Extending causality-related interdependence measures

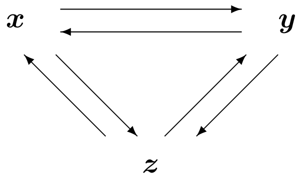
Yuzo Hosoya
Tohoku University
(emeritus professor)

Waseda International Symposium
23rd October, 2018

Outline

Time series interdependence

- Simple relation: Relation focused on a pair of series $x(t)$ and $y(t)$ regardless of third series.
- Partial relation: Relation between $x(t)$ and $y(t)$ where the presence of a third series $z(t)$ is explicitly taken into consideration; the one way effect component of $z(t)$ is eliminated from $x(t)$ and $y(t)$.



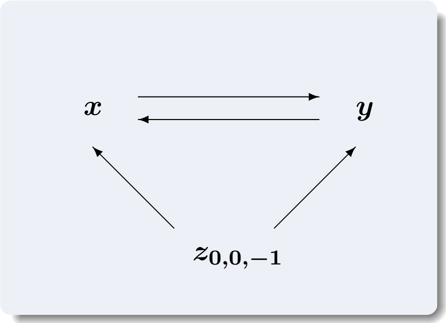


Figure 1. Characterizing causal interdependence quantitatively.

Introduction to empirical causal analysis

- The Wiener-Granger causality of a series $\{y(t)\}$ to another $\{x(t)\}$ is defined by the improvement of the prediction error by addition of $\{y(t)\}$ as predictor.
- Sims (1980)'s innovation accounting: Eliciting the variation componet of $\{x(t)\}$ due to the intrinsic innovation of $\{y(t)\}$ in the framework of the vector MA representation.
- Geweke (1982): Prediction improvement by addition of $\boldsymbol{H}\{y(t-j), j \in \mathbb{Z}^+\}$.
- Hosoya (1991): $\boldsymbol{H}\{y_{0,-1}(t-j), j \in \mathbb{Z}^+\}$ where $y_{0,-1}(t)$ denotes the projection residual of $y(t)$ onto $\boldsymbol{H}\{x(t+1-j), y(t-j), j \in \mathbb{Z}^+\}$.

All of them characterize time series causality by quantifying what constitutes the additional information of $\{y(t)\}$ for prediction improvement of $\{x(t)\}$?

The Szegö condition for multivariate spectrum

- Let $\{x(t), y(t), z(t), t \in \mathbb{Z}\}$ be a zero mean jointly covariance-stationary process, where x, y, z are p_1, p_2, p_3-vectors.
- The process is assumed to have $p \times p$ spectral density matrix $f(\lambda)$, where $p = p_1 + p_2 + p_3$.
- $\int_{-\pi}^{\pi} \ln \det f(\lambda) d\lambda > -\infty \iff f(\lambda)$ satisfying the Szegö condition.
  If the condition is satisfied, the density is canonically factorizable

$$f(\lambda) = \frac{1}{2\pi} \Lambda(e^{-i\lambda}) \Lambda(e^{-i\lambda})^*,$$

  where $\Lambda(z)$ is analytic and of full rank inside the unit disc, having the expansion

$$\Lambda(z) = \sum_{j=0}^{\infty} \Lambda[j] z^j$$

Predicition-error formula (1)

- Denote the covariance matrix of the one step ahead prediction error of the process $\{x(t), y(t), z(t)\}$ by $\Sigma^{\dagger}$ where

$$\Sigma^{\dagger} = Cov\{x_{-1,-1,-1}(t), y_{-1,-1,-1}(t), z_{-1,-1,-1}(t)\}$$

  and $x_{-1,-1,-1}(t)$ indicates the projection residual of $x(t)$ onto the past $x(t), y(t), z(t)$; namely, onto $H\{x(t-j), y(t-j), z(t-j), j \in \mathbb{Z}^{+}\}$.
- If the Szegö condition holds, $\Sigma^{\dagger}$ is positive definite. item

Predection-error formula (2)

$\mathbf{\Lambda}(z)$ is canonical if and only if

$$\begin{aligned}\det\{\mathbf{\Lambda}(0)\mathbf{\Lambda}(0)^*\} &= \det \mathbf{\Sigma}^\dagger \\ &= (2\pi)^p \exp\{\frac{1}{2\pi}\int_{-\pi}^{\pi} \log\det f(\lambda)d\lambda\}.\end{aligned}$$

For related expositions, see Szegö(1939), Helson and Lowdenslager (1958), Rozanov (1967), Hannan (1970), Brockwell and Davis (1987).

Monograph by Hosaya et al.

Hosoya, Y., Oya, K., Takimoto, T. and Kinoshita, R. (2017):

Characterizing Interdependencies of Multiple Time Series: Theory and Application, JSS Research Series in Statistics, Springer.

Chapter 1:	Introduction
Chapter 2:	The Meaures of One-way Effect, Reciprocity and Association
Chapter 3:	Representation of Partial Meaures
Chapter 4:	Inference Baed on the Vector Autoregressive and Moving Average Model
Chapter 5:	Inference on Changes in Interdependence Measures
Appendix:	Technical Supplements.

The partial measures of interdependence

Section 3 of Hosoya et al. (2017) is for exposition of the overall as well as frequency-wise measures of one-way effect, reciprocity and association:

- The equality holds between the partial OMO and FMO:

$$\boldsymbol{PM_{y\to x:z} = \frac{1}{2\pi}\int_{-\pi}^{\pi} PM_{y\to x:z}(\lambda)d\lambda;}$$

- In view of the definitions of the respective measures and the corresponding equality for $\boldsymbol{PM_{x\to y:z}}$, we have

$$\begin{aligned} \boldsymbol{PM_{x,y:z}(\lambda)} &= \boldsymbol{PM_{x\to y:z}(\lambda) + PM_{x.y:z}(\lambda) + PM_{y\to x:z}(\lambda),} \\ \boldsymbol{PM_{x,y:z}} &= \boldsymbol{PM_{x\to y:z} + PM_{x.y:z} + PM_{y\to x:z}.} \end{aligned}$$

Stationary processes in the plane (1)

1. The Szegö theory is extensible to doubly parametrized p-vector stationary processes $\{w(s,t), s,t \in \mathbb{Z}\}$.
2. The extension provides the prediction error formula for linear prediciton of $w(0,0)$ by $H\{w(s,t), s,t \in \mathbb{S}\}$ where $\mathbb{S}$ is a half plane.

References:

Helson and Lowdenslager (1958). "Prediction theory and Fourier series in several variables" , Acta Math. (1961), **106**, 175-213.

Szegö, B.G. (1939). *Orthogonal Polynomials*, Colloquium publications, American Mathematical Society, **23**, New York.

Wiener, N. (1949). *The Extrapolation, Interpolation and Smoothing of Stationary Time Series*, Wiley, New York.

Whittle, P. (1954). On Stationary Processes in the Plane, *Biometrika*, **41**, 434-449.

Example of half plane in $\mathbb{Z}^2$ (the lattice points)

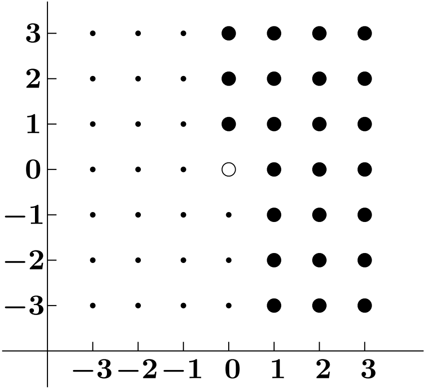


Figure 2: Half planes: Each of the sets of the large dots (•) and the small dots (·) indicates a half plane respectively.

Stationary processes in the plane (2)

A half plane $S \equiv S(0,0)$ in the set of lattice points $\mathbb{Z}^2$ is formally defined by

$$S = \{(s,t), s = 0, t \geq 1\} \cup \{(s,t), s \geq 1, t \in \mathbb{Z}\}.$$

Half plane introduces a total order structure $(s_0, t_0) < (s_1, t_1)$ if $(s_1, t_1) \in S(s_0, t_0)$ in the set of lattice points and it is used for defining prediction and allied concepts.

- Suppose that a p-vector stationary series $\{w(s,t), (s,t) \in \mathbb{Z}^2\}$ has a partition $w(s,t) = \{u(s,t)', v(s,t)'\}'$ where u and v are p_1 and p_2 vectors.
- Suppse that the spectral density of the process $\{w(s,t)\}$ is given by $h(\lambda_1, \lambda_2)$ which has the partition:

$$h(\lambda_1, \lambda_2) = \begin{bmatrix} h_{11}(\lambda_1, \lambda_2) & h_{12}(\lambda_1, \lambda_2) \\ h_{21}(\lambda_1, \lambda_2) & h_{22}(\lambda_1, \lambda_2) \end{bmatrix}, \ -\pi < \lambda_1, \lambda_2 \leq \pi,$$

where $h_{11}(\lambda)$ is the $p_1 \times p_1$ spectral density of $\{u(s,t)\}$.

Stationary processes in the plane (3)

Assumption 1:

The spectral density matrix $h(\lambda_1, \lambda_2)$ satisfies the Szegö condition

$$\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \log\det h(\lambda_1, \lambda_2) d\lambda_1 d\lambda_2 > -\infty.$$

It follows from the assumption that $h(\lambda_1, \lambda_2)$ has a factorization

$$h(\lambda_1, \lambda_2) = \frac{1}{2\pi}\Gamma^{\dagger}(e^{-i\lambda_1}, e^{-i\lambda_2})\Gamma^{\dagger}(e^{-i\lambda_1}, e^{-i\lambda_2})^*,$$

where $\Gamma^{\dagger}$ is a function defined on $\mathbb{C}^2$ and square-integrable on the torus, having the expansion by means of the half plane S

$$\Gamma^{\dagger}(y, z) = \Gamma^{\dagger}[0, 0] + \sum_{(i,j)\in S} \Gamma^{\dagger}[i, j] y^i z^j$$

Stationary processes in the plane (4)

- Define the Hilbert subspace by

  $$H^{\dagger}\{u(s_1,t_1),v(s_2,t_2)\} \equiv H\{u(s_1-S,t_1),v(s_2-S,t_2-S)\}$$

  where $u(s_1-S,t_1-S) \equiv \{u(s_1-i_1,t_1-j_1);(i_1,j_1)\in S\}$.
- Denote by $u_{-1,-1}(s,t)$, $u_{-1,\cdot}(s,t)$ and $u_{-1,0}(s,t)$ respectively the projection residuals of $u(s,t)$ onto $H^{\dagger}\{u(s,t-1),v(s,t-1)\}$, $H^{\dagger}\{u(s,t-1)\}$ and $H^{\dagger}\{u(s,t-1),v(s,t)\}$, whereas denote by $u'_{-1,-1}(s,t)$ the projection residuals of $u(s,t)$ onto $H^{\dagger}\{u(s,t-1),v_{0,-1}(s,t-1)\}$.

We can choose $\Gamma^{\dagger}(y,z)$ so that it has the maximality property:

$$\det\{\Gamma^{\dagger}(0,0)\Gamma^{\dagger}(0,0)^*\} = \det\Sigma$$
$$= (2\pi)^p \exp\{\frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\log\det h(\lambda_1,\lambda_2)d\lambda_1 d\lambda_2\};$$

see Helson and Lowdenslager (1958) Theorem 12 in p.199.

Stationary processes in the plane (5)

Denote the covariance matrix of the one step ahead prediction error of the process $\{u(s,t), v(s,t)\}$ is given in partitioned form by

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

where $\Sigma = Cov\{u_{-1,-1}(t), v_{-1,-1}(t)\} = \Gamma^{\dagger}(0,0)\Gamma(0,0)^*$.
Set $\Sigma_{22:1} \equiv \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$. Denote by $h_{11}(\lambda_1, \lambda_2)$ the $p_1 \times p_2$ upper diagonal block of $h(\lambda_1, \lambda_2)$ and denote by $h_{11:2}(\lambda_1, \lambda_2)$ the $p_1 \times p_1$ matrix defined by

$$h_{11:2}(\lambda_1, \lambda_2) = h_{11}(\lambda_1, \lambda_2) - \tilde{h}_{12}(\lambda_1, \lambda_2)\Sigma_{22:1}^{-1}\tilde{h}_{12}(\lambda_1, \lambda_2)^*.$$

where $\tilde{h}_{12}(\lambda_1, \lambda_2) =$

$$h_{1\cdot}(\lambda_1, \lambda_2)\Gamma^{\dagger}(e^{-i\lambda_1}, e^{-i\lambda_2})\Gamma^{\dagger}(0,0)^*(-\Sigma_{21}\Sigma_{11}^{-1}, I_{p_1})^*.$$

Stationary processes in the plane (6)

Define the (overall) measure of one-way effect in the plane from v to u by

$$M^{\dagger}_{v\to u} = \log[\det\{Cov(u_{-1,\cdot}(s,t))\}/\det\{Cov(u'_{-1,-1}(s,t))\}].$$

Theorem 1:

1. The one-step ahead time-wise prediction is given by

$$\det Cov\{w_{-1}(s,t)\}$$
$$= (2\pi)^p \exp\left\{\frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\log\det h(\lambda_1,\lambda_2)d\lambda_1 d\lambda_2\right\}.$$

2. The one-way effect is repsesented by

$$M^{\dagger}_{v\to u}(s,t) = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\log\left[\frac{\det h_{11}(\lambda_1,\lambda_2)}{\det h_{11:2}(\lambda_1,\lambda_2)}\right]d\lambda_1 d\lambda_2.$$

Whittle's likelihood method (1)

For $\{w(s,t), s = 1,\ldots,m; t = 1,\ldots,n\}$ a second-order stationary series in the plane, define the empirical spectral density by

$$J_{m,n}(\lambda_1,\lambda_2) \equiv w_{m,n}(\lambda_1,\lambda_2)w_{m,n}(\lambda_1,\lambda_2)^*, \ -\pi < \lambda_1,\lambda_2 \leq \pi,$$

where

$$w_{m,n}(\lambda_1,\lambda_2) \equiv \frac{1}{2\pi(mn)^{1/2}}\left\{\sum_{s=1}^{m}\sum_{t=1}^{n} w(s,t)\exp(is\lambda_1 + it\lambda_2)\right\}.$$

Whittle's likelihood method (2)

Suppose that the spectral density of the process $\{w(s,t)\}$ depends on the unknown model parameter ψ and is given by $h(\lambda_1, \lambda_2; \psi)$ where $\psi \in \Psi$ and the parameter space Ψ is assuned a compact subset of $\mathbb{R}^s$ with non-empty interior. In integral form the Whittle log-likelihood function is given by

$$\begin{aligned} \bar{L}_{mn}(\psi) \equiv & -mn \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} [\log \det h(\lambda_1, \lambda_2; \psi) \\ & + tr\{h^{-1}(\lambda_1, \lambda_2; \psi) J_{mn}(\lambda_1, \lambda_2)\}] \, d\lambda_1 d\lambda_2. \end{aligned}$$

The function $\bar{L}_n(\psi)$ is said the whittle log-likelihood function [the approximation was originally proposed by Whittle (1952, 1953) for scalar and multiple stationary processes respectively. The likelihood $\bar{L}_{mn}(\psi)$ was given by Whittle (1954) for scalar processes.

Whittle's likelihood method (3)

Assumption 2:

1. $\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\log deth(\lambda_1,\lambda_2;\psi)d\lambda_1 d\lambda_2$ is differentiable;
2. At almost all point of $(\lambda_1,\lambda_2), h(\lambda_1,\lambda_2;\psi)^{-1}$ is differentiable with respect to ψ.

The partial derivatives are denoted, respectively, by

$$H_j(\psi)=\frac{\partial\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\log deth(\lambda_1,\lambda_2;\psi)d\lambda_1 d\lambda_2}{\partial\psi_j}$$

and

$$h_j(\lambda_1,\lambda_2;\psi)=\frac{\partial h^{-1}(\lambda_1,\lambda_2;\psi)}{\partial\psi_j}$$

and h_j is assumed to be measurable with respect to ψ a.e. λ_1,λ_2.

Whittle's likelihood method (4)

1. The notations $H(\psi)$ and $tr\{h(\lambda_1, \lambda_2; \psi)f(\lambda_1, \lambda_2)\}$ represent, respectively, the s-vectors whose j-th elements are $H_j(\psi)$ and $tr\{h_j(\lambda_1, \lambda_2; \psi)f(\lambda_1, \lambda_2)\}$.
2. Define $S_{nj}(\psi)$ by

$$S_{nj}(\psi) = H_j(\psi) + \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} tr\{h_j(\lambda_1, \lambda_2; \psi)I_n(\lambda_1, \lambda_2)\}d\lambda_1 d\lambda_2, \quad j = 1, \cdots, s$$

   and let $S_n(\psi)$ be the vector $\{S_{nj}(\psi)\}$.
3. A value $\hat{\psi}_n$ satisfying $S_n(\hat{\psi}_n) = 0$ is said to be a maximum Whittle likelihood (MWL) estimate of ψ.

Locally stationary processes and asymptotics (1)

Szegö 's prediction error formula holds approximately to the locally stationary processes introduced by Dahlhaus (1996). His approach is able to be adaptably employed for evaluating the measures of interdependence for the multivariate locally stationary processes.

The locally stationary p-vector process $\{w(t, T)\}$ is represented by

$$w(t, T) = \int_{-\pi}^{\pi} \exp(i\lambda t)\Gamma_{t,T}(\lambda) d\xi \quad \text{for} \quad t = 1, 2, \ldots, T,$$

where $w(t, T) = 0$ for $t \leq 0$.

Locally stationary processes and asymptotics (2)

Assumption 3:

1. $\xi(\lambda), -\pi < \lambda \leq \pi$, is a p-vector random process with cumulants

$$cumulant\{d\xi_{\alpha_1}(\lambda_1), \ldots, d\xi_{\alpha_k}(\lambda_k)\} = \nu\left(\sum_{j=1}^{k} \lambda_j\right) h_{\alpha_1,\ldots,\alpha_k}(\lambda_1, \ldots, \lambda_k) d\lambda_1 \cdots \lambda_k$$

   where $\nu(\lambda)$ is the delta function (mod 2π).

2. $h_{\alpha_1}(\lambda_1) = 0$, and $h_{\alpha_1,\alpha_2}(\lambda_1, \lambda_2) = 1(\alpha_1 = \alpha_2)$, and $|h_{\alpha_1,\ldots,\alpha_k}(\lambda_1, \ldots, \lambda_k)|$ is bounded.

3. $\Gamma_{t,T}(\lambda)$ is a $p \times p$ matrix for which there exists a constant K and a 2π periodic $p \times p$ matrix function such that for all T

$$\sup_{t,\lambda} |\Gamma_{t,T}(\lambda) - \Gamma(t/T, \lambda)| \leq KT^{-1}$$

   and $\Gamma(\tau, \lambda)$ is continuous in τ.

Locally stationary processes and asymptotics (3)

Denote the time-varying spectral density matrix for each $\tau, 0 \leq \tau \leq 1$, by

$$h(\tau, \lambda) = \frac{1}{2\pi}\Gamma(\tau, \lambda)\Gamma(\tau, \lambda)^*.$$

Suppose the p-vector process w is partitioned as $w(t|T) = \{u(t|T)', v(t|T)'\}'$ where u and v are p_1 and p_2 vectors $(p = p_1 + p_2)$. Let the Hilbert subspace $H_T\{u(t|T), v(s|T)\}$ defined by

$H_T\{u(t|T), v(s|T)\} \equiv H\{u(i|T), 1 \leq i \leq t; v(j|T), 1 \leq j \leq s\}$.

Denote by $u_{-1,-1}(t|T), u_{-1}(t|T)$ and $u_{-1,0}(t|T)$ respectively the projection residuals of $u(t|T)$ onto $H_T\{u(t-1|T), v(t-1|T)\}, H_T\{u(t-1|T)\}$ and $H_T\{u(t-1|T), v(t|T)\}$, whereas denote by $u'_{-1,-1}(t|T)$ the projection residuals of $u(t|T)$ onto $H_T\{u(t-1|T), v(t-1|T)\}$.

Locally stationary processes and asymptotics (4)

Define the time-varying (overall) measure of one-way effect from v to u by

$$M_{v\to u}(t|T) = \log[\det\{Cov(u_{-1,\cdot}(t|T))\}/\det\{Cov(u'_{-1,-1}(t|T))\}]$$

and the average measure by

$$\bar{M}_{v\to u}(T) \equiv \frac{1}{T}\sum_{t=1}^{T} M_{v\to u}(t,T).$$

Locally stationary processes and asymptotics (5)

Assumption 4:

For each $\tau, 0 \leq \tau \leq 1$, the spectral density $h(\tau,\lambda)$ satisfies the Szegö condition and has a canonical factorization for each u such that

$$h(\tau,\lambda) = \frac{1}{2\pi}\bar{\Gamma}(\tau, e^{-i\lambda})\bar{\Gamma}(\tau, e^{-i\lambda})^*$$

where $\bar{\Gamma}$ is bounded from below and has uniformly bounded derivative $\partial^2\bar{\Gamma}/\partial u \partial\lambda$.

Set:

- $\Sigma(\tau) = \bar{\Gamma}(\tau, 0)\bar{\Gamma}(\tau, 0)^*$
- $\Sigma_{22:1}(\tau) \equiv \Sigma_{22}(\tau) - \Sigma_{21}(\tau)\Sigma_{11}^{-1}(\tau)\Sigma_{12}(\tau)$.

Locally stationary processes and asymptotics (6)

- Denote by $h_{11}(\tau, \lambda)$ the $p_1 \times p_1$ upper diagonal block of $h(\tau, \lambda)$.
- Denote by $h_{11:2}(\tau, \lambda)$ the $p_1 \times p_1$ matrix defined by

$$h_{11:2}(\tau, \lambda) = h_{11}(\tau, \lambda) - \tilde{h}_{12}(\tau, \lambda)\Sigma_{22:1}^{-1}(\tau)\tilde{h}_{12}(\tau, \lambda)^*$$

where $\tilde{h}_{12}(\tau, \lambda)$
$= h_{1\cdot}(\tau, \lambda)\Gamma(\tau, e^{-i\lambda})\Gamma(\tau, 0)^*[-\Sigma_{21}(\tau)\Sigma_{11}^{-1}(\tau), I_{p_1}]^*$.

Locally stationary processes and asymptotics (7)

The first proposition of the following theorem is a multivariate extension of Dahlhaus (1996, p.149) and the second proposition is a time-dependent version of the one-way effect measure.

Theorem 2:

1. The one-step ahead time-wise prediction is given by

$$\det Cov\{w_{-1}(t,T)\} = (2\pi)^p \exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} \log\det h(t/T,\lambda)d\lambda\right\} + o_t(1) + o_T(1).$$

2. The average measure of one-way effect is given by

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} M_{v\to u}(t,T) = \int_0^1\int_{-\pi}^{\pi}\log\left[\frac{\det h_{11}(\tau,\lambda)}{\det h_{11:2}(\tau,\lambda)}\right]d\lambda d\tau.$$

Nonstationary reducible processes (1)

1. Let $\{w(t),\ t \in \mathbb{Z}\}$ be a second-order stationary p-vector process, and let $\{W(t),\ t \in \mathbb{Z}^{+}\}$ ($W(t) = 0$ for $t \in \mathbb{Z}^{0-}$) be another p-vector possibly non-stationary process with finite second-order moments.
2. Let $w_{-1}(t)$ be the one-step ahead prediction error of $w(t)$ based on $H\{w(t-1)\}$; then, the process $\{W(t)\}$ is said to be reducible to (the generating process) $\{w(t)\}$ if the projection residual of $W(t)$ onto $H\{W(t-1),\ w(0)\}$ is equal to $w_{-1}(t)$ for all $t \in \mathbb{Z}^{+}$.
3. If W is reducible to w, then $H\{W(t),\ w(0)\} = H\{w(t)\}$ for $t \in \mathbb{Z}^{+}$.
4. Namely, if a process $\{W(t)\}$ is reducible to $\{w(t)\}$, the respective prediction errors of $w(t)$ and $W(t)$ are identical if the information $H\{w(0)\}$ is supplemented, .

Nonstationary reducible processes (2)

Let $\{X(t)\}, \{Y(t)\}, \{Z(t)\}$ respectively be p_1, p_2, p_3-vector processes and set $W(t) = (X(t)^*, Y(t)^*, Z(t)^*)^*$. Set

$$A(L) = \sum_{j=0}^{a} A[j]L^j \quad \text{and} \quad B(L) = \sum_{k=0}^{b} B[k]L^k$$

where the $A[j]$ and the $B[j]$ are $(p_1 + p_2 + p_3) \times (p_1 + p_2 + p_3)$ matrices with $A[0] = B[0] = I_{p_1+p_2+p_3}$.

1. Assume that the zeros of $\det A(z)$ and $\det B(z)$ are either on or outside the unit circle, and $\det A(z)$ and $\det B(z)$ do not share common zeros.
2. Suppose $\{W(t)\}$ is generated by the VARMA process

$$A(L)W(t) = B(L)\varepsilon(t) \tag{1}$$

for a white noise process $\{\varepsilon(t)\}$ such that $E(\varepsilon(t)) = 0$ and $Cov(\varepsilon(t)) = \Sigma^\dagger$, where $rank(\Sigma^\dagger) = p_1 + p_2 + p_3$.

Nonstationary reducible processes (3)

Denote by $A(L)^\sharp$ the adjugate matrix of $A(L)$ such that

$$A(L)^\sharp A(L) = \begin{bmatrix} d(L) & & 0 \\ & \ddots & \\ 0 & & d(L) \end{bmatrix} \equiv D(L)$$

where $D(L)$ is the diagonal matrix with $\det A(L)$ as the common diagonal element. Applying the operator $A(L)^\sharp$ to the members of the equation (4), we have

$$\begin{aligned} \begin{bmatrix} d(L) & & 0 \\ & \ddots & \\ 0 & & d(L) \end{bmatrix} \begin{bmatrix} X(t) \\ Y(t) \\ Z(t) \end{bmatrix} &= A(L)^\sharp B(L)\varepsilon(t) \\ &\equiv \begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix} \equiv w(t). \end{aligned} \qquad (2)$$

Nonstationary reducible processes (4)

It follows from the construction that $\{w(t)\}$ is a stationary MA process; moreover, because the zeros of $\det\{A(z)^\sharp B(z)\}$ are either on or outside the unit circle, the covariance matrix of the one-step ahead prediction error of $w(t)$ is equal to $\Sigma^\dagger$ and the joint spectral density matrix $f(\lambda)$ of $\{w(t)\}$ has canonical factorization

$$f(\lambda) = \frac{1}{2\pi}\Lambda(e^{-i\lambda})\Lambda(e^{-i\lambda})^*, \tag{3}$$

where

- $\Lambda(e^{-i\lambda}) = A(e^{-i\lambda})^\sharp B(e^{-i\lambda}){\Sigma^\dagger}^{1/2}$
- ${\Sigma^\dagger}^{1/2}{\Sigma^\dagger}^{1/2} = \Sigma^\dagger$.

Nonstationary reducible processes (5)

Theorem 3:

1. The partial measures of interdependence of the possibly non-stationary processes $\{X(t)\}, \{Y(t)\}, \{Z(t)\}$ in the cointegrated process

$$A(L)W(t) = B(L)\varepsilon(t)$$

   are identified with the correspondng measures of the stationary processes $\{x(t)\}, \{y(t)\}$, and $\{z(t)\}$ which are given by

$$\begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix} \equiv w(t) \equiv A(L)^{\sharp} B(L)\varepsilon(t).$$

2. All the partial measures of the series$\{X(t)\}, \{Y(t)\}, \{Z(t)\}$ can be constructed through the corresponding measures on the basis of the canonical factor $\Lambda(e^{-i\lambda})$ in (3) of the stationary processes $\{x(t)\}, \{y(t)\}$, and $\{z(t)\}$.

Concluding remarks

- Construction of the measure of one-way effect relies crucially on the Szegö theory of approximization, whereas the theory is flexible enough to allow generalization in a variety of directions.
- The talk was limited to theoretical extension of the measures of interdependence originally defiend for stationary vetor time series to random fields, locally stationary processes and nonstationary reducible processes. Implementation by application to practical data is necesssary.
- Computational issues of the causal measures were intensively examined for stationary VARMA model based on simulated and real data sets; see Hosoya et al. (2017).
- Yao and Hosoya (2000) suggested a practical approach of statistical ineference on the measure of one-way effect in the cointegrated AR model set-up which is a typical reducible non-statioanry process.

Thank you for listening !