Bivariate long-range dependent time series models with general phase

Stefanos Kechagias
(joint work with Vladas Pipiras)

University of North Carolina at Chapel Hill
SAS Institute

July 6, 2015

European Meeting of Statisticians
Amsterdam 2015

Outline

1. Motivation from real data
2. Definitions and models for bivariate long-range dependent (LRD) time series
3. Inference under a parametric noncausal bivariate LRD model
4. Application to U.S. inflation rates
5. Future work

Bivariate LRD - Motivation from real data

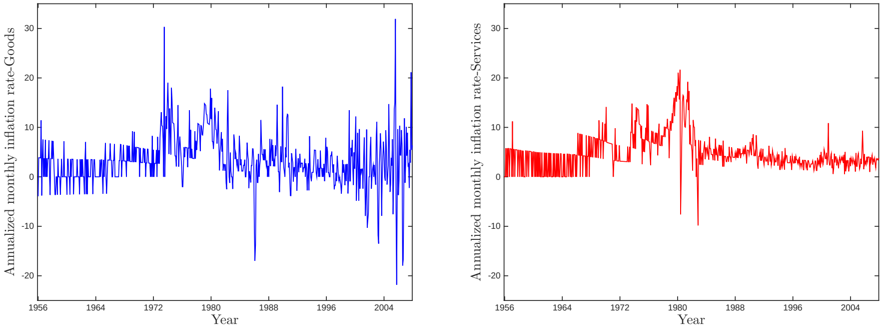

Figure: *Annualized monthly U.S. inflation rates for goods (left) and services (right) from February 1956 to January 2008.*

Bivariate LRD - Motivation from real data

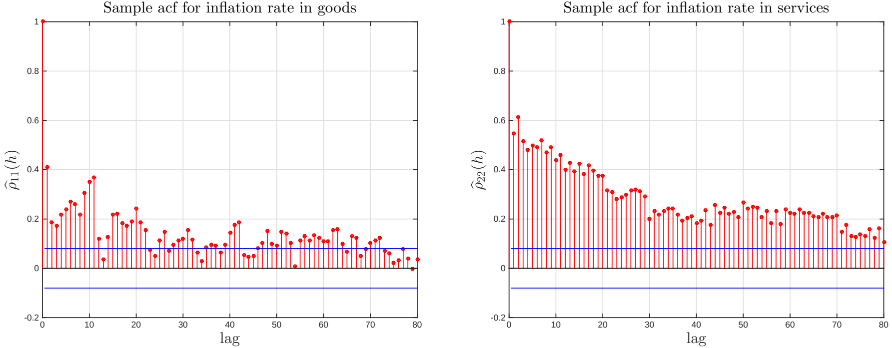

Figure: *Sample autocorrelation functions of U.S. inflation rates in goods (left) and in services (right).*

- Slow decay of the two acf's hints towards long-range dependence.
- Services inflation appears to have longer memory than goods inflation.

Bivariate LRD - Motivation from real data

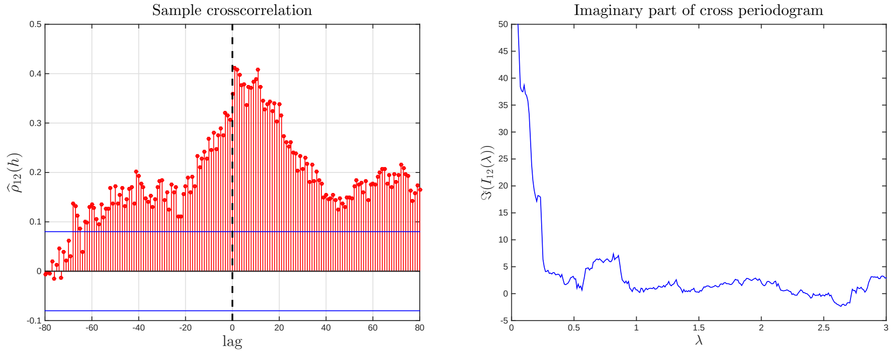

Figure: *Left: Sample crosscorrelation function of U.S. inflation rates in goods and services. Right: Imaginary part of cross periodogram.*

- $\Im(I_{12}(\lambda)) > 0$ (for λ close to 0) implies asymmetry in the series
- We need LRD models that allow for general asymmetry behavior

Definitions of bivariate long-range dependent series

We start with some notation:

- $\{X_n\}_{n\in\mathbb{Z}}=\{(X_{1,n},X_{2,n})'\}_{n\in\mathbb{Z}}$ is a bivariate, zero mean, second-order stationary time series
- $\gamma(n)=\mathbb{E}X_0X_n'$ is the autocovariance matrix of $\{X_n\}_{n\in\mathbb{Z}}$
- $f(\lambda)$ is the spectral density matrix of $\{X_n\}_{n\in\mathbb{Z}}$ satisfying
  $\gamma(n)=\int_{-\pi}^{\pi}e^{in\lambda}f(\lambda)d\lambda.$
- $\{\epsilon_n\}_{n\in\mathbb{Z}}$ is a bivariate WN with $\mathbb{E}\epsilon_n\epsilon_n'=I$.
  $\{\eta_n\}_{n\in\mathbb{Z}}$ is a bivariate WN with $\mathbb{E}\eta_n\eta_n'=\Sigma$

The definitions and models of the time series we will discuss involve the so-called *long-range dependent* parameters $d_1,d_2\in(0,1/2)$.

Definitions of bivariate LRD series

A bivariate stationary time series is LRD if

Time domain: As $k \to \infty$, its autocovariance matrix $\gamma(k)$ satisfies

$$\gamma(k) = \begin{pmatrix} \gamma_{11}(k) & \gamma_{12}(k) \\ \gamma_{21}(k) & \gamma_{22}(k) \end{pmatrix} \sim \begin{pmatrix} R_{11} k^{2d_1 - 1} & R_{12} k^{d_{12}-1} \\ R_{21} k^{d_{12}-1} & R_{22} k^{2d_2 - 1} \end{pmatrix},$$

where $R = (R_{jk})_{j,k=1,2}$ is some 2×2 real matrix and $d_{12} = d_1 + d_2$.

Spectral domain: As $\lambda \to 0^+$, its spectral density matrix $f(\lambda)$ satisfies

$$f(\lambda) = \begin{pmatrix} f_{11}(\lambda) & f_{12}(\lambda) \\ f_{21}(\lambda) & f_{22}(\lambda) \end{pmatrix} \sim \begin{pmatrix} g_{11}\lambda^{-2d_1} & g_{12} e^{i\phi} \lambda^{-d_{12}} \\ g_{12} e^{-i\phi} \lambda^{-d_{12}} & g_{22}\lambda^{-2d_2} \end{pmatrix},$$

where $g_{11}, g_{12}, g_{22} \in \mathbb{R}$ and the phase parameter $\phi \in (-\pi, \pi]$.

Note: The spectral domain definition contains 6 parameters.

Definitions of bivariate LRD series

Remark 1: The phase parameter is unique to LRD. Indeed, for *short-range dependent* series $f(\lambda) = (2\pi)^{-1} \sum_{k=-\infty}^{\infty} e^{-ik\lambda}\gamma(k)$ and $f(0) = (2\pi)^{-1} \sum_{k=-\infty}^{\infty} \gamma(k)$ has real entries.

Remark 2: Under mild assumptions (and letting $G_{12} = g_{12}e^{i\phi}$)

$$f_{12}(\lambda) \underset{\lambda \to 0^+}{\sim} G_{12}\lambda^{-2d_{12}} \quad \Leftrightarrow \quad \gamma_{12}(k) \underset{k \to \infty}{\sim} R_{12}k^{d_{12}-1},$$

with

$$\phi = -\text{atan}\left\{ \frac{R_{12} - R_{21}}{R_{12} + R_{21}} \tan\left(\frac{\pi d_{12}}{2}\right) \right\}.$$

Remark 3: $\phi = 0 \Leftrightarrow R_{12} = R_{21}$. This corresponds to $\gamma_{12}(k)$ being symmetric at the two tails, that is $\gamma_{12}(k) \underset{k \to \infty}{\sim} \gamma_{21}(k) = \gamma_{12}(-k)$.

Bivariate LRD models

- A common model for bivariate LRD series is the VARFIMA$(0, D, 0)$ defined as

$$X_n = \begin{pmatrix} X_{1,n} \\ X_{2,n} \end{pmatrix} = \begin{pmatrix} (I - B)^{-d_1}\eta_{1,n} \\ (I - B)^{-d_2}\eta_{2,n} \end{pmatrix} = (I-B)^{-D}\eta_n = (I-B)^{-D}Q_+\epsilon_n,$$

  where $D = \mathrm{diag}(d_1, d_2)$, $\eta_n \sim \mathrm{WN}(0, \Sigma)$, $\Sigma = Q_+ Q_+'$, $BX_n = X_{n-1}$

- **Fact:** The spectral density matrix of the VARFIMA series above satisfies

$$f(\lambda) \sim \begin{pmatrix} g_{11}\lambda^{-2d_1} & g_{12}e^{-i\phi}\lambda^{-d_{12}} \\ g_{12}e^{i\phi}\lambda^{-d_{12}} & g_{22}\lambda^{-2d_2} \end{pmatrix}, \quad \text{as} \quad \lambda \to 0^+,$$

  with the special phase parameter $\phi = \frac{\pi}{2}(d_1 - d_2)$.

Question: *Can one define a bivariate LRD model that allows for general phase parameter? Will such a model yield better predictions?*

Bivariate LRD models

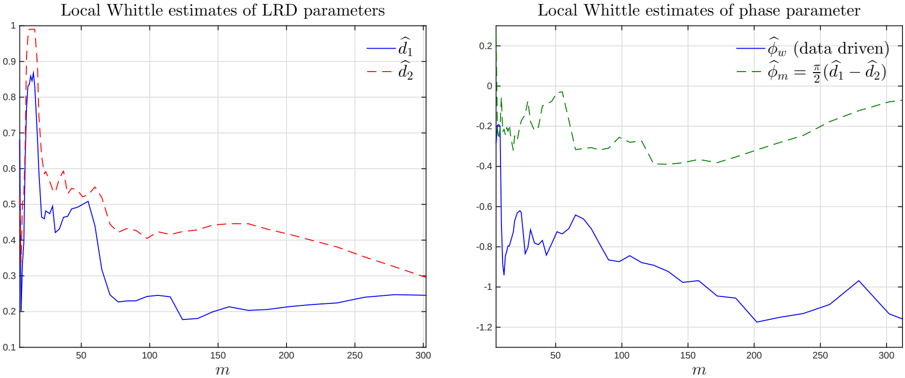

Figure: Left: Local Whittle estimates of d_1, d_2 for the inflation data plotted as functions of a tuning parameter $m = N^{0.25}, \ldots, N^{0.9}$, where N is the sample size. Right: Local Whittle phase estimates, one corresponding to the VARFIMA (dashed line) and one estimated directly from the data (solid line).

Causal VARFIMA$(0, D, 0)$

- The VARFIMA$(0, D, 0)$ series $X_n = (I - B)^{-D} Q_+ \epsilon_n$ has a *causal or one-sided* linear representation of the form

$$X_n = \sum_{m \in I} \Psi_m \epsilon_{n-m}, \qquad (1)$$

where $I = \mathbb{Z}^+$ and the entries $(\psi_{jk,m})_{j,k=1,2}$ of $\{\Psi_m\}_{j \in \mathbb{Z}}$ satisfy

$$\psi_{jk,m} \underset{m \to \infty}{\sim} \alpha_{jk}^+ |m|^{d_j - 1}, \quad \text{for some} \quad \alpha_{jk}^+ \in \mathbb{R}. \qquad (2)$$

- **Fact:** The causal series (1) with power-law coefficients (2) always has the special phase $\phi = \frac{\pi}{2}(d_1 - d_2)$.

Question: *How can we modify* (1) *to obtain a series with general phase?*

General Phase LRD models

- **Result 1:** A bivariate LRD series with general phase can be constructed by taking $I = \mathbb{Z}$ in (1) and Ψ_m as in (2) with $m \to \infty$ and $m \to -\infty$ (noncausal or two-sided series).

- **Result 2:** A causal bivariate LRD series with general phase can be constructed by taking *trigonometric power law* coefficients

$$\psi_{jk,m} = \alpha_{jk} m^{-b_k} \cos(2\pi m^a) + \beta_{jk} m^{-b_k} \sin(2\pi m^a), \quad m \geq 0,$$

  where $\alpha_{jk}, \beta_{jk} \in \mathbb{R}$, $0 < a < 1$, $\frac{1}{2} < b_k \leq 1 - \frac{1}{2}a$, $j = 1, 2$.

Question: *What about a parametric bivariate LRD model with general phase?*

Noncausal VARFIMA$(0, D, 0)$

- Define the bivariate noncausal VARFIMA$(0, D, 0)$ series as

$$X_n = \left((I - B)^{-D} Q_+ + (I - B^{-1})^{-D} Q_-\right) \epsilon_n,$$

where Q_+, Q_- are two real-valued 2×2 matrices.

- **Result 3:** The noncausal VARFIMA$(0, D, 0)$ series has a general phase. Moreover, its autocovariance function has an explicit form.

- The noncausal VARFIMA$(0, D, 0)$ series has 10 parameters. This causes identifiability problems as the same ϕ can be obtained by more than one choice of Q_+, Q_-.

Noncausal VARFIMA$(0, D, 0)$

- Taking

$$Q_- = CQ_+, \quad C = \begin{pmatrix} c & 0 \\ 0 & -c \end{pmatrix},$$

leads to the noncausal VARFIMA$(0, D, 0)$ series

$$X_n = \Delta_c(B)^{-1}\eta_n,$$
$$\Delta_c(B)^{-1} := (I - B)^{-D} + (I - B^{-1})^{-D} C.$$

- **Result 4:** For any $\phi_c \in (-\pi/2, \pi/2)$, $\exists!$ $c \in (-1, 1)$ such that X_n has the phase parameter $\phi = \phi_c$. Moreover, c has a closed form given by

$$c = \frac{2(a_1 + a_2) - \sqrt{\Delta}}{2(a_1 - a_2 - \tan(\phi_c))(1 + a_1 a_2)},$$

where $a_k = \tan\left(\frac{\pi d_k}{2}\right)$, and $\Delta = 16 a_1 a_2 + 4(1 + a_1 a_2)^2 \tan^2(\phi_c)$.

Models with SRD components

- Define the noncausal VARFIMA$(0, D, q)$ series as

$$Y_n = \Delta_c(B)^{-1}\Theta(B)\eta_n,$$

where $\Theta(B) = I_2 + \Theta_1 B + \ldots + \Theta_q B^q$ is the MA matrix polynomial.

- **Remark 4:** The noncausal VARFIMA$(0, D, q)$ series has a general phase parameter, and is identifiable.

- **Result 5:** The autocovariance matrix function of the noncausal VARFIMA$(0, D, q)$ series has an explicit form.

Models with SRD components

Define the noncausal VARFIMA(p, D, q) and FIVARMA(p, D, q) series as

$$\Phi(B)X_n = \Delta_c(B)^{-1}\Theta(B)\eta_n,$$
$$\Phi(B)\Delta_c(B)X_n = \Theta(B)\eta_n,$$

where $\Phi(B) = I_2 + \Phi_1 B + \ldots + \Phi_q B^q$ is the AR matrix polynomial.

Remark 5: The noncausal VARFIMA(p, D, q) has a general phase parameter, and is identifiable if the same VARMA(p, q) model is also identifiable.

Focus on models with diagonal Φ.

- Motivation from VARMA literature
- FIVARMA series can be written as VARFIMA series with diagonal Φ.
- If Φ is nondiagonal, X_n can be thought to exhibit a form of *fractional cointegration*.

CLDL algorithm for noncausal VARFIMA(p, D, q)

Let $\theta = (d_1, d_2, c, U, \Theta')'$.[1] Write the VARFIMA$(p, D, q)$ series as

$$\Phi(B)X_n = Y_n, \qquad Y_n = \Delta_c(B)\Theta(B)\eta_n.$$

The likelihood function of $\{Y_n\}_{n=p+1,\dots,N}$ conditional on $X_1, \dots, X_p$, Φ is

$$L(\Phi, \theta; X_n | X_1, \dots, X_p) \equiv L(\theta; \Phi(B)X_n), \quad n = p+1, \dots, N.$$

The conditional likelihood estimators of Φ and θ are then given by

$$(\widehat{\Phi}, \widehat{\theta}) = \underset{\Phi, \theta \in S}{\operatorname{argmax}} \, L(\Phi, \theta; X_n | X_1, \dots, X_p),$$

where $S = \{\theta : 0 < d_1, d_2 < 0.5, \ -1 < c < 1\}$ denotes the parameter space for θ. For fixed Φ, the likelihood is computed through the multivariate Durbin-Levinson algorithm.

[1] $\Sigma = U'U$, where U is upper triangular

Simulation-VARFIMA$(0, D, 0)$

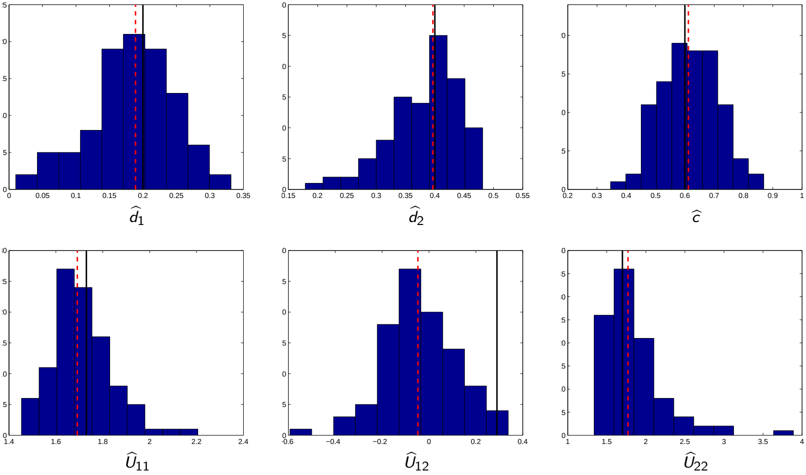

Figure: Sample size $N = 200$, 100 replications. Dotted lines indicate median over all replications while black lines indicate true parameter values.

Application to inflation data

- Causal VARFIMA$(1, D, 0)$, Sela 2010

$$\begin{aligned} g_t &= 0.30 g_{t-1} + 0.43 s_{t-1} + \eta_{1t}, \\ s_t &= -0.02 g_{t-1} - 0.31 s_{t-1} + \eta_{2t}, \end{aligned}$$

with $\widehat{d}_1 = 0.21, \widehat{d}_2 = 0.48$ and $\widehat{\Delta}_0(B)\eta_t \sim N(0, \widehat{\Sigma}_\eta)$.

- Noncausal VARFIMA$(1, D, 0)$

$$\begin{aligned} g_t &= 0.18 g_{t-1} + 0.03 s_{t-1} + e_{1t}, \\ s_t &= 0.09 g_{t-1} - 0.49 s_{t-1} + e_{2t}, \end{aligned}$$

with $\widehat{d}_1 = 0.18, \widehat{d}_2 = 0.36, \widehat{c} = 0.53$ and $\widehat{\Delta}_c(B)e_t \sim N(0, \widehat{\Sigma}_e)$. The corresponding phase estimate is $\widehat{\phi} = -1$.

References

Robinson, P. M. (2008), 'Multiple local Whittle estimation in stationary systems', *AoS* 36(5), 2508–2530.

Pai, J. & Ravishanker, N. (2009), 'A multivariate preconditioned conjugate gradient approach for MLE in vector long memory processes', *SPL* 79, 1282–1289.

Sela, R. J. & Hurvich, C. M. (2009), 'Computationally efficient methods for two multivariate fractionally integrated models', *JTSA* 30(6), 631–651.

Tsay, W. J. (2010), 'Maximum likelihood estimation of stationary multivariate ARFIMA processes', *JSCS* 80(7), 729745.

Kechagias, S. & Pipiras, V. (2015), 'Definitions and representations of multivariate long-range dependent time series', *JTSA* 36(1), 1–25.

Kechagias, S. & Pipiras, V., 'Inference and applications of a bivariate long-range dependent time series model with general phase', *Preprint*, http://stefanos.web.unc.edu/.

Future work

- Fractional cointegration models
- Trigonometric power law coefficients and other causal models
- Multivariate identifiable LRD models with general phase
- Invertibility of $\Delta_c(B)^{-1}$
- Assess the forecasting performance of general phase models