Optimal linear prediction of stochastic trends (and cycles)
PRELIMINARY AND INCOMPLETE VERSION. PLEASE DO NOT CIRCULATE

Tommaso Proietti
Università di Roma "Tor Vergata" and CREATES

Alessandro Giovannelli
Università di Roma "Tor Vergata"

February 15, 2016

Abstract

A recent strand of the time series literature has considered the problem of estimating high-dimensional autocovariance matrices, for the purpose of out of sample prediction. For an integrated time series, the Beveridge-–Nelson trend is defined as the current value of the series plus the sum of all forecastable future changes. For the optimal linear projection of all future changes into the space spanned by the past of the series, we need to solve a high-dimensional Toeplitz system involving n autocovariances, where n is the sample size. The paper proposes a non parametric estimator of the trend that relies on banding, or tapering, the autocovariance sequence to achieve consistency. We derive the properties of the estimator and compare it with alternative parametric estimators based on the direct and indirect finite order autoregressive predictors. We then consider the estimation of trends within a multivariate system composed of a target series (e.g. gross domestic product) and a set of observable dynamic factors.

Keywords: High-dimensional autocovariance matrices; Toeplitz systems; Beveridge–Nelson decomposition; Factor models.

JEL Codes: C22, C52, C58.

1 Introduction

The paper deals with the estimation of the stochastic trend that drives the dynamics of a stochastic process that is integrated of order one, or difference stationary. The trend definition arises from the well-known Beveridge-Nelson decomposition (Beveridge and Nelson (1981)), which is based on the projection of the sum of the future changes of the series onto the available sample. This projection requires the estimation of the complete autocovariance function (ACV), thereby raising an incidental or high-dimensional problem. In the past, this has been solved by fitting a model of the ARIMA class to a finite realization, but here we take a nonparametric approach based on regularizing the sample ACV function in a novel way.

There is a vast literature concerning the estimation of high-dimensional covariance matrices in the multivariate cross-sectional setting, where the number of random variables is large, compared to the number of observations available. This literature is exposed in detail in (Pourahmadi, 2013).

We contribute to the literature on high-dimensional covariance matrix estimation by proposing a new method based on a tapered Durbin-Levinson recursion. The proposed tapered ACV estimator regularizes the partial autocorrelation function (PACF), so that the coefficients of the linear predictor are shrunk towards the values implied by a finite order autoregressive predictor. In particular, our estimator is based on a regularized Durbin-Levinson set of recursions. For processes with high dynamic range, such as a cyclical process, regularising the PACF may be more effective than tapering the ACV.

2 The Beveridge-Nelson decomposition

Assume that $\{X_t, t = 0, 1, \ldots\}$ is a difference stationary process, i.e. $\Delta X_t = X_t - X_{t-1}, t = 1, 2, \ldots$, is a stationary process that we assume to be zero mean, $\mathrm{E}(\Delta X_t) = 0$, $\forall t$, and characterised by the autocovariance function $\gamma(k) = \mathrm{E}(\Delta X_t \cdot \Delta X_{t-k}), k = 0, \pm 1, \pm 2, \ldots$, which is absolutely summable, $\sum_{k=-\infty}^{\infty} |\gamma(k)| < \infty$.

The Beveridge-Nelson decomposition (Beveridge and Nelson (1981)) is $X_t = m_t + c_t$, where m_t is the trend and c_t is the cycle- The BN trend is defined in terms of the long run prediction of the series at time t: m_t equals the current level plus all forecastable future changes:

$$\begin{aligned} \lim_{h\to\infty} \mathrm{E}[X_{t+h}|\mathcal{F}_t] &= \lim_{h\to\infty} \mathrm{E}[X_t + \textstyle\sum_{j=1}^{h} \Delta X_{t+j}|\mathcal{F}_t] \\ &= X_t + \textstyle\sum_{j=1}^{\infty} \mathrm{E}[\Delta X_{t+j}|\mathcal{F}_t]. \end{aligned}$$

Here $\mathcal{F}_t$ is the information set available at time t.

In fact, writing

$$X_{t+h} = X_t + \sum_{j=1}^{h} \Delta X_{t+j},$$

the h-step ahead predictor, $\tilde{X}_{t+h|t} = \mathrm{E}[X_{t+h}|\mathcal{F}_t]$, is obtained by adding to the current X_t all forecastable future changes up to time $t + h$, i.e.:

$$\tilde{X}_{t+h|t} = X_t + \sum_{j=1}^{h} \widetilde{\Delta X}_{t+j|t}, \tag{1}$$

where $\widetilde{\Delta X}_{t+j|t} = \mathrm{E}(\Delta X_{t+j}|\mathcal{F}_t)$. If h is allowed to go to infinity in (1) and we assume that the drift is zero, then $\tilde{X}_{t+h|t}$ tends to the BN trend, or permanent, component, and $\lim_{h\to\infty}\sum_{j=1}^{h}\widetilde{\Delta X}_{t+j|t}$ is minus the BN cycle (transitory component). Hence. we define

$$\begin{aligned} X_t &= m_t + c_t, \\ m_t &= \lim_{h\to\infty}\tilde{X}_{t+h|t} \\ &= X_t + \textstyle\sum_{j=1}^{\infty}\widetilde{\Delta X}_{t+j|t}, \\ c_t &= -\textstyle\sum_{j=1}^{\infty}\widetilde{\Delta X}_{t+j|t}. \end{aligned} \tag{2}$$

In the case when the drift is nonzero, $\mathrm{E}(\Delta X_t) = \beta \neq 0$, the BN trend is redefined as the limit of $\tilde{X}_{t+h|t} - \beta h$, as $h \to \infty$, which equals the current value of the series plus "all forecastable future changes beyond the mean rate of drift" (Beveridge and Nelson, 1981):

$$m_t = X_t + \sum_{j=1}^{\infty}(\widetilde{\Delta X}_{t+j|t} - \beta).$$

The optimal linear predictor of ΔX_{t+j} at time t is

$$\widetilde{\Delta X}_{t+j|t} = \sum_{i=1}^{t}\phi_{it}^{(j)}\Delta X_{t-i-1}.$$

Defining the $t \times 1$ vector $\phi_t^{(h)} = [\phi_{1t}^{(j)}, \phi_{2t}^{(j)}, \dots, \phi_{tt}^{(j)}]'$,

$$\Gamma_t = \begin{bmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(t-1) \\ \gamma(1) & \gamma(0) & \ddots & \gamma(t-2) \\ \vdots & \cdots & \ddots & \vdots \\ \gamma(t-1) & \gamma(t-2) & \cdots & \gamma(0) \end{bmatrix}, \quad \gamma_t^{(j)} = \begin{bmatrix} \gamma(j) \\ \gamma(j+1) \\ \vdots \\ \gamma(j+t-1) \end{bmatrix}, \ j = 1, 2, \dots,$$

the coefficients of the j-step-ahead predictor of ΔX_{t+j} are obtained from the solution of the Yule-Walker linear system:

$$\boldsymbol{\phi}_t^{(h)} = \Gamma_t^{-1}\gamma_t^{(h)}.$$

Then, let us set

$$\gamma_{t,h} = \begin{bmatrix} \gamma(1) + \cdots + \gamma(h) \\ \gamma(2) + \cdots + \gamma(h+1) \\ \vdots \\ \gamma(t) + \cdots + \gamma(h+t-1) \end{bmatrix},$$

so that

$$\gamma_{t,h} = \gamma_{t,h-1} + \gamma_t^{(h)}, h > 1, \ \ \gamma_{t,1} = \gamma_t^{(1)},$$

and $\gamma_t = \lim_{h\to\infty}\gamma_{t,h}$. Then, the optimal linear predictor of the sum of all future changes is

$$\sum_{j=1}^{\infty}\widetilde{\Delta X}_{t+j|t} = \sum_{i=1}^{t}\phi_{it}\Delta X_{t-i+1},$$

where the coefficient vector $\phi_t = [\phi_{1t}, \phi_{2t}, \dots, \phi_{tt}]'$ satisfies the Yule-Walker system

$$\Gamma_t\phi_t = \gamma_t. \tag{3}$$

The components of the BN decomposition are thus computed as follows:

$$\begin{aligned} m_t &= (1+\phi_{1t})X_t + (\phi_{2t} - \phi_{1t})X_{t-1} + \cdots + (\phi_{tt} - \phi_{t-1,t})X_2 - \phi_{tt}X_1, \\ c_t &= -\phi_{1t}\Delta X_t - \phi_{2t}\Delta X_{t-1} - \cdots - \phi_{tt}\Delta X_1. \end{aligned} \tag{4}$$

3 The BN decomposition of a linear process

Let us consider the linear process:

$$\Delta X_t = \psi(B)\xi_t, \quad \xi_t \sim \mathrm{WN}(0,\sigma^2), \qquad t = 1,...,n, \tag{5}$$

where $\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j, \psi_0 = 1, \sum_j |\psi_j| < \infty$, and B is the backward shift operator, B, $\Delta = 1 - B$. We further assume that the coefficients of $[\psi(B)]^{-1}$ are summable.

The h-step ahead prediction of X_t at time t is related to the $h+1$ prediction at time $t-1$ by the following (forecast revision) equation:

$$\tilde{X}_{t+h|t} = \tilde{X}_{t+h|t-1} + (1 + \psi_1 + \cdots + \psi_{h-1})\xi_t.$$

Letting $m_t = \lim_{h\to\infty} \tilde{X}_{t+h|t}$, and taking the limit of both sides as $h \to \infty$ we have the trend generating equation

$$m_t = m_{t-1} + \psi(1)\xi_t,$$

a random walk.

Subtracting side by side from $X_t = X_{t-1} + \psi(B)\xi_t$, yields

$$c_t = \psi^*(B)\xi_t, \quad \Delta\psi^*(B) = \psi(B) - \psi(1),$$

where $\psi^*(B) = \sum_{j=0}^{\infty} \psi_j^* B^j$ has $\psi_0^* = 1 - \psi(1)$.

As a result $\mathrm{Cov}(\Delta m_t, c_t) = \psi(1)[1 - \psi(1)]\sigma^2$ is positive if $\psi(1) < 1$ (low persistence) - notice that our assumptions imply $\psi(1) > 0$; it is negative if $\psi(1) > 1$ (high persistence). It can also be zero, of course, when $\psi(1) = 1$, but still the trend and the cycle still cross-correlated, that is $\mathrm{Cov}(\Delta m_t, c_{t-j})$ is non zero for $j < 0$.

The components are measurable with respect to the information set available at time t, via the following filters

$$m_t = \frac{\psi(1)}{\psi(B)} X_t, \quad c_t = \frac{\psi^*(B)}{\psi(B)} \Delta X_t.$$

As an illustration, consider the BN decomposition of an ARIMA(2,1,2) process, $\phi(B)\Delta X_t = \beta + \theta(B)\xi_t, \xi_t \sim \mathrm{WN}(0,\sigma^2)$ fitted by (Morley, Nelson, and Zivot, 2003) to the US GDP, where $\theta(L) = 1 + \theta_1 L + \theta_2 B^2$ and $\phi(B) = 1 - \phi_1 B - \phi_2 B^2$, such that: $m_t = m_{t-1} + \beta + \frac{\theta(1)}{\phi(1)}\xi_t$ and c_t has the ARMA(2,1) representation:

$$c_t = \phi_1 c_{t-1} + \phi_2 c_{t-2} + \left[1 - \frac{\theta(1)}{\phi(1)}\right]\xi_t + \left[\phi_2 \frac{\theta(1)}{\phi(1)} + \theta_2\right]\xi_{t-1}. \tag{6}$$

Interestingly, $\phi_1 + \phi_2 = -(\theta_1 + \theta_2)$ implies $\Delta m_t = \beta + \xi_t$ and $c_t = \phi_1 c_{t-1} + \phi_2 c_{t-2} + (\phi_2 + \theta_2)\xi_{t-1}$. When $\phi_1 + \phi_2 \neq -(\theta_1 + \theta_2)$, the ARMA(2,1) representation for the cycle is noninvertible if

$$\left|\frac{\phi_2\theta(1) + \theta_2\phi(1)}{\phi(1) - \theta(1)}\right| > 1;$$

if we set $\theta_1 = \theta_2 = 0$, i.e. for an ARIMA(2,1,0) process, the cycle is non-invertible if $|\phi_2/(\phi_1 + \phi_2)| > 1$.

The BN decomposition also provides the optimal real time estimator of the components of an unobserved components model, where the trend is a random walk and the cycle is a stationary ARMA process, driven by correlated disturbances. For this interpretation see (Proietti and Harvey, 2000), (Morley et al., 2003), (Proietti, 2004), (Proietti, 2006), (Morley, 2011), and more recently, (Weber, 2011) and (Dungey, Jacobs, Tian, and van Norden, 2015).

3.1 Estimation of the BN components

This paper deals with nonparametric estimation of the components in (4). Given a sample time series $\{x_t, t = 0, \ldots, n\}$, with associated changes $\{y_t = \Delta x_t, t = 1, \ldots, n\}$ (if ΔX_t has a non zero mean, then we define $y_t = \Delta x_t - n^{-1}\sum_t \Delta x_t$), we can estimate the lag j autocovariance using the estimator

$$\hat{\gamma}(j) = \frac{1}{n}\sum_{t=j+1}^{n} y_t y_{t-j}, j = 0, 1, \ldots, n-1.$$

If the aim is estimating the BN components at time n, the direct nonparametric estimator based on replacing the theoretical autocovariances with the sample ones is cleary unfeasible: the maximum lag sample autocovariance is available at lag $n-1$. Furthermore, $\hat{\Gamma}_t = \{\hat{\gamma}(|r-s|), r, s = 1, 2, \ldots, t\}$ is not a consistent estimate of Γ_t, as it is reviewed in the next section. Increasing the number of observations does not aid the estimation as higher order autocovariances come into play.

Hence, the estimation of the BN components poses an incidental or high-dimensional inferential problem. A parametric model can be used that expresses all the covariance structure, $\{\gamma(j), j = 0, \pm 1, \pm 2, \cdots\}$, or the spectral density of ΔX_t, as a function of a parsimonious set of parameters. Indeed, the ARIMA model class has been considered, see (Morley et al., 2003) and the references therein. However, one may argue that a parsimonious model fitted so as to maximise the overall score of the one-step-ahead predictive likelihoods is inadequate for estimating long-run trends. As (Newbold, Agiakloglou, and Miller, 1993) conclude,

> *"Parsimoniously parametrized time series models were developed as aids to short-term forecasting, where the fiction that the analyst has discovered the 'true' model is innocuous. Such fiction, however, is far from innocuous when attempting to base inference about long-run behavior on these fitted models."*

4 The Estimation of High-dimensional Autocovariance Matrices

Let $\{y_t, t = 1, 2, \ldots\}$ denote a time series realization of a stationary short memory stochastic process Y_t. The sample mean and autocovariance are respectively

$$\bar{y} = \frac{1}{n}\sum_{t=1}^{n} y_t \quad \text{and} \quad \hat{\gamma}(j) = \frac{1}{n}\sum_{t=j+1}^{n}(y_t - \bar{y})(y_{t-j} - \bar{y})$$

Let $\hat{\Gamma}_n = \{\hat{\gamma}(|i-j|), i, j = 1, \ldots, n\}$, which is a positive definite Toeplitz matrix. (Wu and Pourahmadi, 2009) showed that $\hat{\Gamma}_n$ is not consistent for Γ_n, as the operator norm (the largest eigenvalue) of the estimation error matrix, $\rho(\hat{\Gamma}_{n,l} - \Gamma_n)$, does not converge to zero. They proposed the banded estimator

$$\hat{\Gamma}_{n,l} = \{\hat{\gamma}(|i-j|) \cdot I(|i-j| \le l), i, j = 1, \ldots, n\}, \tag{7}$$

where l is the band parameter and $I(\cdot)$ is the indicator function.

(Hannan and Deistler, 1988) showed that, for linear ARMA processes and for $l \le (\log n)^\alpha, \alpha < \infty$, the infinity norm of $\hat{\Gamma}_{n,l} - \Gamma_n$ is $O\left(n^{-1/2}\sqrt{\log\log n}\right)$. (Wu and Pourahmadi, 2009) proved the consistency of the banded estimator for non-linear short-range dependent processes of the form $Y_t = g(\dots, \xi_{t-1}, \xi_t)$, such that $g(\cdot)$ is a measurable function and $\mathrm{E}(Y_t^2) < \infty$, and obtained an explicit upper bound for the operator norm of $\hat{\Gamma}_{n,l} - \Gamma_n$,

$$\rho(\hat{\Gamma}_{n,l} - \Gamma_n) = O\left(r_n\right), r_n = ln^{2/\alpha - 1} + \sum_{j=l}^{\infty} |\gamma(j)|$$

where $\alpha \in (2, 4]$ is the moment of the physical dependence measure introduced by (Wu, 2005). (Bickel and Gel, 2011) obtained the consistency of (7) under the Frobenius norm,

$$\parallel \hat{\Gamma}_{n,l} - \Gamma_n \parallel_F = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} [\hat{\gamma}(|i-j|) - \gamma(|i-j|)]^2 \cdot I(|i-j| \leq l)}$$

provided that l increases with n sufficiently slowly.

Banding has the advantage of preserving the Toeplitz structure, but the estimator (7) is not necessarily positive definite. WP propose estimating the band l by minimizing the average estimation error infinity norm computed over subsamples of the series, where the true value is replaced by the sample autocovariance matrix banded at a large value, such as $K = 30$. The estimated band is decreased by one if the corresponding autocovariance matrix fails to be p.d., until

The selection of the optimal banding parameter is considered in (Bickel and Gel, 2011). They propose a cross-validation method which divides the time series into two consecutive segments of length n_1 (typically $n/3$) and n_2, respectively. The banded autocovariance matrices estimates in (7) are computed as well as target quantities, such as the h-step-ahead mean square forecast error of the optimal linear (Yule-Walker) predictor, or an autocovariance estimate, (McMurry and Politis, 2010) proposed the tapered autocovariance matrix estimator

$$\hat{\Gamma}_{n,l,MP} = \{\hat{\gamma}(|i-j|) \cdot w(|i-j|), i, j = 1, \dots, n\} \tag{8}$$

where $w(|i-j|) = \kappa((i-j)/l)$, where l is the banding parameter and $\kappa(u)$ is a kernel function of the class of tapered weight functions in (McMurry and Politis, 2010), given by

$$\kappa(u) = \begin{cases} 1, & |u| \leq 1 \\ g(|u|), & 1 < |u| \leq c_\kappa \\ 0, & |u| > c_\kappa \end{cases}$$

- The trapezoidal kernel, denoted $\kappa_T(u)$, arises for $c_\kappa = 2$ and $g(|u|) = 2 - |u|$.
- The infinitely differentiable kernel, denoted $\kappa_I(u)$, arises for $c_\kappa = 2$ and

$$g(|u|) = \exp\left\{ -\frac{b}{(|u|-2)^2} \exp\left[-\frac{b}{(|u|-1)^2} \right] \right\}$$

MP propose a data-based selection criterion for the banding parameter l, which is chosen as the smallest $\hat{l}$ such that

$$|\hat{\rho}(\hat{l} + k)| < c\sqrt{\log n / n},\ k = 1, 2, \dots, K_n,\ K_n = o(\log n). \tag{9}$$

For the sample sizes typically available in economics, MP recommend $c = 2$ and $K_n = 5$. The rule amounts to conducting an approximate 95% simultaneous test of $\rho(\hat{l}+k) = 0, k = 1, \dots, K_n$.

The estimator in (8) is only asymptotically p.d. and in finite samples MP suggested correcting the eigenvalues of $\hat{\Gamma}_{n,l,MP} = \sum_{i=1}^{n} d_i v_i v_i'$, where d_i is the i-th eigenvalue are v_i is the corresponding eigenvector, by replacing the negative eigenvalues $\hat{\Gamma}^*_{n,l,MP} = \sum_{i=1}^{n} d_i^* v_i v_i'$, with $d_i^* = \max\{d_i, \epsilon\hat{\gamma}(0)/n^{\beta}\}$ with $\epsilon = \beta = 1$. In they argue that the rescaled estimator, using a larger ϵ, e.g. $\epsilon = 20$ and $\beta = 1$,

$$\hat{\Gamma}^{\dagger}_{n,l,MP} = \frac{n\hat{\gamma}(0)}{\sum_i d_i^*}\hat{\Gamma}^*_{n,l,MP},$$

ensuring that the average eigenvalue of $\hat{\Gamma}^{\dagger}_{n,l,MP}$ equals $\hat{\gamma}(0)$, is preferable for the purposes of prediction. Other correction methods are presented in (McMurry, Politis, et al., 2015).

5 A Regularized Durbin-Levinson Algorithm

Rather than banding or tapering the sample autocovariance sequence, we propose tapering the sample partial autocorrelations so that the coefficients of the linear predictor are shrunk towards the values implied by the finite order AR predictor. This amounts to shrinking the high order autocovariances towards $\gamma(p + k) = \sum_{j=1}^{p} \phi_{jp}^{(1)}\gamma(p + k - j)$, using the above notation, for some $p = o(n)$. For processes with high dynamic range, such as a cyclical process, regularising the PACF may be more effective than tapering the ACV.

Consider the following regularised Durbin-Levinson recursions (see Brockwell and Davis (1991)): let $v_0 = w_0\gamma(0)$, $\pi_{11} = w_1\gamma(1)/\gamma(0)$, $v_1 = (1 - \pi_{11}^2)v_0$. For $k = 2, \dots, n$,

$$\begin{array}{rcl} \pi_{kk} & = & w_k\left[\gamma(k) - \sum_{j=1}^{k-1}\pi_{k-1,j}\gamma(k-j)\right]/v_{k-1} \\ \pi_{kj} & = & \pi_{k-1,j} - \pi_{kk}\pi_{k-1,k-j}, \quad j = 1, 2, \dots, k-1, \\ v_k & = & (1 - \pi_{kk}^2)v_{k-1} \end{array} \tag{10}$$

where $\{w_k, k = 0, 1, \dots, n-1\}$ is a weight sequence. The sequence $\{\pi_{kk}, k = 1, \dots, n-1\}$ is the sequence of tapered partial autocorrelations. If $w_k = 1$, for all k, we recover the usual DL recursion yielding the raw partial autocorrelations. If $w_k = 1$ for $k \le p$, and $w_k = 0$ for $k > p$, then (10) amounts to fitting an autoregressive model of order p. Notice that the numerator of π_{kk} is the deviation of the k-th autocovariance from that predicted from an AR$(k-1)$ process. Notice also that π_{kj} is shrunk towards $\pi_{k,j-1}$.

We consider shrinking towards zero the discrepancy by using either a trapezoidal kernel as in (McMurry and Politis, 2010) and (McMurry et al., 2015), or a Poisson kernel. The Poisson kernel (Zygmund (2002), p. 96) (normalized so that $\kappa_P(0) = 1$) is

$$\kappa_P(u) = \frac{(1-r)^2}{1 + r^2 - 2r\cos u},$$

where $r \in (0, 1)$ is the bandwidth parameter. When applied to the j-th partial autocorrelation, $w_j = \kappa_P(j/n)$. Notice that this is the correct way of tapering the PACF: if a rectangular kernel is applied, i.e. $w_k = 1, k \ge p, w_k = 0, k > p$, the system 10 computes the usual Yule-Walker predictor based on p lagged values.

The implied autocovariance sequence is obtained from the PACF sequence $\{\pi_{kk}, k = 1, 2, \dots, n\}$, by reversing the DL recursions Tunnicliffe Wilson (1979). For $k = n-1, 2, \dots, 1$,

$$\begin{aligned} \pi_{kj} &= \left[\pi_{k+1,j} - \pi_{k+1,k+1}\pi_{k+1,k+1-j}\right]/(1-\pi_{k+1,k+1}^2), \quad j = k-1, \dots, 1, \\ v_k &= v_{k+1}/(1-\pi_{k+1,k+1}^2) \end{aligned} \tag{11}$$

$v_0 = v_1/(1-\pi_{1,1}^2)$.
Then, setting $\gamma(0) = v_0$ and $\gamma(1) = v_0\pi_{11}$, the lag k autocovariances for $k > 1$ are obtained as follows

$$\gamma(k) = \sum_{j=1}^{k-1} \pi_{k-1,j}\gamma(k-j) + v_{k-1}\pi_{kk}. \tag{12}$$

Our estimator of the covariance matrix is thus obtained by running the tapered Durbin-Levinson recursions (10) on the raw sample autocovariances, $\{\hat{\gamma}(k), k = 0, \dots, n-1\}$ and obtained the regularized autocovariance sequence from (11)-(12). In the sequel, we will denote the autocovariance sequence by $\{\tilde{\gamma}(k), k = 0, \dots, n-1\}$ and the corresponding Toeplitz autocovariance matrix by $\tilde{\Gamma}_n$.

The weighting function depends on a bandwidth parameter, which corresponds to the banding parameter l in the case of the trapezoidal kernel, and to the parameter r for the Poisson kernel. In the sequel we denote as PG the estimator obtained by the trapezoidal kernel, whereas PGp the estimator obtained by the Poisson kernel. In the next section we propose to select the bandwidth by leave-one-out crossvalidation.

The above recursions deliver the elements of the UDU decomposition of the inverse autocovariance function

$$\tilde{\Gamma}_n^{-1} = C_n' D_n C_n,$$

where

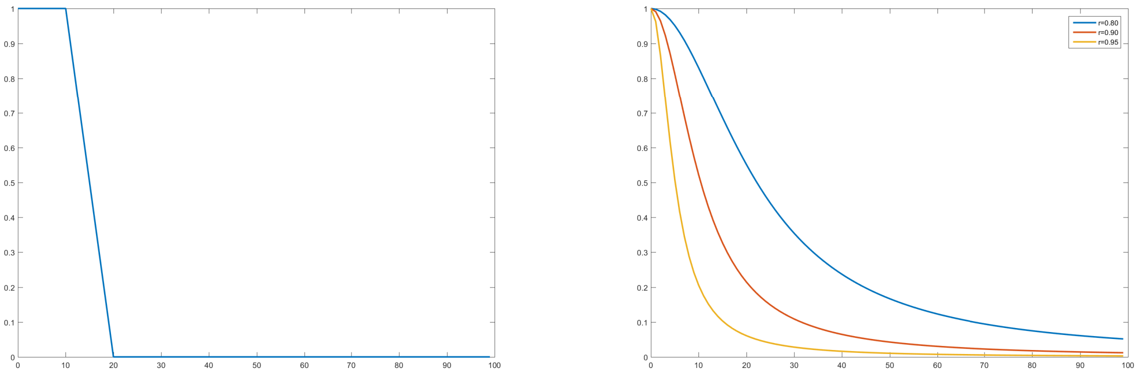
$$C_n = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\tilde{\pi}_{11} & 1 & 0 & \cdots & 0 \\ -\tilde{\pi}_{22} & -\tilde{\pi}_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \ddots & \vdots \\ -\tilde{\pi}_{n-1,n-1} & -\tilde{\pi}_{n-1,n-2} & -\tilde{\pi}_{n-1,n-3} & \cdots & 1 \end{bmatrix}, \quad D_n = \mathrm{diag}\left(\tilde{v}_0^{-1}, \tilde{v}_1^{-1}, \dots, \tilde{v}_{n-1}^{-1}\right). \tag{13}$$

5.1 Crossvalidatory Estimate of Bandwidth

Let $\tilde{Y}_{t\backslash t}$ denote the optimal linear prediction of Y_t using all the available remaining observations, which is the set $\{y_1, \dots, y_{t-1}, y_{t+1}, \dots, y_n\}$, and let $U_t = (Y_t - \tilde{Y}_{t\backslash t})$ denote the interpolation, or crossvalidation error, and let σ_u denote its standard error.

The optimal interpolator weights can be computed from the elements of the inverse Γ_t^{-1} and thus can be estimated from the output of the regularized Durbin-Levinson system as follows. If $\tilde{\Gamma}_n^{-1} = \{\tilde{\gamma}^{(ij)}, i, j = 1, \dots, n\}$, then

$$y_t - \tilde{y}_{t\backslash t} = \sum_{j=1}^{n} \frac{\tilde{\gamma}^{(tj)}}{\tilde{\gamma}^{(tt)}} y_j;$$



(a) Trapezoidal kernel **(b)** Poisson kernel

Figure 1

morever, $\tilde{\sigma}^2_{u,t} = 1/\tilde{\gamma}^{(tt)}$.

In terms of the elements of the regularized Durbin-Levinson recursions,

$$\tilde{\sigma}^2_{u,t} = \left[\frac{1}{\tilde{v}_{t-1}} + \sum_{j=1}^{n-t} \frac{\tilde{\pi}^2_{t+j-1,j}}{\tilde{v}_{t+j-1}} \right]^{-1},$$

$$u_t = \tilde{\sigma}^2_{u,t} c'_{t,n} D_n C_n y,$$

where $c_{t,n} = [0, \dots, 0, 1, -\tilde{\pi}_{t,1}, -\tilde{\pi}_{t+1,2}, \dots, -\tilde{\pi}_{n-1,n-t}]'$ is the t-th column of the matrix C_n and $y = [y_1, \dots, y_n]$.

The leave-one-out crossvalidatory estimate of the bandwidth parameter is the value minimising the crossvalidation score

$$CV = \sum_t \frac{u_t^2}{\tilde{\sigma}^2_{u,t}}.$$

5.2 Bootstrap

- Compute $e = D_n^{1/2} C_n y$, where e is the vector with elements

$$e_t = \left(y_t - \sum_{j=1}^{t-1} \tilde{\pi}_{t-1,j} y_{t-j} \right) / \sqrt{\tilde{v}_{t-1}}, \quad e_1 = y_1 / \sqrt{\tilde{v}_0}.$$

- Center and sample e with replacement; let e^* denote the bootstrap sample.
- Compute the bootstrap sample $y^* = C_n^{-1} D_n^{-1/2} e^*$, where the generic element is

$$y_t^* = \sum_{j=1}^{t-1} \tilde{\pi}_{t-1,j} y^*_{t-j} + \sqrt{\tilde{v}_{t-1}} e_t^*, \quad y_1^* = \sqrt{\tilde{v}_0} e_1^*.$$

6 A simulated example: ARMA Models

Following (Wu and Pourahmadi, 2009) and (McMurry and Politis, 2010), we conducted several simulations to assess the performance of our estimators compared to with the method of (McMurry and Politis, 2010) reported in (8). Results are based on 1000 replications. For our estimators (PG and PGp) the choice of the optimal bandwidth parameters is performed by the cross validation approach proposed in Section 5.1, whereas for the estimator proposed by MP we use their rule reported in (9), where $c = 2$ and $K_n = 5$. Each model we considered is tested with sample sizes $n = 100$, 250, 500, and 750. We compare performances by computing a matrix norm of the difference between the estimated matrix and the true, namely $\hat{\Gamma}_{n,l}$ and Γ_n respectively, using the infinity matrix norm defined as $|||A|||_\infty := \max\limits_{i \in 1,\ldots n} \sum_{j=1}^{n} |a_{ij}|$ and the operator norm defined as the largest eigenvalue of the estimation error matrix $\rho(\hat{\Gamma}_{n,l} - \Gamma_n)$.

6.1 AR(1)

Following (McMurry and Politis, 2010), in the first simulation the data were generated by the autoregressive process $X_t = \phi X_{t-1} + \epsilon_t$, with ϵ_t an i.i.d. sequence of $N(0,1)$ random variables. For each simulation ϕ assumes the values $\pm 0.9, \pm 0.7, \pm 0.5, \pm 0.3, \pm 0.1$. The results are reported in Table 1. The shaded entries in the table indicate that our estimators outperform the MP approach. The results show the PG as the best estimator especially when the autoregressive coefficient ϕ takes values equals to $\pm 0.9, \pm 0.7, \pm 0.5$. This is an expected result since PG and PGp operate directly on the PACF.

6.2 MA(1)

In the second simulation the data were generated by a moving average process $X_t = \theta \epsilon_{t-1} + \epsilon_t$, with ϵ_t an i.i.d. sequence of $N(0,1)$ random variables. For each simulation θ assumes the values $\pm 0.9, \pm 0.7, \pm 0.5, \pm 0.3, \pm 0.1$. The results are reported in Table 2. In this case the MP estimator has the best performances since it operates directly on ACF.

6.3 ARMA(1,1)

In the third simulation the data were generated by an autoregressive moving average process ARMA(1,1), $X_t = \phi X_{t-1} + \theta \epsilon_{t-1} + \epsilon_t$, with ϵ_t an i.i.d. sequence of $N(0,1)$ random variables. For each simulation we set the value of $\phi = \pm 0.9, \pm 0.7, \pm 0.5, \pm 0.3, \pm 0.1$. Once ϕ is fixed we let vary $\theta = \pm 0.9, \pm 0.7, \pm 0.5, \pm 0.3, \pm 0.1$. Cases in which common factors are present are not included in the tables. The results are presented in Tables 3 - 12. Simulation results corresponding to $\pm 0.9, \pm 0.7, \pm 0.5$ and $\theta = \pm 0.3, \pm 0.1$ show that our estimators have a significant improvement with respect to the MP one. This confirms the previously findings in the case of an AR(1) see Table 1.

6.4 Beveridge-Nelson Decomposition

In the fourth simulation, we generate X_t according to a trend plus cycle model with orthogonal components, such that the trend is a random walk with drift and the cycle is an ARMA(2,1) process with complex AR roots. In particular, $X_t = T_t + C_t$, $\Delta T_t = 0.1 + \eta_t, \eta_t \sim \text{NID}(0, 0.05)$ and $(1 - 1.78L + 0.81L^2)C_t = (1 - 0.80L)\kappa_t, \kappa_t \sim \text{NID}(0,1)$, with η_t independent of κ_t. The cycle has a period of 40 observations.

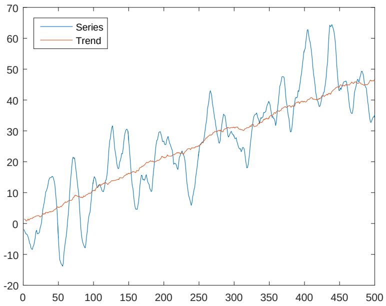
A simulated series $x_t = \tau_t + c_t$ of length $n = 500$ is presented in panel (a) of Figure 2, along with τ_t. For this particular simulation, we can estimate the efficiency of the MP and PG estimator as

$$\hat{G}(h) = 100 \times \left(1 - \frac{\widehat{\text{MSFE}}_{PG}(h)}{\widehat{\text{MSFE}}_{MP}(h)}\right). \tag{14}$$

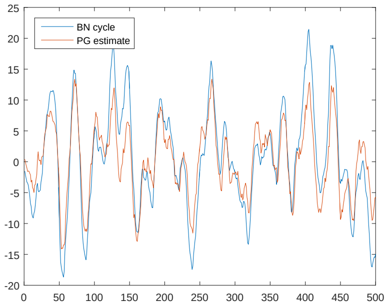
The underlying BN trend can be shown to be

$$m_t = \frac{\theta(1)\phi(L)}{\phi(1)\theta(L)} x_t,$$

see (Proietti, 2006), where $\phi(L) = 1 - 1.78L + 0.81L^2$ and $\theta(L)$ is the MA(2) polynomial of the reduced form model $\phi(L)\Delta X_t = 0.1 + \theta(L)\epsilon_t$. Table 13 reports the quantiles for the relative



(a) Simulated series and trend



(b) True BN cycle and PG estimate

Figure 2

efficiency of the PGp estimator of the trend (notice that this is the same as the efficiency of the cycle estimator). The main evidences are that the gains from using the PGp estimator are substantial for much of the distribution of $\hat{G}(h)$ over 1000 replications. This result suggests that for processes with high dynamic range, such as a cyclical process, regularising the PACF may be more effective than tapering the ACV.

7 Conclusions

- We have proposed an alternative regularized estimator of the autocovariance function of a random process, based on a modification of the Durbin-Levinson algorithm.
- This is more efficient than available tapered estimators for processes with high spectral dynamic range (cyclical processes).
- We estimate the bandwidth of the tapering function by crossvalidation.
- Crossvalidation and bootstrap arise quite naturally as a by product of the regularized DL algorithm.
- We believe that the multivariate extension, providing additional cross-sectional information, can reduce the uncertainty surrounding the cycle estimates.

Table 1: AR1

ϕ	∞-MP	∞-PG	∞-PGp	Op-MP	Op-PG	Op-PGp
			$n = 100$			
-0.9	14.554	12.341	13.460	12.362	11.188	11.959
-0.7	3.125	2.537	3.148	2.769	2.394	2.715
-0.5	1.419	1.215	1.674	1.220	1.117	1.332
-0.3	0.631	0.709	1.064	0.529	0.620	0.754
-0.1	0.321	0.551	0.897	0.299	0.452	0.569
0.1	0.318	0.521	0.872	0.293	0.425	0.540
0.3	0.620	0.670	1.029	0.525	0.583	0.712
0.5	1.374	1.110	1.469	1.200	1.018	1.158
0.7	2.968	2.484	2.839	2.675	2.358	2.480
0.9	12.261	11.267	11.268	10.730	10.304	10.095
			$n = 250$			
ϕ	∞-MP	∞-PG	∞-PGp	Op-MP	Op-PG	Op-PGp
-0.9	11.967	8.926	11.482	10.361	8.613	9.944
-0.7	2.126	1.644	2.663	1.889	1.588	2.093
-0.5	0.936	0.710	1.384	0.817	0.660	0.973
-0.3	0.483	0.433	1.030	0.405	0.384	0.632
-0.1	0.201	0.308	0.892	0.182	0.263	0.469
0.1	0.203	0.317	0.879	0.185	0.266	0.462
0.3	0.477	0.433	1.045	0.407	0.386	0.612
0.5	0.946	0.738	1.433	0.835	0.691	0.974
0.7	2.163	1.672	2.546	1.958	1.621	1.990
0.9	9.995	7.807	9.791	8.903	7.578	8.573
			$n = 500$			
ϕ	∞-MP	∞-PG	∞-PGp	Op-MP	Op-PG	Op-PGp
-0.9	9.712	6.177	9.784	8.447	6.096	8.163
-0.7	1.613	1.151	2.562	1.419	1.116	1.781
-0.5	0.698	0.537	1.521	0.602	0.502	0.942
-0.3	0.402	0.306	1.088	0.338	0.272	0.580
-0.1	0.146	0.212	0.946	0.131	0.181	0.424
0.1	0.148	0.217	0.940	0.133	0.185	0.415
0.3	0.402	0.305	1.094	0.340	0.271	0.572
0.5	0.713	0.537	1.486	0.623	0.505	0.907
0.7	1.623	1.192	2.578	1.437	1.156	1.788
0.9	8.523	6.030	9.217	7.590	5.958	7.527
			$n = 750$			
ϕ	∞-MP	∞-PG	∞-PGp	Op-MP	Op-PG	Op-PGp
-0.9	7.998	4.858	9.102	7.002	4.813	6.993
-0.7	1.372	0.968	2.715	1.200	0.941	1.752
-0.5	0.585	0.429	1.544	0.502	0.402	0.865
-0.3	0.360	0.254	1.096	0.305	0.226	0.533
-0.2	0.213	0.210	1.054	0.180	0.183	0.463
-0.1	0.122	0.175	1.042	0.107	0.149	0.423
0.1	0.125	0.178	0.971	0.110	0.152	0.395
0.3	0.365	0.256	1.153	0.309	0.228	0.555
0.5	0.577	0.430	1.513	0.492	0.402	0.847
0.7	1.366	0.970	2.699	1.214	0.942	1.730
0.9	7.403	5.108	9.042	6.568	5.055	6.937

Notes: Each entry in the table represents losses in the matrix infinity norm for the AR(1) processes. Losses are calculated for the trapezoidal kernel as proposed in (McMurry and Politis, 2010) denoted by MP, and for our estimators PG and PGp reported in (10)-(12)

Table 2: MA1

θ	∞-MP	∞-PG	∞-PGp	Op-MP	Op-PG	Op-PGp
			$n = 100$			
-0.9	0.483	1.656	1.499	0.426	1.140	0.962
-0.7	0.418	1.237	1.153	0.380	0.910	0.778
-0.5	0.392	0.931	1.005	0.373	0.725	0.705
-0.3	0.329	0.663	0.938	0.320	0.541	0.618
-0.1	0.302	0.534	0.877	0.293	0.436	0.543
0.1	0.278	0.505	0.871	0.275	0.413	0.537
0.3	0.347	0.676	0.989	0.337	0.554	0.681
0.5	0.382	0.935	1.066	0.361	0.724	0.751
0.7	0.456	1.260	1.194	0.416	0.925	0.823
			$n = 250$			
θ	∞-MP	∞-PG	∞-PGp	Op-MP	Op-PG	Op-PGp
-0.9	0.287	1.353	1.451	0.253	0.886	0.805
-0.7	0.292	0.943	0.788	0.271	0.668	0.525
-0.5	0.227	0.624	0.705	0.223	0.483	0.478
-0.3	0.201	0.417	0.843	0.198	0.342	0.502
-0.1	0.181	0.310	0.883	0.179	0.259	0.460
0.1	0.181	0.310	0.874	0.179	0.260	0.460
0.3	0.195	0.400	0.839	0.192	0.331	0.503
0.5	0.239	0.637	0.710	0.232	0.489	0.500
0.7	0.266	0.917	0.780	0.244	0.652	0.522
0.8	0.287	1.125	1.002	0.256	0.767	0.615
			$n = 500$			
θ	∞-MP	∞-PG	∞-PGp	Op-MP	Op-PG	Op-PGp
-0.9	0.198	1.231	1.279	0.174	0.768	0.666
-0.7	0.186	0.786	0.545	0.178	0.552	0.364
-0.5	0.157	0.472	0.490	0.154	0.357	0.342
-0.3	0.136	0.299	0.706	0.135	0.247	0.393
-0.1	0.128	0.220	0.925	0.126	0.186	0.418
0.1	0.122	0.216	0.948	0.121	0.183	0.425
0.3	0.145	0.300	0.745	0.144	0.250	0.411
0.5	0.160	0.466	0.505	0.157	0.353	0.358
0.7	0.192	0.764	0.543	0.182	0.531	0.368
0.9	0.199	1.255	1.345	0.176	0.774	0.692
			$n = 750$			
θ	∞-MP	∞-PG	∞-PGp	Op-MP	Op-PG	Op-PGp
-0.9	0.154	1.156	1.165	0.138	0.700	0.588
-0.7	0.139	0.675	0.430	0.134	0.466	0.289
-0.5	0.130	0.400	0.369	0.127	0.301	0.270
-0.3	0.113	0.250	0.591	0.112	0.207	0.320
-0.1	0.100	0.179	0.979	0.100	0.151	0.401
0.1	0.103	0.179	0.958	0.102	0.151	0.398
0.3	0.109	0.245	0.677	0.109	0.199	0.351
0.5	0.131	0.401	0.397	0.129	0.299	0.289
0.7	0.138	0.673	0.435	0.133	0.459	0.293
0.9	0.172	1.182	1.166	0.151	0.704	0.587

Notes: see Table 1.

Table 3: ARMA: $\phi = -0.1$

ϕ	θ	∞-MP	∞-PG	∞-PGp	Op-MP	Op-PG	Op-PGp	∞-MP	∞-PG	∞-PGp	Op-MP	Op-PG	Op-PGp
		$n = 100$						$n = 250$					
-0.1	-0.9	0.520	1.559	1.401	0.419	1.032	0.871	0.363	1.280	1.377	0.297	0.825	0.749
-0.1	-0.7	0.502	1.160	1.102	0.406	0.837	0.727	0.326	0.826	0.762	0.265	0.576	0.485
-0.1	-0.5	0.427	0.801	0.964	0.361	0.623	0.648	0.297	0.560	0.745	0.247	0.430	0.472
-0.1	-0.3	0.343	0.583	0.875	0.299	0.457	0.538	0.226	0.354	0.866	0.195	0.289	0.481
-0.1	0.3	0.511	0.798	1.109	0.441	0.663	0.787	0.304	0.466	0.819	0.262	0.393	0.543
-0.1	0.5	0.549	1.086	1.118	0.461	0.855	0.822	0.366	0.702	0.754	0.308	0.547	0.561
-0.1	0.7	0.620	1.461	1.310	0.512	1.070	0.929	0.411	1.015	0.797	0.328	0.735	0.560
-0.1	0.9	0.700	1.833	1.649	0.572	1.285	1.086	0.460	1.554	1.512	0.361	1.034	0.879
		$n = 500$						$n = 750$					
-0.1	-0.9	0.268	1.121	1.300	0.219	0.688	0.642	0.235	1.058	1.182	0.194	0.630	0.568
-0.1	-0.7	0.255	0.657	0.529	0.211	0.458	0.343	0.220	0.591	0.426	0.182	0.404	0.272
-0.1	-0.5	0.208	0.406	0.530	0.172	0.300	0.332	0.193	0.348	0.386	0.161	0.259	0.264
-0.1	-0.3	0.170	0.253	0.810	0.145	0.209	0.391	0.152	0.218	0.777	0.129	0.179	0.349
-0.1	0.3	0.233	0.343	0.643	0.199	0.285	0.414	0.204	0.294	0.546	0.176	0.242	0.348
-0.1	0.5	0.291	0.556	0.489	0.251	0.429	0.376	0.248	0.484	0.411	0.215	0.371	0.322
-0.1	0.7	0.340	0.863	0.555	0.263	0.607	0.393	0.296	0.791	0.441	0.228	0.537	0.311
-0.1	0.9	0.362	1.393	1.336	0.270	0.858	0.706	0.326	1.325	1.175	0.240	0.793	0.626

Notes: xxxxxx

Table 4: ARMA: $\phi = -0.3$

ϕ	θ	∞-MP	∞-PG	∞-PGp	Op-MP	Op-PG	Op-PGp	∞-MP	∞-PG	∞-PGp	Op-MP	Op-PG	Op-PGp
					n=100						**n=250**		
-0.3	-0.9	0.671	1.332	1.267	0.507	0.896	0.761	0.519	1.073	1.353	0.398	0.669	0.699
-0.3	-0.7	0.607	0.955	1.044	0.458	0.685	0.654	0.456	0.658	0.768	0.351	0.460	0.448
-0.3	-0.5	0.463	0.633	0.905	0.363	0.486	0.563	0.332	0.407	0.813	0.261	0.313	0.453
-0.3	-0.1	0.501	0.623	0.958	0.418	0.539	0.650	0.378	0.373	0.988	0.313	0.327	0.57
-0.3	0.1	0.722	0.806	1.169	0.610	0.701	0.850	0.559	0.489	0.986	0.480	0.436	0.647
-0.3	0.5	1.139	1.423	1.439	0.985	1.137	1.178	0.768	0.997	0.845	0.665	0.795	0.719
-0.3	0.7	1.154	1.818	1.562	0.99	1.368	1.176	0.788	1.454	0.964	0.678	1.057	0.742
-0.3	0.9	1.173	2.258	1.991	1.016	1.596	1.388	0.814	2.013	1.645	0.695	1.341	1.018
					n=500						**n=750**		
-0.3	-0.9	0.424	0.925	1.261	0.330	0.553	0.595	0.367	0.874	1.133	0.284	0.513	0.515
-0.3	-0.7	0.397	0.524	0.586	0.306	0.352	0.336	0.348	0.456	0.445	0.270	0.300	0.272
-0.3	-0.5	0.283	0.312	0.728	0.222	0.238	0.369	0.256	0.271	0.641	0.201	0.205	0.318
-0.3	-0.1	0.311	0.261	1.042	0.258	0.228	0.498	0.281	0.210	1.161	0.233	0.185	0.503
-0.3	0.1	0.464	0.357	0.989	0.400	0.323	0.582	0.387	0.285	0.921	0.333	0.257	0.519
-0.3	0.5	0.508	0.801	0.626	0.436	0.627	0.540	0.397	0.737	0.549	0.344	0.567	0.474
-0.3	0.7	0.536	1.283	0.651	0.467	0.905	0.513	0.436	1.188	0.534	0.380	0.846	0.428
-0.3	0.9	0.546	1.883	1.396	0.468	1.197	0.845	0.422	1.767	1.201	0.367	1.106	0.739

Notes: xxxxxx

Table 5: ARMA: $\phi = -0.5$

ϕ	θ	∞-MP	∞-PG	∞-PGp	Op-MP	Op-PG	Op-PGp	∞-MP	∞-PG	∞-PGp	Op-MP	Op-PG	Op-PGp
					n=100						**n=250**		
-0.5	-0.9	0.820	1.120	1.187	0.589	0.757	0.697	0.662	0.869	1.244	0.491	0.553	0.630
-0.5	-0.7	0.634	0.740	0.933	0.464	0.518	0.567	0.511	0.544	0.845	0.380	0.376	0.459
-0.5	-0.3	0.784	0.732	1.109	0.626	0.638	0.786	0.626	0.457	1.122	0.501	0.398	0.669
-0.5	-0.1	1.223	0.991	1.431	1.040	0.907	1.101	0.890	0.649	1.442	0.773	0.602	0.961
-0.5	0.1	1.537	1.407	1.737	1.347	1.291	1.455	1.018	0.810	1.337	0.901	0.752	1.024
-0.5	0.3	1.714	1.607	1.992	1.512	1.420	1.755	1.114	1.076	1.430	0.982	0.958	1.260
-0.5	0.7	1.876	2.603	2.185	1.658	2.027	1.803	1.211	2.186	1.318	1.071	1.666	1.105
-0.5	0.9	1.964	3.183	2.574	1.736	2.344	1.957	1.290	2.748	1.885	1.147	1.930	1.347
					n=500						**n=750**		
-0.5	-0.9	0.559	0.760	1.257	0.407	0.448	0.571	0.474	0.713	1.080	0.341	0.407	0.476
-0.5	-0.7	0.454	0.437	0.795	0.345	0.301	0.400	0.419	0.391	0.672	0.319	0.268	0.345
-0.5	-0.3	0.544	0.345	1.202	0.444	0.295	0.625	0.481	0.290	1.181	0.394	0.244	0.568
-0.5	-0.1	0.633	0.448	1.576	0.540	0.417	0.912	0.543	0.379	1.786	0.459	0.352	0.927
-0.5	0.1	0.755	0.589	1.026	0.655	0.549	0.760	0.620	0.484	0.921	0.539	0.449	0.654
-0.5	0.3	0.819	0.868	1.192	0.726	0.754	1.075	0.719	0.791	1.133	0.637	0.686	1.027
-0.5	0.5	0.883	1.359	1.062	0.782	1.096	0.970	0.737	1.302	0.975	0.654	1.072	0.889
-0.5	0.7	0.863	1.984	0.858	0.765	1.492	0.743	0.754	1.840	0.730	0.660	1.397	0.644
-0.5	0.9	0.899	2.586	1.564	0.794	1.784	1.096	0.749	2.552	1.285	0.659	1.731	0.928

Notes: xxxxxx

Table 6: ARMA: $\phi = -0.7$

ϕ	θ	∞-MP	∞-PG	∞-PGp	Op-MP	Op-PG	Op-PGp	∞-MP	∞-PG	∞-PGp	Op-MP	Op-PG	Op-PGp
					n=100						**n=250**		
-0.7	-0.9	0.926	1.059	1.112	0.665	0.673	0.643	0.809	0.841	1.191	0.618	0.552	0.613
-0.7	-0.5	1.467	1.148	1.418	1.203	0.982	1.111	1.237	0.901	1.370	1.019	0.748	0.930
-0.7	-0.3	2.435	1.840	2.394	2.138	1.722	1.996	1.705	1.320	2.323	1.482	1.239	1.731
-0.7	-0.1	2.990	2.472	3.148	2.638	2.352	2.682	2.023	1.546	2.957	1.787	1.498	2.177
-0.7	0.1	3.433	2.746	3.310	3.044	2.580	2.953	2.265	1.788	2.404	2.020	1.704	2.123
-0.7	0.3	3.644	3.105	3.973	3.227	2.835	3.647	2.589	2.291	3.096	2.321	2.104	2.900
-0.7	0.5	3.599	3.732	4.038	3.200	3.245	3.696	2.776	3.069	3.127	2.464	2.666	2.952
-0.7	0.9	3.871	4.870	4.155	3.438	3.930	3.545	2.807	4.631	2.924	2.520	3.670	2.464
					n=500						**n=750**		
-0.7	-0.9	0.741	0.720	1.188	0.584	0.483	0.557	0.686	0.658	1.076	0.534	0.422	0.487
-0.7	-0.5	0.973	0.756	1.278	0.792	0.631	0.806	0.813	0.655	1.221	0.651	0.548	0.748
-0.7	-0.3	1.242	1.007	2.225	1.076	0.933	1.481	1.030	0.833	2.239	0.877	0.768	1.361
-0.7	-0.1	1.604	1.128	3.334	1.418	1.093	2.173	1.313	0.947	3.544	1.152	0.918	2.171
-0.7	0.1	1.760	1.204	1.780	1.558	1.146	1.544	1.493	1.095	1.543	1.306	1.046	1.301
-0.7	0.3	1.891	1.830	2.836	1.663	1.630	2.662	1.648	1.712	2.423	1.475	1.530	2.251
-0.7	0.5	1.897	2.785	2.504	1.668	2.429	2.355	1.600	2.643	2.421	1.416	2.337	2.269
-0.7	0.9	2.031	4.397	2.269	1.792	3.480	1.901	1.628	4.326	1.780	1.425	3.484	1.510

Notes: xxxxxx

Table 7: ARMA: $\phi = -0.9$

ϕ	θ	∞-MP	∞-PG	∞-PGp	Op-MP	Op-PG	Op-PGp	∞-MP	∞-PG	∞-PGp	Op-MP	Op-PG	Op-PGp
					n=100						**n=250**		
-0.9	-0.7	6.297	5.323	4.926	5.236	4.531	4.248	5.021	4.324	4.369	4.356	3.909	3.863
-0.9	-0.5	10.730	9.877	9.859	9.083	8.761	8.735	8.753	6.788	8.492	7.705	6.372	7.370
-0.9	-0.3	13.172	11.761	12.685	11.163	10.604	11.155	10.261	8.389	11.769	9.025	7.993	9.853
-0.9	-0.1	14.966	13.756	14.267	12.688	12.423	12.567	11.949	8.936	12.939	10.357	8.620	10.799
-0.9	0.1	15.148	12.982	14.403	12.831	11.730	12.855	11.561	8.465	10.495	10.043	8.151	9.718
-0.9	0.3	15.216	12.034	16.918	12.776	10.751	14.967	12.730	9.402	14.429	11.069	8.817	13.380
-0.9	0.5	14.925	13.047	16.118	12.604	11.561	14.236	13.610	11.243	14.658	11.735	10.339	13.496
-0.9	0.7	14.656	14.316	15.348	12.302	12.631	13.603	13.051	12.979	11.886	11.280	11.877	10.993
					n=500						**n=750**		
-0.9	-0.7	3.673	3.459	4.181	3.214	3.177	3.755	3.154	2.836	4.099	2.714	2.570	3.689
-0.9	-0.5	6.381	4.945	6.856	5.654	4.665	5.691	5.507	4.319	6.546	4.876	4.088	5.318
-0.9	-0.3	8.422	6.184	11.075	7.468	6.006	8.675	7.164	5.265	11.183	6.375	5.101	8.430
-0.9	-0.1	9.055	6.397	12.685	7.949	6.306	9.592	8.092	5.334	12.887	7.091	5.267	9.300
-0.9	0.1	10.227	6.181	7.698	8.921	6.062	7.209	8.517	5.147	6.303	7.464	5.056	6.003
-0.9	0.3	10.582	7.966	11.410	9.272	7.549	10.891	8.453	7.232	9.419	7.357	6.809	8.995
-0.9	0.5	9.902	10.237	12.893	8.661	9.569	12.189	8.487	10.129	11.589	7.414	9.568	10.973
-0.9	0.7	10.270	12.484	9.755	9.065	11.743	9.200	8.897	12.367	8.319	7.823	11.819	7.805

Notes: xxxxxx

Table 8: ARMA: $\phi = 0.1$

ϕ	θ	∞-MP	∞-PG	∞-PGp	Op-MP	Op-PG	Op-PGp	∞-MP	∞-PG	∞-PGp	Op-MP	Op-PG	Op-PGp
				n=100						**n=250**			
0.1	-0.9	0.649	1.818	1.589	0.527	1.234	1.036	0.461	1.541	1.463	0.364	1.011	0.852
0.1	-0.7	0.647	1.439	1.275	0.541	1.052	0.900	0.423	1.053	0.793	0.345	0.761	0.557
0.1	-0.5	0.523	1.008	1.056	0.448	0.802	0.760	0.382	0.732	0.705	0.332	0.571	0.522
0.1	-0.3	0.455	0.757	0.985	0.397	0.621	0.698	0.307	0.476	0.823	0.267	0.402	0.535
0.1	0.3	0.357	0.608	0.933	0.317	0.489	0.606	0.234	0.370	0.832	0.204	0.305	0.472
0.1	0.5	0.425	0.824	1.025	0.363	0.635	0.696	0.281	0.549	0.762	0.238	0.419	0.484
0.1	0.7	0.492	1.200	1.178	0.409	0.865	0.781	0.319	0.807	0.793	0.262	0.566	0.499
0.1	0.9	0.538	1.551	1.456	0.442	1.053	0.919	0.357	1.264	1.422	0.290	0.824	0.778
				n=500						**n=750**			
0.1	-0.9	0.358	1.414	1.311	0.270	0.887	0.708	0.333	1.334	1.173	0.250	0.816	0.625
0.1	-0.7	0.334	0.891	0.543	0.266	0.625	0.388	0.293	0.798	0.438	0.232	0.554	0.312
0.1	-0.5	0.282	0.535	0.482	0.242	0.416	0.370	0.247	0.477	0.400	0.216	0.365	0.313
0.1	-0.3	0.227	0.342	0.627	0.197	0.286	0.396	0.203	0.298	0.526	0.175	0.246	0.342
0.1	0.3	0.171	0.257	0.811	0.146	0.211	0.399	0.148	0.207	0.842	0.126	0.171	0.383
0.1	0.5	0.228	0.424	0.522	0.190	0.316	0.339	0.189	0.352	0.414	0.157	0.261	0.276
0.1	0.7	0.254	0.668	0.530	0.208	0.455	0.332	0.220	0.586	0.427	0.185	0.401	0.273
0.1	0.9	0.272	1.145	1.353	0.224	0.711	0.675	0.240	1.065	1.207	0.198	0.637	0.577

Notes: xxxxxx

Table 9: ARMA: $\phi = 0.3$

ϕ	θ	∞-MP	∞-PG	∞-PGp	Op-MP	Op-PG	Op-PGp	∞-MP	∞-PG	∞-PGp	Op-MP	Op-PG	Op-PGp
					n=100						**n=250**		
0.3	-0.9	1.160	2.331	1.939	1.002	1.646	1.346	0.815	1.973	1.594	0.708	1.339	1.022
0.3	-0.7	1.158	1.875	1.498	0.999	1.405	1.126	0.808	1.452	0.947	0.697	1.075	0.725
0.3	-0.5	1.050	1.404	1.277	0.911	1.139	1.037	0.768	0.990	0.829	0.669	0.805	0.686
0.3	-0.1	0.733	0.788	1.109	0.640	0.699	0.813	0.567	0.498	1.018	0.491	0.446	0.654
0.3	0.1	0.504	0.636	0.983	0.422	0.543	0.649	0.374	0.377	0.959	0.314	0.331	0.539
0.3	0.5	0.462	0.643	0.921	0.362	0.490	0.578	0.339	0.411	0.875	0.265	0.314	0.478
0.3	0.7	0.588	0.927	1.068	0.452	0.677	0.680	0.462	0.678	0.835	0.353	0.477	0.495
0.3	0.9	0.664	1.338	1.300	0.503	0.903	0.794	0.519	1.054	1.370	0.399	0.665	0.715
					n=500						**n=750**		
0.3	-0.9	0.551	1.846	1.405	0.478	1.197	0.861	0.428	1.755	1.181	0.373	1.119	0.746
0.3	-0.7	0.516	1.254	0.643	0.447	0.906	0.510	0.429	1.186	0.521	0.378	0.847	0.419
0.3	-0.5	0.524	0.812	0.614	0.454	0.642	0.528	0.401	0.737	0.550	0.350	0.579	0.473
0.3	-0.3	0.497	0.523	0.663	0.441	0.448	0.540	0.378	0.446	0.569	0.334	0.380	0.477
0.3	-0.1	0.469	0.358	0.845	0.408	0.323	0.502	0.382	0.282	0.973	0.329	0.253	0.510
0.3	0.1	0.316	0.269	1.080	0.263	0.235	0.523	0.290	0.218	1.093	0.242	0.191	0.490
0.3	0.5	0.281	0.312	0.754	0.219	0.237	0.379	0.258	0.273	0.658	0.203	0.206	0.321
0.3	0.7	0.389	0.534	0.585	0.300	0.359	0.344	0.351	0.464	0.463	0.271	0.310	0.281
0.3	0.9	0.431	0.926	1.312	0.332	0.556	0.619	0.377	0.878	1.220	0.289	0.525	0.551

Notes: xxxxxx

Table 10: ARMA: $\phi = 0.5$

ϕ	θ	∞-MP	∞-PG	∞-PGp	Op-MP	Op-PG	Op-PGp	∞-MP	∞-PG	∞-PGp	Op-MP	Op-PG	Op-PGp
					n=100						**n=250**		
0.5	-0.9	1.873	3.188	2.458	1.663	2.388	1.913	1.281	2.885	1.917	1.136	2.061	1.394
0.5	-0.7	1.830	2.640	2.036	1.629	2.087	1.663	1.201	2.184	1.310	1.063	1.710	1.111
0.5	-0.3	1.659	1.689	1.773	1.484	1.497	1.539	1.129	1.158	1.357	1.000	1.034	1.219
0.5	-0.1	1.473	1.280	1.558	1.305	1.171	1.271	0.995	0.838	1.250	0.880	0.777	0.962
0.5	0.1	1.222	1.013	1.420	1.058	0.925	1.093	0.885	0.636	1.383	0.775	0.590	0.910
0.5	0.2	1.017	0.843	1.244	0.865	0.764	0.915	0.811	0.545	1.283	0.684	0.500	0.803
0.5	0.4	0.516	0.586	0.900	0.419	0.490	0.588	0.397	0.376	0.959	0.316	0.313	0.514
0.5	0.7	0.632	0.761	0.976	0.460	0.538	0.603	0.510	0.543	0.889	0.376	0.376	0.488
0.5	0.9	0.800	1.089	1.202	0.585	0.730	0.712	0.673	0.916	1.258	0.496	0.587	0.654
					n=500						**n=750**		
0.5	-0.9	0.914	2.688	1.545	0.810	1.890	1.115	0.753	2.583	1.346	0.667	1.799	1.006
0.5	-0.7	0.913	1.981	0.897	0.811	1.538	0.777	0.743	1.877	0.690	0.659	1.486	0.607
0.5	-0.3	0.845	0.891	1.173	0.752	0.792	1.046	0.674	0.803	1.078	0.598	0.695	0.967
0.5	-0.1	0.743	0.595	1.051	0.645	0.557	0.761	0.608	0.495	1.033	0.533	0.465	0.696
0.5	0.1	0.654	0.474	1.520	0.561	0.441	0.877	0.533	0.367	1.667	0.456	0.340	0.865
0.5	0.3	0.551	0.346	1.170	0.452	0.299	0.611	0.490	0.299	1.104	0.406	0.255	0.539
0.5	0.7	0.453	0.449	0.750	0.340	0.308	0.383	0.423	0.394	0.641	0.318	0.271	0.334
0.5	0.9	0.559	0.795	1.299	0.415	0.469	0.593	0.485	0.718	1.123	0.346	0.414	0.493

Notes: xxxxxx

Table 11: ARMA: $\phi = 0.7$

ϕ	θ	∞-MP	∞-PG	∞-PGp	Op-MP	Op-PG	Op-PGp	∞-MP	∞-PG	∞-PGp	Op-MP	Op-PG	Op-PGp
				n=100						**n=250**			
0.7	-0.9	3.501	5.003	3.960	3.132	4.181	3.434	2.536	4.802	2.885	2.272	3.911	2.473
0.7	-0.5	3.398	3.923	3.435	3.056	3.532	3.149	2.386	3.035	2.737	2.126	2.712	2.559
0.7	-0.3	3.432	3.365	3.462	3.079	3.108	3.194	2.330	2.257	2.804	2.103	2.089	2.604
0.7	-0.1	3.114	2.703	2.848	2.802	2.544	2.566	2.236	1.810	2.213	2.002	1.734	1.901
0.7	0.1	2.745	2.230	2.689	2.447	2.120	2.284	1.984	1.506	2.881	1.769	1.464	2.129
0.7	0.3	2.328	1.777	2.308	2.056	1.661	1.925	1.624	1.233	2.226	1.434	1.151	1.609
0.7	0.5	1.463	1.188	1.418	1.245	1.034	1.126	1.218	0.890	1.352	1.023	0.750	0.921
0.7	0.9	0.921	1.046	1.110	0.664	0.672	0.652	0.809	0.839	1.188	0.615	0.554	0.610
				n=500						**n=750**			
0.7	-0.9	1.960	4.525	2.193	1.748	3.712	1.908	1.630	4.384	1.820	1.443	3.585	1.588
0.7	-0.5	1.906	2.812	2.399	1.689	2.519	2.235	1.562	2.742	2.319	1.370	2.469	2.153
0.7	-0.3	1.851	1.915	2.563	1.661	1.747	2.380	1.523	1.685	2.355	1.346	1.520	2.167
0.7	-0.1	1.698	1.290	1.722	1.518	1.246	1.473	1.441	1.086	1.470	1.271	1.043	1.244
0.7	0.1	1.518	1.107	3.147	1.348	1.075	2.079	1.270	0.904	3.327	1.126	0.875	1.973
0.7	0.3	1.214	0.970	2.085	1.053	0.899	1.376	1.039	0.826	2.266	0.886	0.756	1.381
0.7	0.5	0.995	0.762	1.289	0.838	0.653	0.822	0.825	0.671	1.198	0.677	0.571	0.737
0.7	0.9	0.738	0.736	1.142	0.572	0.494	0.559	0.684	0.655	1.160	0.530	0.420	0.511

Notes: xxxxxx

Table 12: ARMA: $\phi = 0.9$

ϕ	θ	∞-MP	∞-PG	∞-PGp	Op-MP	Op-PG	Op-PGp	∞-MP	∞-PG	∞-PGp	Op-MP	Op-PG	Op-PGp
				n=100						**n=250**			
0.9	-0.7	12.952	14.981	12.916	11.319	13.632	11.653	10.572	13.339	10.586	9.355	12.560	9.858
0.9	-0.5	12.732	13.896	12.666	11.145	12.674	11.342	10.692	12.076	12.136	9.598	11.453	11.151
0.9	-0.3	12.878	12.924	11.783	11.336	11.878	10.626	10.435	9.916	11.786	9.269	9.407	10.940
0.9	-0.1	12.549	11.599	11.596	11.003	10.624	10.495	10.174	8.346	9.009	9.092	8.086	8.364
0.9	0.1	12.102	10.885	11.657	10.655	9.951	10.355	9.883	8.303	11.563	8.847	8.047	9.607
0.9	0.3	10.918	9.975	10.748	9.506	9.086	9.457	8.904	7.856	10.439	7.993	7.511	8.728
0.9	0.5	9.047	8.523	8.537	7.889	7.669	7.587	7.063	6.319	7.791	6.325	5.963	6.734
0.9	0.7	5.756	5.341	5.005	5.023	4.745	4.450	4.653	4.417	4.491	4.125	4.051	4.059
				n=500						**n=750**			
0.9	-0.7	8.764	13.204	8.530	7.783	12.688	8.025	7.811	12.846	7.801	6.883	12.358	7.275
0.9	-0.5	8.945	10.947	11.501	7.958	10.422	10.805	7.836	10.583	10.323	6.933	10.143	9.689
0.9	-0.3	8.815	8.126	9.972	7.824	7.746	9.443	7.717	7.455	8.407	6.846	7.104	7.949
0.9	-0.1	8.760	6.362	7.235	7.784	6.249	6.852	7.279	5.279	5.644	6.369	5.193	5.383
0.9	0.1	8.308	6.122	11.819	7.405	6.026	8.979	7.354	4.960	12.521	6.570	4.891	8.718
0.9	0.3	7.338	5.824	10.083	6.563	5.658	7.767	6.646	4.961	10.479	5.861	4.805	7.594
0.9	0.5	6.009	4.837	6.935	5.377	4.578	5.872	5.232	4.283	6.611	4.684	4.055	5.427
0.9	0.7	3.653	3.411	4.251	3.225	3.125	3.848	3.070	2.915	4.073	2.634	2.636	3.710

Notes: xxxxxx

Table 13

Quantile	MSE ratio c_t	MSE ratio ψ_t	ℓ	r
0.10	0.0009	0.0009	4	0.83
0.25	0.0033	0.0031	11	0.91
0.50	0.3360	0.3245	17	0.96
0.75	0.9499	0.9703	20	0.97
0.90	1.7332	1.7306	22	0.97

References

Beveridge, S., and Nelson, C. R. (1981). A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle'. *Journal of Monetary economics*, *7*(2), 151–174.

Bickel, P. J., and Gel, Y. R. (2011). Banded regularization of autocovariance matrices in application to parameter estimation and forecasting of time series. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *73*(5), 711–728.

Brockwell, P. J., and Davis, R. A. (1991). *Time series: theory and methods.* Springer Series in Statistics.

Dungey, M., Jacobs, J. P., Tian, J., and van Norden, S. (2015). Trend in cycle or cycle in trend? new structural identifications for unobserved-components models of us real gdp. *Macroeconomic Dynamics*, *19*(04), 776–790.

Hannan, E. J., and Deistler, M. (1988). *The statistical theory of linear systems* (Vol. 70). SIAM.

McMurry, T. L., and Politis, D. N. (2010). Banded and tapered estimates for autocovariance matrices and the linear process bootstrap. *Journal of Time Series Analysis*, *31*(6), 471–482.

McMurry, T. L., Politis, D. N., et al. (2015). High-dimensional autocovariance matrices and optimal linear prediction. *Electronic Journal of Statistics*, *9*, 753–788.

Morley, J. C. (2011). The two interpretations of the beveridge–nelson decomposition. *Macroeconomic Dynamics*, *15*(03), 419–439.

Morley, J. C., Nelson, C. R., and Zivot, E. (2003). Why are the beveridge-nelson and unobserved-components decompositions of gdp so different? *Review of Economics and Statistics*, *85*(2), 235–243.

Newbold, P., Agiakloglou, C., and Miller, J. (1993). Long-term inference based on short-term forecasting models. *Time Series Analysis*, 9–25.

Pourahmadi, M. (2013). *High-dimensional covariance estimation: With high-dimensional data.* John Wiley & Sons.

Proietti, T. (2004). Unobserved components models with correlated disturbances. *Statistical Methods and Applications*, *12*(3), 277–292.

Proietti, T. (2006). Trend–cycle decompositions with correlated components. *Econometric Reviews*, *25*(1), 61–84.

Proietti, T., and Harvey, A. (2000). A beveridge–nelson smoother. *Economics Letters*, *67*(2), 139–146.

Tunnicliffe Wilson, G. (1979). Some efficient computational procedures for high order arma models. *Journal of Statistical Computation and Simulation*, *8*(3-4), 301-309.

Weber, E. (2011). Analyzing us output and the great moderation by simultaneous unobserved components. *Journal of Money, Credit and Banking*, *43*(8), 1579–1597.

Wu, W. B. (2005). Nonlinear system theory: Another look at dependence. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(40), 14150–14154.

Wu, W. B., and Pourahmadi, M. (2009). Banding sample autocovariance matrices of stationary processes. *Statistica Sinica*, *19*(4), 1755.

Zygmund, A. (2002). *Trigonometric series* (Vol. 1 and 2). Cambridge university press.