Polyspectral Factorization and Prediction for Quadratic Processes

T.S. McElroy, D. Ghosh, S. Lahiri

Contact: Tucker McElroy
Research and Methodology Directorate
U.S. Census Bureau,
Washington, DC, USA
email: tucker.s.mcelroy@census.gov

United States Census Bureau

 This work partially supported by NSF research grant 1811998, 2019-2021 (DMS Co-PI). *Development of a general framework for nonlinear prediction using autocumulants: Theory, methodology and computation.*

1 Overview of Results

1. Polyspectral factorization: bijection between cepstral coefficients and autocumulants
2. Wiener-Hopf factorization for infinite array spectral densities
3. Quadratic h-step ahead forecast filter
4. Conditions for quadratic filter to be linear

2 Main Objective

For a nonlinear stationary time series $\{X_t\}$ with autocovariance function $\{\gamma_k\}$, we provide explicit formulas for optimal filters $\Pi(z) = \sum_{k \geq 0} \pi_k z^k$ and $\Pi(z,y) = \sum_{j,k \geq 0} \pi_{j,k} z^j y^k$, where (for $h \geq 1$)

$$\widehat{X}_{t+h|t} = \sum_{k \geq 0} \pi_k X_{t-k} + \sum_{k \geq j \geq 0} \pi_{j,k} (X_{t-j} X_{t-k} - \gamma_{k-j}). \quad (1)$$

Upper triangular form: by setting $\Pi^{(r)}(x) = \sum_{j \geq 0} \pi_j^{(r)} x^j$ with $\pi_j^{(r)} = \pi_{j,j+r}$,

$$\Pi(z,y) = \sum_{j \geq 0} \sum_{r \geq 0} \pi_{j,j+r} z^j y^{j+r} = \sum_{r \geq 0} \Pi^{(r)}(zy) y^r.$$

3 Framework and Notation

Let $\{X_t\}$ be a $k+1$th order stationary time series with $k+1$ moments for given $k \geq 1$; the $k+1$th order autocumulant function is defined by

$$\gamma_{h_1,\ldots,h_k} = \text{cum}\left(X_{t+h_1}, \ldots, X_{t+h_k}, X_t\right).$$

Under absolute summability, define the $k+1$th order polyspectral density by

$$f(\lambda_1, \ldots, \lambda_k) = \sum_{\underline{h} \in \mathbb{Z}^k} \gamma_{\underline{h}} \exp\{-i \underline{h}' \underline{\lambda}\},$$

where we set $\underline{\lambda} = [\lambda_1, \ldots, \lambda_k]'$, and each of these are frequencies in $[-\pi, \pi]$. The coefficients are recovered via integration over the unit torus:

$$\gamma_{\underline{h}} = \frac{1}{(2\pi)^k} \int_{[-\pi,\pi]^k} \exp\{i \underline{h}' \underline{\lambda}\} f(\underline{\lambda}) \, d\underline{\lambda}.$$

Shorthand: $\langle z^h g(z) \rangle_z = (2\pi)^{-1} \int_{-\pi}^{\pi} e^{-i\lambda h} g(e^{-i\lambda}) \, d\lambda$, and for any $-\infty \leq r \leq s \leq \infty$ and Laurent series $f(z)$, $[f(z)]_r^s = \sum_{j=r}^s \langle z^{-j} f(z) \rangle_z z^j$.

4 Linear Prediction of Quadratic Variables

In the case of linear h-step ahead forecasting, the optimal filter is given by $\eta(B)$ (where B is the backshift operator):

$$\eta(z) = [z^{-h} \Psi_2(z)]_0^\infty \, \Psi_2(z)^{-1} = \sum_{j \geq 0} \psi_{j+h}^{(2)} z^j \, \Psi_2(z)^{-1}, \quad (2)$$

where $\sigma^2 \Psi_2(z) \Psi_2(z^{-1}) = f(z)$, the spectral density.

For any $r \geq 0$ define the process $\{Y_t^{(r)}\}$ by $Y_t^{(r)} = X_t X_{t-r} - \gamma_r$, which is stationary with mean zero. With $\widehat{Y_{t-j}^{(r)}} = \Phi^{(j,r)}(B) X_t$, we have (**new**):

$$\Phi^{(j,r)}(z) = \sigma^{-2} [z^j \Psi_2(z^{-1})^{-1} \langle y^r f(zy^{-1}, y) \rangle_y]_0^\infty \Psi_2(z)^{-1}. \quad (3)$$

With (3) we can rewrite (1) as follows, with both summands orthogonal:

$$\widehat{X}_{t+h|t} = \eta(B) X_t \oplus \sum_{r \geq 0} \Pi^{(r)}(B) (Y_t^{(r)} - \widehat{Y_t^{(r)}}). \quad (4)$$

To derive $\Pi^{(r)}(z)$ we need a factorization result for polyspectra.

5 Factorization of Polyspectra

For a linear process $X_t = \Psi(B) Z_t$ with $\{Z_t\}$ i.i.d. and $\Psi(z) = \sum_{j \in \mathbb{Z}} \psi_j z^j$, it is known that (where μ_{k+1} is the $k+1$th cumulant of Z_t)

$$f(\underline{z}) = \mu_{k+1} \prod_{j=1}^k \Psi(z_j) \, \Psi(z_1^{-1} \cdots z_k^{-1}). \quad (5)$$

We want to generalize this. Let $\mathcal{S}_{k+1}$ be the group of permutations on $k+1$ elements, and ρ is a group representation for $\sigma \in \mathcal{S}_{k+1}$ to $k \times k$ matrices $\rho(\sigma)$; the autocumulant symmetry is then denoted by $\gamma_{\underline{h}} = \gamma_{\rho(\sigma)\underline{h}}$. Also, (**new**) for any $\sigma \in \mathcal{S}_{k+1}$ the value of the polyspectra at variables z_m for $1 \leq m \leq k$ is the same if z_m is replaced by

$$\prod_{j=1}^k z_j^{[\rho(\sigma^{-1})]_{jm}}.$$

The general polyspectral factorization (**new**) is

$$f(\underline{z}) \propto \prod_{\sigma \in \mathcal{S}_{k+1}} \Psi_{k+1} \left(\left\{ \prod_{j=1}^k z_j^{[\rho(\sigma) D]_{jm}} \right\}_{m=1}^k \right), \quad (6)$$

where D is an upper triangular matrix of ones, and $\Psi_{k+1}(\underline{u}) = \exp\{\sum_{\underline{h} > 0} \tau_{D\underline{h}} \, \underline{u}^{\underline{h}}\}$, with the shorthand $\underline{z}^{\underline{j}} = \prod_{\ell=1}^k z_\ell^{j_\ell}$; the coefficients τ are real numbers supported on the tetrahedral cone

$$R = \{h_1, h_2, \ldots, h_k \in \mathbb{Z} : h_1 \geq h_2 \geq \ldots \geq h_k \geq 0\}.$$

Figure 1: Absolute value of the GARCH(1,1) bispectrum.

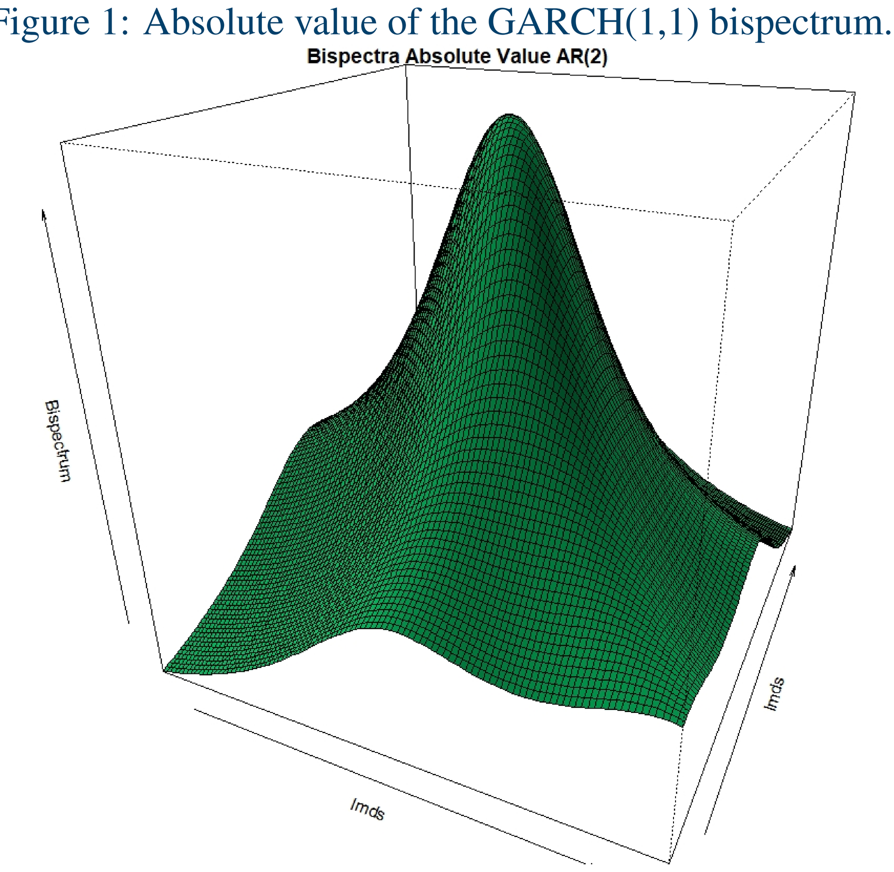
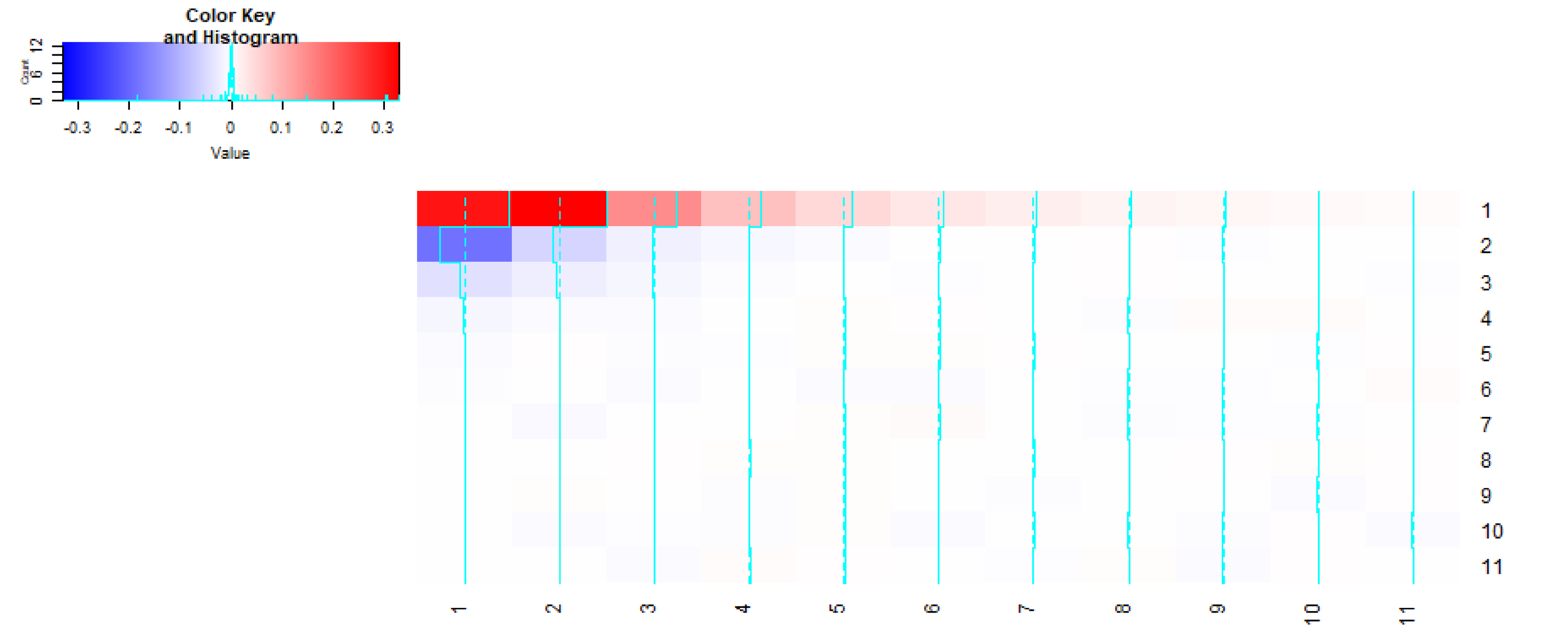


Figure 2: τ matrix for the GARCH(1,1) process.



We say a process is order $k+1$ quasi-linear if (5) holds for a Laurent series $\Psi(z)$; defining the cone's edges via $R_j = \{D \, \underline{h} : h_i = 0 \text{ for all } i \neq j\}$, a process is order $k+1$ quasi-linear (**new**) iff τ is supported on $R_1 \cup R_k$. An example is the GARCH(1,1) process, which is not third order quasi-linear (Figure 1); τ is not supported on the edges (Figure 2).

6 Wiener-Hopf Factorization of Infinite Arrays

Define a 1-array Laurent series $\underline{\xi}(z)$ with rth component $\xi^{(r)}(z)$ for $r \geq 0$; define a 2-array Laurent series $B(z)$ with component r, s given by $B^{(r,s)}(z)$ for $r, s \geq 0$. A power series array is special case where Laurent series is causal in z. Linear algebra carries over: multiplication, transpose, identity array, invertibility. By definition, a Laurent series 2-array $B(z)$ has a forward Wiener-Hopf factorization if there exist power series 2-arrays $B^-(z)$ and $B^+(z)$ such that

$$B(z) = B^-(z^{-1}) B^+(z)'. \quad (7)$$

$B(z)$ has a backward Wiener-Hopf factorization if

$$B(z) = B^+(z) B^-(z^{-1})'. \quad (8)$$

With $\underline{\tau}(z)$ a Laurent series 1-array, $\underline{\xi}(z)$ a power series 1-array, and $B(z)$ a Laurent series 2-array, define the system

$$[\underline{\tau}(z)]_0^\infty = [B(z) \, \underline{\xi}(z)]_0^\infty.$$

If $B(z)$ has a forward Wiener-Hopf factorization (7) such that $B^+(z)$ and $B^-(z)$ are invertible, then (**new**) the system is solved by

$$\underline{\xi}(z) = B^+(z)^{-1\prime} [B^-(z^{-1})^{-1} \underline{\tau}(z)]_0^\infty.$$

A Hermitian Laurent series 2-array $B(z)$ is positive definite if every Schur complement is a positive definite (scalar) function. Suppose that $B(z)$ is a Hermitian Laurent series 2-array that is positive definite. Then (**new**) there exist power series 2-arrays $B^-(z)$ and $B^+(z)$ such that (8) holds.

7 Quadratic Forecast Formula

With (4) our objective is to compute the Laurent series 1-array $\underline{\Pi}(z)$ with rth component $\Pi^{(r)}(z)$ for $r \geq 0$. Let $A(z)$ be spectrum of the 1-array process $\{\underline{Y}_t\}$, i.e., for any $r, s \geq 0$, $\langle z^{-k} A^{(r,s)}(z) \rangle_z = \text{Cov}[Y_{t+k}^{(r)}, Y_t^{(s)}]$. **Assumption P**: The autocovariance generating function for the linear prediction of $Y_t^{(r)}$ from $\{Y_t^{(r-1)}, \ldots, Y_t^{(0)}\}$ is positive definite for all $r \geq 1$. This implies that $A(z)$ is positive definite. Define $\underline{L}(z)$ as the "forward-looking" portion of the bispectrum, via $L^{(s)}(z) = \langle y^{-s} f(zy, y^{-1}) \rangle_y$ for $s \geq 0$. **Assumption L**: $\underline{L}(z)$ is non-zero. If Assumption L is violated, the quadratic filter reduces trivially to a linear filter, because $\Phi^{(j,r)}(z) = 0$. If $\{X_t\}$ satisfies Assumptions P and L, then (**new**)

$$\underline{\Pi}(z) = [0, I] \, B^+(z)^{-\prime} \left[B^-(z^{-1})^{-1} \, \underline{R}(z) \right]_0^\infty, \quad (9)$$

where the $[0, I]$ operator denotes a forward row shift acting on a Laurent series 2-array, and

$$B(z) = \begin{bmatrix} 1 & -\Psi_2(z)^{-1} \underline{L}(z^{-1})' \\ -\Psi_2(z^{-1})^{-1} \underline{L}(z) & \sigma^2 A(z) \end{bmatrix} \qquad \underline{R}(z) = \begin{bmatrix} 0 \\ \sigma^2 z^{-h} [\Psi_2(z)]_0^{h-1} \, \Psi_2(z)^{-1} \underline{L}(z^{-1}) \end{bmatrix}.$$

The MSE of the quadratic filter is equal to the linear MSE minus the quantity $\langle \underline{Q}'(z^{-1}) \underline{Q}(z) \rangle$, where

$$\underline{Q}(z) = \sigma^{-1} \left[B^+(z^{-1})^{-1} \, \underline{R}(z) \right]_0^\infty.$$

When $\underline{R}(z) = 0$ the quadratic filter reduces to a linear filter. This is equivalent to

$$0 = [(z^{-h} - \eta(z)) \underline{L}(z^{-1})]_0^\infty. \quad (10)$$