Sigma}_v\mathbf{b}_* = c$ (*a constant*), we approximate L_{1n} of the function of $\mathbf{\Sigma}_v$ and $\mathbf{b}_*$. If we use

$$(3.12) \qquad \mathbf{\Sigma}_v = \frac{1}{n}\sum_{k=n}^{n} a_{kn}^{-1}\mathbf{z}_k\mathbf{z}_k^{'} ,$$

and $1/[a_{kn}(a_{kn}+c)] = (1/c)[1/a_{kn} - 1/(a_{kn}+c)]$, then given $\mathbf{b}_*^{'}\mathbf{\Sigma}_v\mathbf{b}_* = c$, the problem becomes to find a minimum of

$$(3.13) \qquad \mathbf{b}_*^{'}\sum_{k=1}^{n}[\frac{1}{a_{kn}+c}\mathbf{z}_k\mathbf{z}_k^{'}]\mathbf{b}_* .$$

Then it may be natural to use the minimization of

$$(3.14) \qquad \frac{1}{c}\mathbf{b}_*^{'}\mathbf{A}_m\mathbf{b}_* \ \ (\text{s.t. } \mathbf{b}_*^{'}\mathbf{\Sigma}_v\mathbf{b}_* = c, \mathbf{A}_m = \sum_{k=1}^{m_n}\mathbf{z}_k\mathbf{z}_k^{'}),$$

where we use the fact that $c < a_{kn} + c < 4 + c$ and $a_{kn} \sim O(1/n)$ when $k_n = o(1)$. It is straightforward to extend the above likelihood analysis to the cases for more general q $(q \leq p)$.

4. Macro-SIML Estimation

The exact ML estimator of unknown parameters is a rather complicated function of all observations and it may depend crucially on the underlying assumptions including the Gaussianity. Then we need a simple robust procedure such that the assumptions of Gaussianity and the specifications of each components are not crucial for the estimation results.

Let denote $a_{k_n,n}$ and then we can evaluate that $a_{k_n,n} \to 0$ as $n \to \infty$ when $k_n = O(n^\alpha)$ $(0 < \alpha < 1)$ since $\sin x \sim x$ as $x \to 0$. When k_n is small, we expect that $a_{k_n,n}$ is small. Then the separating information maximum likelihood (SIML) estimator of $\hat{\mathbf{\Sigma}}_x$ is defined by

$$\hat{\mathbf{\Sigma}}_{x,SIML} = \frac{1}{m_n} \sum_{k=1}^{m_n} \mathbf{z}_k \mathbf{z}_k^{'} \ . \tag{4.1}$$

(We need to use a consistent estimator for $\mathbf{\Sigma}_v$.) For $\hat{\mathbf{\Sigma}}_x$, the number of terms m_n should be dependent on n. Then we need the order requirement that $m_n = O(n^\alpha)$ $(0 < \alpha < 1)$.

It may be natural to use the sample quantities $\hat{\mathbf{\Sigma}}_x$ $(= ((1/m_n)\sum_{k=1}^{m_n} z_{ik} z_{jk}))$ order to make statistical inference on $\mathbf{\Sigma}_x$. The estimation of the Pearson-correlation coefficent is a typical case, which is given by

$$\hat{\rho}_{ij} = \frac{\sum_{k=1}^{m_n} z_{ik} z_{jk}}{\sqrt{\sum_{k=1}^{m_n} z_{ik}^2}\sqrt{\sum_{k=1}^{m_n} z_{jk}^2}} \ . \tag{4.2}$$

Furthermore, we consider the estimation of the structural relationships in the non-stationary time series process satisfying (3.1). Hewe we notice that the present statistical problem could be regarded as the estimation of structural relationships with the covariance matrix $\mathbf{\Sigma}_x(\boldsymbol{\theta})$ with $\boldsymbol{\theta}$ being the vector of parameters. In the standard statistical multivariate analysis Anderson (1984) discussed the statistical models of estimating structural relationships among a set of variables and we have n independent observations on the underlying variables.

We return to the case when $q = 1$. By using the arguments on the likelihood function, it may be natural to consider the characteristic equation

$$\left[\mathbf{\Sigma}_v^{-1}\hat{\mathbf{\Sigma}}_x\mathbf{\Sigma}_v^{-1} - \lambda\mathbf{\Sigma}_v^{-1}\right]\mathbf{b} = \mathbf{0} \ , \tag{4.3}$$

where λ is the (scalar) characteristic root. By using the relation $\mathbf{b}_* = \mathbf{\Sigma}_v^{-1}\mathbf{b}$ and we can represent

$$\left[\hat{\mathbf{\Sigma}}_x - \lambda\mathbf{\Sigma}_v\right]\mathbf{b}_* = \mathbf{0} \ . \tag{4.4}$$

If we take the smallest eigen-value and we take $\hat{\mathbf{\Sigma}}_{v,SIML}$, we have the $\mathbf{b}_{*,SIML}$.

Also under a set of regularity conditions we have that the smallest eigen-value

$$\lambda_1 \longrightarrow 0 \text{ (in probability)} \tag{4.5}$$

as $n \to \infty$. Then we may use the SILS (Separating Information Least Squares) method by solving

$$\hat{\mathbf{\Sigma}}_x \mathbf{b}_{*,SILS} = \mathbf{0}\ . \tag{4.6}$$

When $p = 2, q = 1$, $\boldsymbol{\beta} = (1, -\beta_2)'$ and $\mathbf{\Pi} = (\beta_2, 1)'$ ($\mathbf{\Sigma}_v = \sigma_v^2 \mathbf{I}_2$) in Example 1, then the SILS estimation becomes

$$\hat{\beta}_2 = \frac{\sum_{k=1}^{m_n} z_{1k} z_{2k}}{\sum_{k=1}^{m_n} z_{2k}^2}\ , \tag{4.7}$$

which is the regression coefficient of the first transformed variable on the second transformed variable in $\mathbf{z}_k$ $(= (z_{1k}, z_{2k})')$ $(k = 1, \cdots, m_n)$.

Some Remarks :
By using the SIML procedure, it is also straightforward to investigate the several structural relationships among trend variables at the same time. The SIML estimation can be defined by the smaller q $(\leq p)$ roots and the corresponding q $(\leq p)$ vectors of the characteristic equation. It may correspond to the standard situation in the statistical multivariate analysis except the fact that the classical multivariate analysis was based on the case when the observations are realizations of i.i.d. random variables without seasonality as well as non-stationarity in time series data sets.

We have done several simulations which are given in Appendix. The data length is 80, the number of simulations is 3000, $\alpha = 0.6$, and $m_n = n^{\alpha}$ in each case. In fugures *cor* $= 0.9$ means the true correlation coefficient $(= .9)$ among the trend components and *cor* is the SIML estimate. (vol1 is the correlation estimate based on the first differenced data and vol4 is the correlation estimate based on the seasonal differenced data with $s = 4$.)

When we have the basic model with the trend and noise components without the seasonal and cycle components, the optimal choice of $m_n = n^{\alpha}$ would be $\alpha = 0.8$, but it seems that the choice of $\alpha = 0.6$ would appropriate for the robustness of the results when we have seasonality as well as non-stationary trends.

5. Further Thoughts

Let $\mathbf{s}_i' = (s_{1i}, \cdots, s_{pi})$ be the vector of the seasonal components, which are independent of $\mathbf{v}_i$. Then we have

$$\mathbf{y}_i = \mathbf{s}_i + \mathbf{v}_i \tag{5.1}$$

where $\mathbf{s}_i$ are a sequence of seasonal components with $\mathcal{E}(\mathbf{w}_i^{(s)}) = \mathbf{0}$ and $\mathcal{E}(\mathbf{w}_i^{(s)} \mathbf{w}_i^{(s)'}) = \mathbf{\Sigma}_s$ with $\Delta_s \mathbf{s}_i = \mathbf{s}_i - \mathbf{s}_{i-s} = \mathbf{w}_i^{(s)}$ $(s \geq 2)$. We assume that $\mathbf{\Sigma}_s$ is positive definite and finite. When we transform the observed data by using the seasonal difference operator, we have

(5.2) $$(1-\mathcal{L}^s)\mathbf{y}_i = \mathbf{w}_i^{(s)} + (1-\mathcal{L}^s)\mathbf{v}_i \ .$$

For the seasonal transformation let $n = ms$ ($s \geq 2$) and transform $\mathbf{Y}_n$ to $\mathbf{Z}_n^{(s)}$ ($= (\mathbf{z}_k^{(s)'})$) by

(5.3) $$\mathbf{Z}_n^{(s)} = \mathbf{P}_n^{(s)}\mathbf{C}_n^{(s)-1}\left(\mathbf{Y}_n - \bar{\mathbf{Y}}_0\right)$$

where $\mathbf{C}_n^{(s)-1} = \mathbf{C}_m^{-1} \otimes \mathbf{I}_s$ and

(5.4) $$\mathbf{C}_m^{-1} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & 0 & \cdots \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}_{m\times m} ,$$

(5.5) $$\mathbf{P}_n^{(s)} = \mathbf{P}_m \otimes \mathbf{I}_s \ .$$

By using the spectral decomposition $\mathbf{C}_n^{(s)-1}\mathbf{C}_n^{(s)'-1} = (\mathbf{P}_m \otimes \mathbf{I}_s)(\mathbf{D}_m^{(s)} \otimes \mathbf{I}_s)(\mathbf{P}_m^{'} \otimes \mathbf{I}_s)$ and $\mathbf{D}_m^{(s)}$ is a diagonal matrix with

$$a_{km} = 2[1-\cos(\pi(\frac{2k-1}{2m+1}))] = 4\sin^2(\pi/2)[(2k-1)/(2m+1)] \ (k=1,\cdots,m)$$

with the s multiplicities.
Then we can develop the likelihood analysis as the case when $s=1$.

We consider the case when we have

(5.6) $$\mathbf{y}_i = \mathbf{x}_i + \mathbf{s}_i + \mathbf{v}_i$$

where $\mathbf{x}_i$ are a sequence of trend components and $\mathbf{s}_i$ are a sequence of seasonal components. When we transform the observed data by using the seasonal difference operator, we have

(5.7) $$(1-\mathcal{L}^s)\mathbf{y}_i = (1+\mathcal{L}+\cdots+\mathcal{L}^{s-1})\Delta\mathbf{x}_i + (1-\mathcal{L}^s)\mathbf{s}_i + (1-\mathcal{L}^s)\mathbf{v}_i \ .$$

Then there can be alternative possibilities of transformations of $\mathbf{Y}_n$, but we may use $\mathbf{Z}_n^{(s)*}$ ($= (\mathbf{z}_k^{(s)*'})$) by

(5.8) $$\mathbf{Z}_n^{(s)*} = \mathbf{P}_n\mathbf{C}_n^{(s)-1}\left(\mathbf{Y}_n - \bar{\mathbf{Y}}_0\right) \ .$$

It seems that the likelihood analysis can be extended by using the orthogonality relations with components of $p_{jk}^{(n)}$. Let

(5.9) $$\mathbf{\Lambda}_n = (\lambda_{jk}) = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ \vdots & 1 & 0 & \cdots & 0 \\ 1 & 1 & 1 & \cdots & 0 \\ 0 & \cdots & 1 & 1 & 0 \\ 0 & \cdots & 1 & \cdots & 1 \end{pmatrix}_{n\times n}$$

and

(5.10) $$\mathbf{P}_n^* = (p_{jk}^*) = \mathbf{P}_n \mathbf{\Lambda}_n \ .$$

Then

(5.11) $$\sum_{j=1}^{n} p_{kj}^* p_{k',j}^* \sim \delta(k,k') \frac{\sin^2 \frac{\pi}{2}\frac{2k-1}{2n+1}s}{\sin^2 \frac{\pi}{2}\frac{2k-1}{2n+1}} \sim s^2 \delta(k,k')$$

when $k/n \to 0$.

We consider the general case when we have

(5.12) $$\mathbf{y}_i = \mathbf{x}_i + \mathbf{s}_i + \mathbf{c}_i + \mathbf{v}_i$$

where $\mathbf{x}_i$ are a sequence of trend components, $\mathbf{s}_i$ are a sequence of seasonal components and $\mathbf{c}_i$ are a sequence of cycle components. When we transform the observed data by using the seasonal difference operator, we have

(5.13) $$(1-\mathcal{L}^s)\mathbf{y}_i = (1+\mathcal{L}+\cdots+\mathcal{L}^{s-1})\Delta \mathbf{x}_i + \mathbf{w}_i^{(s)} + (1-\mathcal{L}^s)\mathbf{c}_i + (1-\mathcal{L}^s)\mathbf{v}_i$$

with $(1-\mathcal{L}^s)\mathbf{s}_i = \mathbf{w}_i^{(s)}$.
Then again there are alternative possibilities of transformations of $\mathbf{Y}_n$, but we may use $\mathbf{Z}_n^{(s)*}$ $(= (\mathbf{z}_k^{(s)*'}))$ by

(5.14) $$\mathbf{Z}_n^{(s)*} = \mathbf{P}_n \mathbf{C}_n^{(s)-1}\left(\mathbf{Y}_n - \bar{\mathbf{Y}}_0\right) \ .$$

It seems that the likelihood analysis can be extended in the second case by using the orthogonality relations with components of $p_{jk}^{(n)}$. Let

(5.15) $$\mathbf{P}_n^{**} = (p_{jk}^{**}) = \mathbf{P}_n(\mathbf{C}_m^{-1} \otimes \mathbf{I}_s) \ .$$

Then

(5.16) $$\sum_{j=1}^{n-s} p_{kj}^{**} p_{k',j}^{**} \sim 4\sin^2[\frac{\pi}{2}\frac{2k-1}{2n+1}s]$$

when $k = k'$. When $k \neq k'$,

(5.17) $$\sum_{j=1}^{n-s} p_{kj}^{**} p_{k',j}^{**} \sim O(\frac{1}{n}) \ .$$

Also for any finite l we have

(5.18) $$\sum_{j=1}^{n-s-l} p_{kj}^{**} p_{k',j+l}^{**} \sim O(\frac{1}{n})$$

we find that the covariances of the transformed cycle compotents $(\mathbf{c}_k^*) = \mathbf{P}_n^{**}(\mathbf{c}_j)$ are

(5.19) $$Cov(\mathbf{c}_j^*, \mathbf{c}_{j'}^*) = \rho^{|j-j'|} \times O(\frac{1}{n})$$

with $|\rho| < 1$.

References

[1] Anderson, T.W. (1984), "Estimating Linear Statistical Relationships," *Annals of Statistics*, 12, 1-45.

[2] Anderson, T.W. (2003), *An Introduction to Statistical Multivariate Analysis*, 3rd Edition, John-Wiley.

[3] Fuller, W. (1987), *Measurement Error Models*, John-Wiley.

[4] Kitagawa, G. (2010), *Introduction to Time Series Modeling*, CRA Press.

[5] Kunitomo, N. and S. Sato (2008, 2013), "Separating Information Maximum Likelihood Estimation of Realized Volatility and Covariance with Micro-Market Noise," Discussion Paper CIRJE-F-581, Graduate School of Economics, University of Tokyo, (http://www.e.u-tokyo.ac.jp/cirje/research/dp/2008), North American Journal of Economics and Finance, Elsevier, 26, 282-309.

[6] Kunitomo, N. and S. Sato (2011), "The SIML Estimation of Realized Volatility of Nikkei-225 Futures and Hedging Coefficient with Micro-Market Noise," *Mathematics and Computers in Simulation*, Vol.81, 1272-1289.

[7] 国友直人 (2011),「構造方程式モデルと計量経済学」, 朝倉書店。

[8] 国友直人 (2014), 「計測誤差と統計学」, 日本統計学会誌, Vol. 43-2, 157-183.

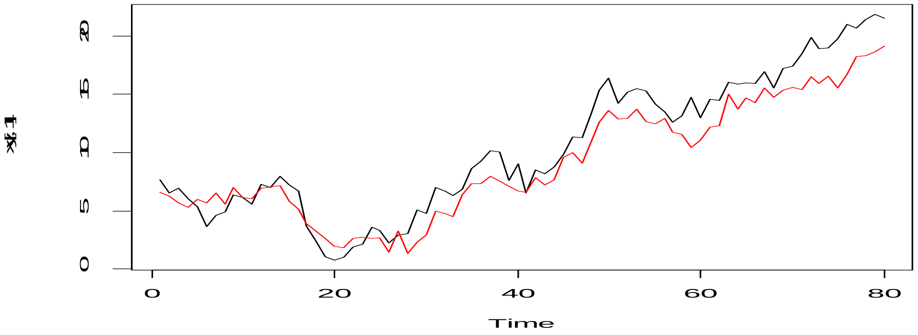
[9] 国友直人 (2015), 「(応用を目指す) 数理統計学」, 朝倉書店 (近刊)。

Appendix: Some Figures

Quarterly Data alpha=0.6
n=80 nsim=3000

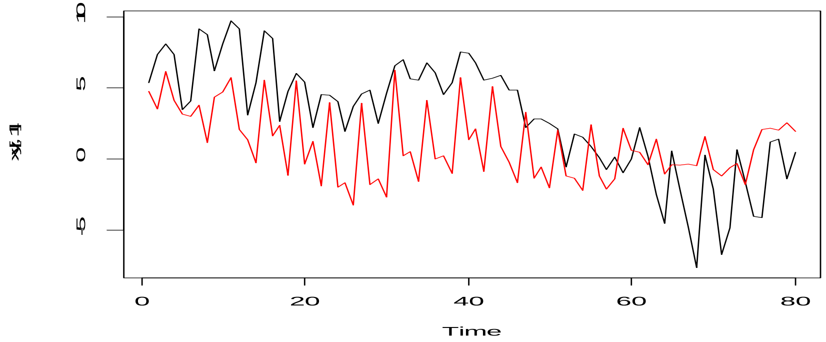
With Noise, no Seaonality

cor=0.9	cor	vol4	vol1
mean	0.852227	0.732646	0.491137
SD	0.087745	0.075951	0.094544
cor=0.0	cor	vol4	vol1
mean	0.007154	0.003388	0.001042
SD	0.277539	0.16811	0.119081



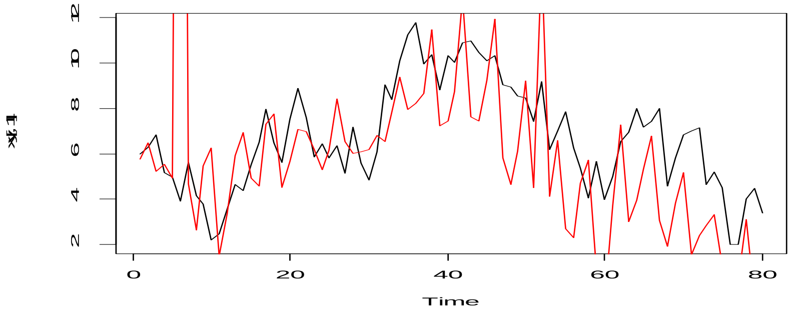
With noise and seasonality

cor=0.9	cor	vol4	vol1
mean	0.805092	0.662977	0.133043
SD	0.118186	0.088369	0.295366
cor=0.0	cor	vol4	vol1
mean	−0.00689	2.59E−03	0.004784
SD	0.278001	1.62E−01	0.287285



With noise, seasonality and outlier

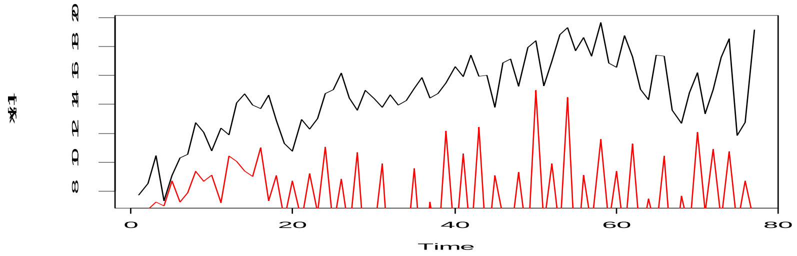
cor=0.9	cor	vol4	vol1
mean	0.584384	0.308292	0.063187
SD	0.209764	0.13995	0.217537
cor=0.0	cor	vol4	vol1
mean	0.005651	0.004289	−0.00455
SD	0.275338	0.128497	0.210767



With noise and changing seasonality

cor=0.9	cor	vol4	vol1
mean	0.672386	0.344135	0.033535
SD	0.196482	0.184532	0.190742

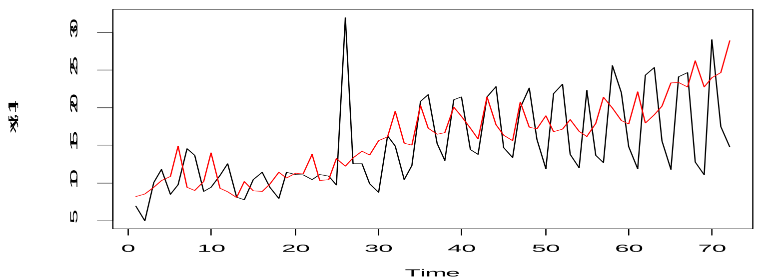
cor=0.0	cor	vol4	vol1
mean	0.002128	0.001937	0.002259
SD	0.283569	0.148525	0.183541



With noise, seasonality,outlier and changing seasonality

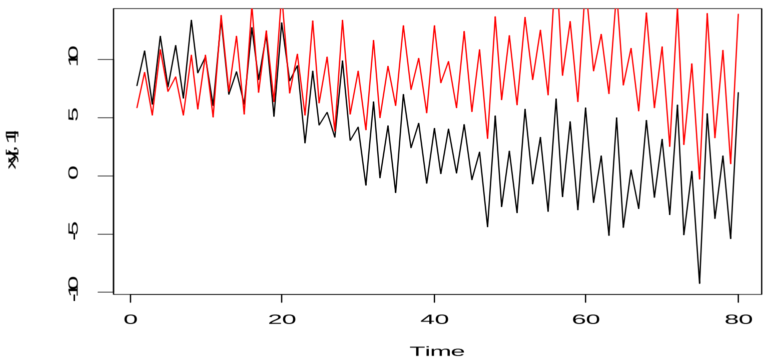
cor=0.9	cor	vol4	vol1
mean	0.491676	0.184133	0.019022
SD	0.244269	0.152994	0.160183

cor=0.0	cor	vol4	vol1
mean	−0.00467	0.000966	0.002659
SD	0.283498	0.133407	0.162206



With noise and seasonality

cor=0.0	cor	vol4	vol1
mean	0.044796	0.083241	0.673487
SD	0.282269	0.165334	0.237343



地域統計の季節調整問題

川崎能典（統計数理研究所）
国友直人（東京大学大学院経済学研究科）

科学研究費(基盤研究(A))研究集会
「経済統計・政府統計の基礎と応用」
2015年1月30日(金)
東京大学大学院経済学研究科・小島ホール

問題の提起

- 都道府県別時系列の季節調整
- 都道府県別時系列を集計して定義される地域別時系列(cf. 東北ブロック)の季節調整
- 両者は独立した統計処理
- 都道府県別季節調整済系列の合算と、地域別時系列に対する季調結果の不一致
- 両者の季調済系列間に何らかの整合性を持たせる必要があるのでは?

自然な制約～2変量を例に

- 第1県の時系列と季調 $y_{1t} = n_{1t} + s_{1t}$
- 第2県の時系列と季調 $y_{2t} = n_{2t} + s_{2t}$
  - n_{*t}はnon-seasonal partを、s_{*t}はseasonal partを表す。以下のN_t, S_tも同じ。
- 合算系列(より上位の地域ブロック統計に相当) $Y_t = y_{1t} + y_{2t}$ の季調 $Y_t = N_t + S_t$
- 自然な制約(期待) $N_t \approx n_{1t} + n_{2t}$, これを裏(除去したいもの)から見ると$S_t \approx s_{1t} + s_{2t}$

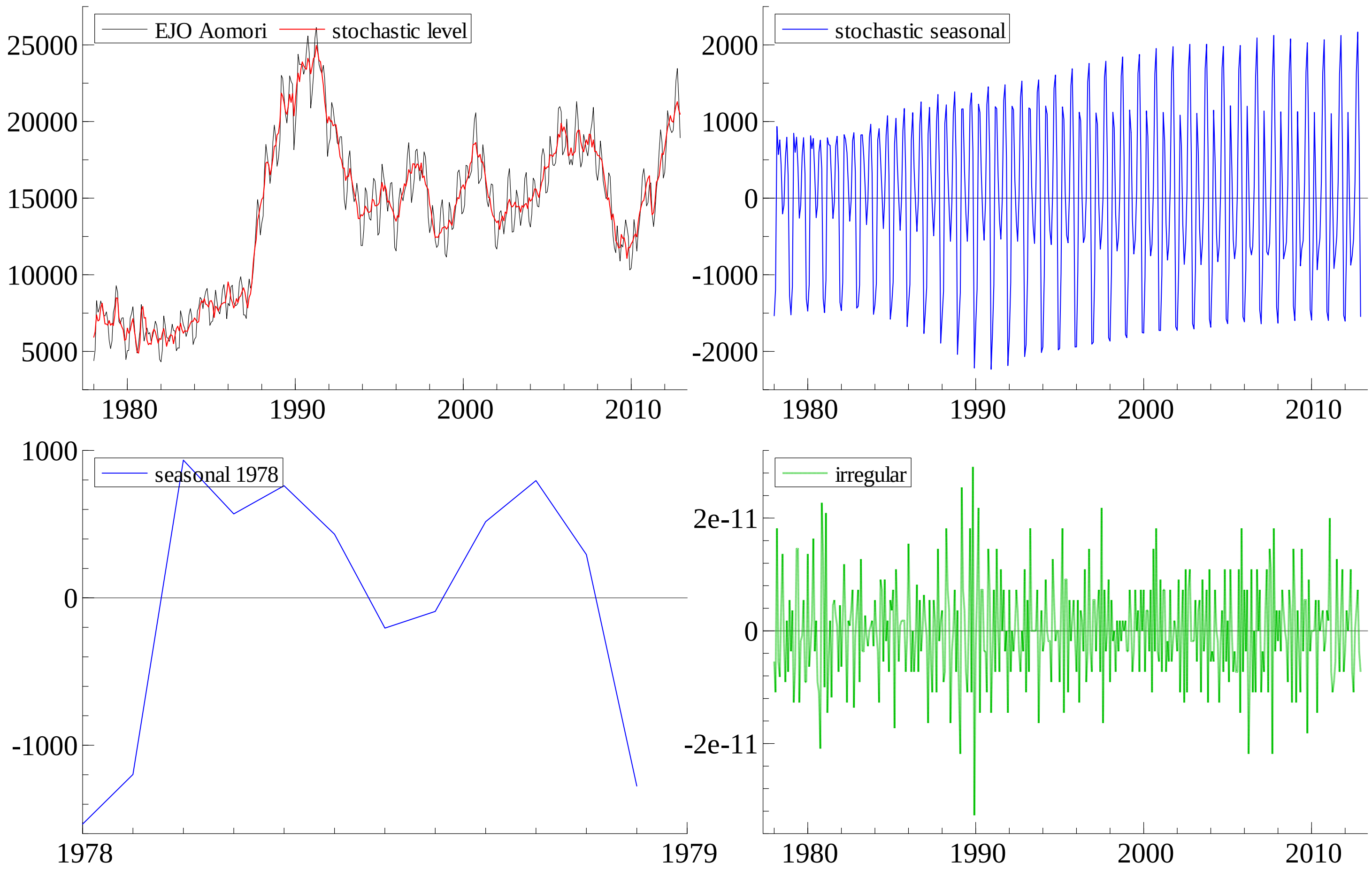
自然な制約(続)

- 例えば$N_t = TC_t + I_t$と分けて考えた時、$TC_t \approx TC_{1t} + TC_{2t}$を課すモデリングは可能だが、$I_t \approx I_{1t} + I_{2t}$は考えづらい
  - 前者は状態変数だが、後者は観測ノイズ
- $S_t \approx s_{1t} + s_{2t}$が成り立つ限りは$N_t \approx n_{1t} + n_{2t}$が期待できるので、集計系列(地域ブロックのデータ)の季節成分だけに、下位系列との整合性を課すモデリングをまず試す。

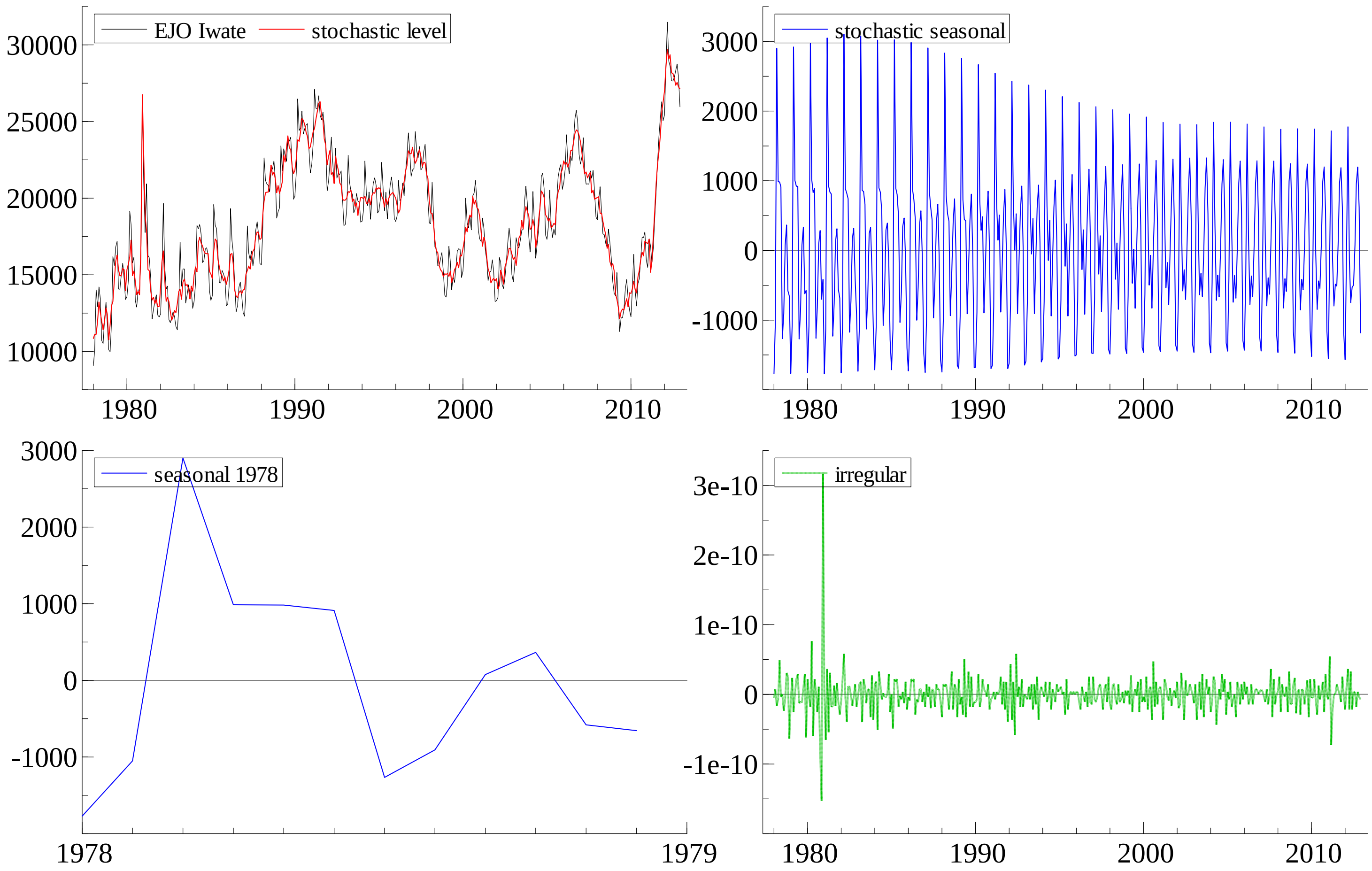
データ

- 職業安定業務統計 有効求人数(原系列)
  - 厚生労働省
- 青森県、岩手県、その合算を考える
- 1978年1月～2012年12月(時系列長420)

青森県の一変量解析

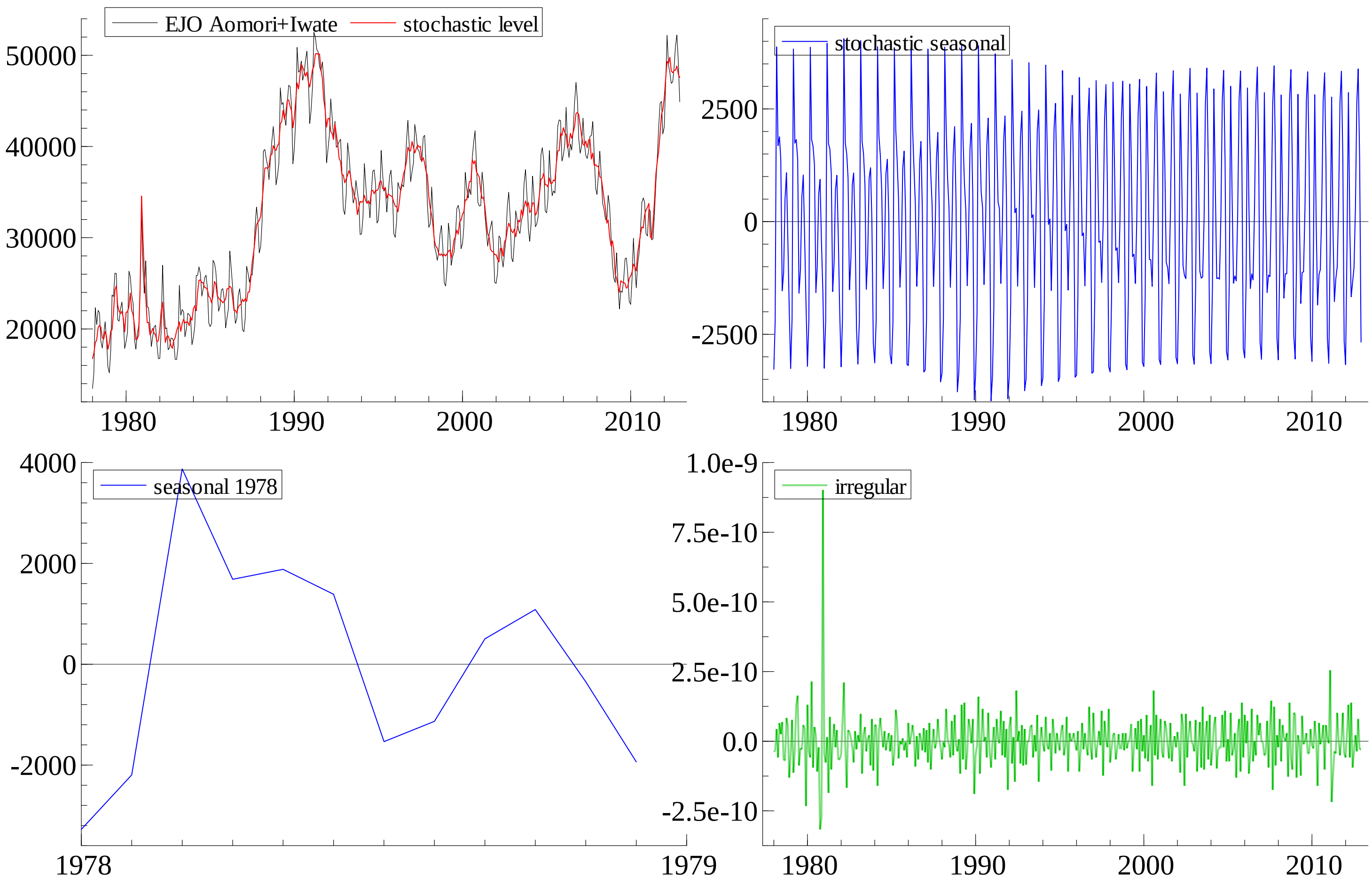

EJO Aomori
stochastic level
25000
20000
15000
10000
5000
1980
1990
2000
2010
stochastic seasonal
2000
1000
0
-1000
-2000
1980
1990
2000
2010
seasonal 1978
1000
0
-1000
1978
1979
irregular
2e-11
0
-2e-11
1980
1990
2000
2010

岩手県の一変量解析



EJO Iwate
stochastic level
30000
25000
20000
15000
10000
1980
1990
2000
2010
stochastic seasonal
3000
2000
1000
0
-1000
1980
1990
2000
2010
seasonal 1978
3000
2000
1000
0
-1000
1978
1979
irregular
3e-10
2e-10
1e-10
0
-1e-10
1980
1990
2000
2010

青森+岩手合算系列の一変量解析



多変量状態空間モデル

観測方程式

$$Y_t = y_{1t} + y_{2,t} = \mu_t + s_{1t} + s_{2t} + \epsilon_t$$

$$y_{1t} = \mu_{1,t} + s_{1,t} + \epsilon_{1,t}$$

$$y_{2t} = \mu_{2,t} + s_{2,t} + \epsilon_{2,t}$$

状態方程式

システムノイズ、観測ノイズの分散共分散行列は対角を仮定

$$\mu_t = \mu_{t-1} + \eta_t, \mu_{i,t} = \mu_{i,t-1} + \eta_{i,t} \ (i = 1,2)$$

$$s_{i,t} = -\sum_{j=1}^{s-1} s_{i,t-j} + \omega_{i,t} \ (i = 1,2)$$

(暫定的な)解析結果

- 個別季節調整: 個別の対数尤度の総和
$$\hat{\ell} = -10256.17$$
- 多変量季節調整
$$\hat{\ell} = -10543.26$$
- まだ多変量状態空間モデルのパラメータ推定(最適化)がうまくいっておらず、暫定的な結果だが、多変量モデルは奏功していない…。
  - AIC流のバイアス補正で逆転できるような尤度の差ではない

考察

- 集計系列と、基礎となる個別系列間で整合性ある季節調整を考えたいが、観測ベクトル自体はおのずと完全多重共線状態。
- 個別モデルを束ねて形式的に多変量状態空間モデルとして推定してもパラメータが非現実的に大きくなりうまくいかない。完全多重共線かつ自由度が高すぎ。
- 制約(ファクター型の構造)が足りない?

考察(続):まとめに代えて

- システムノイズの次元落ちに関する丁寧な探索が必要だと、手間がかかりすぎて実用的とは言えない
- 不規則変動部分は(状態変数でないので)ファクター流にモデル化しづらい
- 観測ノイズは対角を仮定しているが推定してみると対角とは言いがたい。直交化は可能だが季節調整の解釈に問題は生じないか?
  - Nishio (2002) FTSM3