

Forecasting and Signal Extraction with Misspecified Models

Previsione e stima dei segnali da modelli parzialmente specificati

Tommaso Proietti

Dipartimento di Scienze Statistiche

Università di Udine

proietti@dss.uniud.it

Riassunto: Il lavoro considera la stima adattiva (o multiperiodale) e la validazione incrociata per la stima dei parametri di un modello parzialmente specificato, con l'intento di misurare e prevedere la componente ciclica e tendenziale di una serie temporale. Dopo aver introdotto un'opportuna metrica per valutare l'efficienza nella stima delle componenti, si conclude che soltanto la prima opzione può rivelarsi molto efficiente.

Keywords: Business cycles, Cross-validation, Multistep estimation.

1. Introduction

This paper focusses on forecasting and signal extraction in the presence of model misspecification. The role of multistep estimation for long range forecasting with AR models has been illustrated by Clements and Hendry (1996). We consider time series models that are simple, in that the dimension of the parameter space is limited, and generate predictors and signal extraction filters that are well understood. Even though they are suboptimal, it may be the case that they produce efficient forecasts and signal estimates, when the parameter estimation criteria are modified so as to enhance the features of interest for the specific problem at hand. Due to space constraint we illustrate some results concerning the local level model (LLM).

2. ME and CV for the Local Linear Model

The local level model is formulated as follows:

$$\begin{aligned} y_t &= \mu_t + \epsilon_t, & t = 1, 2, \dots, T, & \epsilon_t \sim \text{WN}(0, \sigma_\epsilon^2) \\ \mu_{t+1} &= \mu_t + \eta_t, & \eta_t &\sim \text{WN}(0, \sigma_\eta^2), \end{aligned} \quad (1)$$

$E(\eta_t \epsilon_{t-j}) = 0, \forall j$. The reduced form is the IMA(1,1) model: $\Delta y_t = \xi_t + \theta \xi_{t-1}, \xi_t \sim \text{WN}(0, \sigma^2)$, with $\theta \leq 0$. In the steady state, the one-step-ahead prediction errors are $\tilde{v}_t = y_t - \tilde{y}_{t|t-1} = \Delta y_t / (1 + \theta L)$, and the level predictions, filtered and smoothed estimates depend on θ in a simple fashion:

$$\tilde{\mu}_{t+1|t} = \tilde{\mu}_{t|t} = \frac{(1 + \theta)}{(1 + \theta L)} y_t, \quad \tilde{\mu}_{t|\infty} = \frac{(1 + \theta)^2}{|1 + \theta L|^2} y_t = \frac{(1 + \theta)}{(1 + \theta L^{-1})} \tilde{\mu}_{t|t}$$

the signal extraction filters are an exponentially weighted average of current and past observations and a two-sided version of it. A constant drift can be added to the right hand side of the level equation yielding a local linear trend model (LLTM).

The l -step ahead forecast error, $\tilde{\nu}_{t+l|t}$, can be written:

$$\tilde{\nu}_{t+l|t} = \left[1 + (1 + \theta)L + (1 + \theta)L^2 + \cdots + (1 + \theta)L^{l-1}\right] \tilde{\nu}_t, \quad l > 1.$$

$$\tilde{\nu}_{t+l|t} = v(L)\Delta y_t, \quad v(L) = S_{l-1}(L) + \frac{1}{1 + \theta L}L^{l-1} \quad (2)$$

where $S_j(L) = 1 + L + L^2 + \cdots + L^{j-1}$.

Multistep estimation determines θ as the minimiser of

$$ME(\theta, l) = \text{Var}(\tilde{\nu}_{t+l|t}) = \frac{1}{\pi} \int_0^\pi |v(e^{-i\omega})|^2 g(\omega) d\omega, \quad (3)$$

where $|v(e^{-i\omega})|^2 = v(e^{-i\omega})v(e^{i\omega})$ is the squared gain the filter $v(L)$, and $g(\omega)$ is the spectral generating function of Δy_t .

ME can thus be viewed as minimising the variance of a *filtered* series; see Haywood and Tunnicliffe Wilson (1997). The results will depend on the properties of the series, an important feature being its order of integration. If y_t is stationary the zero frequency is not informative on θ . On the other hand, for difference stationary series, the filter $v(L)$ emphasises the spectral density around the zero frequency; the concentration of power around the zero frequency increases with l . This suggests that for $I(1)$ series the ME estimate with l large will give more relevance to the long run features of the series (a lowpass filter).

Cross-validation is based on the minimisation of the sum of squares of the interpolation error:

$$y_t - E[y_t | Y_{\setminus t}] = \frac{u_t}{M} = \frac{1}{2} \frac{1 - \theta}{|1 + \theta L|^2} |1 - L|^2 y_t = u(L)\Delta y_t,$$

with $|1 - L|^2 = (1 - L)(1 - L^{-1})$, $|1 + \theta L|^2 = (1 + \theta L)(1 + \theta L^{-1})$ and $u(L) = \frac{1}{2} \frac{1 - \theta}{|1 + \theta L|^2} (1 - L^{-1})$. The CV estimator minimises:

$$CV(\theta) = \frac{1}{\pi} \int_0^\pi |u(e^{-i\omega})|^2 g(\omega) d\omega \quad (4)$$

where $g(\omega)$ is the SGF of Δy_t and $|u(e^{-i\omega})|^2$ is the squared gain of $u(L)$. The filter $u(L)$ has a bandpass, rather than lowpass, nature.

3. Forecast and signal extraction relative efficiency

The relative forecast accuracy at horizon l is the basic measure of performance:

$$RelEff(l) = \frac{ME(\hat{\theta}, l)}{\text{Var}(\nu_{t+l|t})}$$

where the denominator is the true forecast error variance.

As far as the estimation of unobserved components is concerned, suppose the true generating model is $y_t = \mu_t + \epsilon_t$ with orthogonal components, so that, if $g_\mu(L)$ and $g_\epsilon(L)$ denote the autocovariance generating functions (ACGF) of the trend and the cycle,

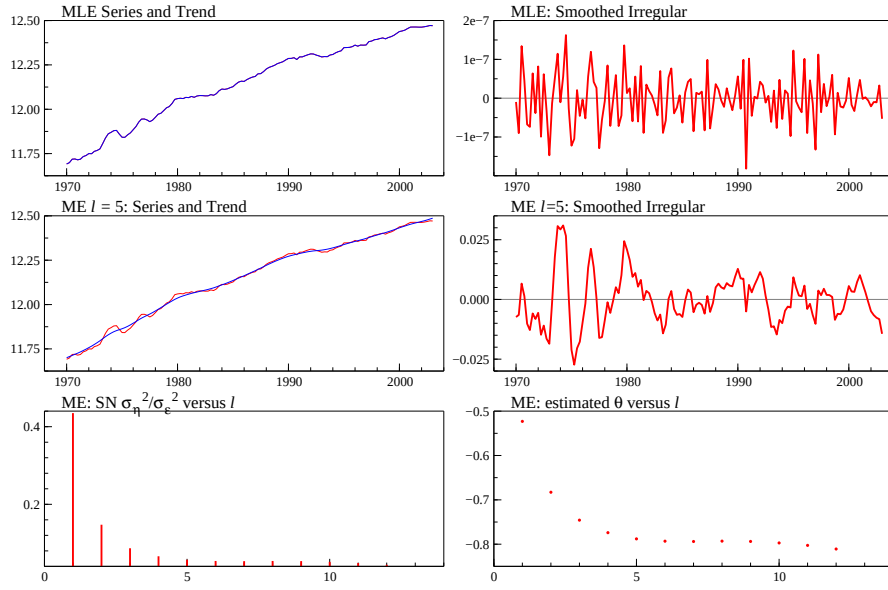


Figure 1: Trend and cycle in Italian real GDP extracted by the LLTM

$g_y(L) = g_\mu(L) + g_\epsilon(L)$. Let also $w_{\mu*}(L)$ and $w_{\epsilon*}(L)$ denote the W-K signal extraction filters for the approximating model $y_t = \mu_t^* + \epsilon_t^*$, where

$$w_{\mu*}(L) = \frac{g_{\mu*}(L)}{g_{\mu*}(L) + g_{\epsilon*}(L)}, \quad w_{\epsilon*}(L) = \frac{g_{\epsilon*}(L)}{g_{\mu*}(L) + g_{\epsilon*}(L)}$$

The ACGF of the unobserved components estimation error, $g_e(L)$, can be expressed in terms of the squared gains of the signal extraction filters and the ACGF of the true components in the following manner:

$$g_e(L) = g_\mu(L)|w_{\epsilon*}(L)|^2 + g_\epsilon(L)|w_{\mu*}(L)|^2. \quad (5)$$

The minimum is the Whittle's true estimation error ACGF: $g_\mu(L)g_\epsilon(L)/g_y(L)$ and its integral over $[0, \pi)$ is the denominator of the relative measure of efficiency employed in the next section, the numerator being the integral of (5) over the same range.

4. Illustrations

The first illustration deals with an application of the LLTM for the extraction of the trend component from the logarithms of Italian GDP, plotted in figure 1. If Gaussian disturbances are assumed for this series, maximum likelihood estimation (MLE) yields an estimate of σ_ϵ^2 that is practically zero, so that all the variation is absorbed by the trend component; see the top row of figure 1. The decomposition is clearly inadequate for the series, as the pattern of residual autocorrelation shows, the latter being suggestive of the presence of a cyclical component. Minimising the variance of the l -step-ahead prediction errors yields different results: the variance of the irregular component grows with l , relative to that of the changes in the trend. The central panels display the smoothed estimates of the components for $l = 5$. The estimated irregular more closely resembles

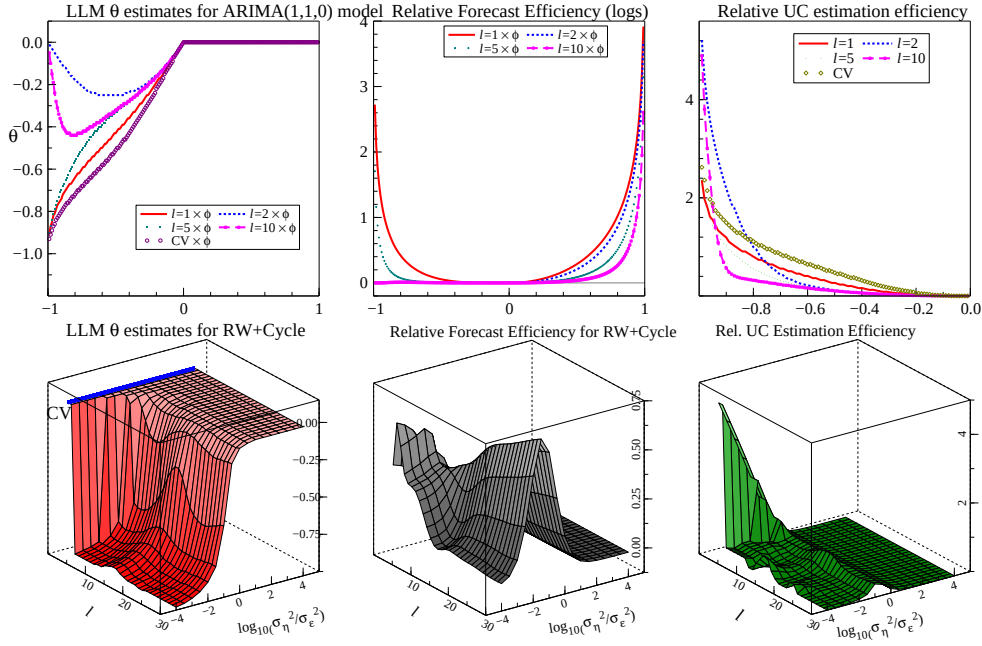


Figure 2: θ estimates and relative efficiency (logs) of $ME(\theta, l)$ and $CV(\theta)$

the cyclical component of GDP. Cross-validation yields the same estimate as maximum likelihood.

Figure 2 presents the θ ME and CV estimates and the relative efficiency for two cases: in the first the true model is $y_t \sim \text{ARIMA}(1,1,0)$: $\Delta y_t = \phi \Delta y_{t-1} + \xi_t$, $\xi_t \sim \text{WN}(0, \sigma^2)$. When ϕ is negative the process can be decomposed into a RW trend plus a stationary AR(1) component, $y_t = \mu_t + \epsilon_t$, where $\Delta \mu_t = \eta_t$, $\eta_t \sim \text{WN}\left(0, \frac{\sigma^2}{(1-\phi)^2}\right)$, $\epsilon_t = \phi \epsilon_t + \kappa_t$, $\kappa_t \sim \text{WN}\left(0, -\frac{\phi \sigma^2}{(1-\phi)^2}\right)$. The W-K estimator of the trend is $\tilde{\mu}_{t|l} = (1 - \phi)^{-2} |1 - \phi L|^2 y_t$, and thus it is provided by a filter with finite impulse response. On the other hand, the trend extraction filter for the LLM has an infinite impulse response.

The second generating process is $y_t = \mu_t + \epsilon_t$, where ϵ_t is the cyclic process:

$$\begin{bmatrix} \epsilon_{t+1} \\ \epsilon_{t+1}^* \end{bmatrix} = \rho \begin{bmatrix} \cos \lambda_c & \sin \lambda_c \\ -\sin \lambda_c & \cos \lambda_c \end{bmatrix} \begin{bmatrix} \epsilon_t \\ \epsilon_t^* \end{bmatrix} + \begin{bmatrix} \kappa_t \\ \kappa_t^* \end{bmatrix}, \quad (6)$$

where $\kappa_t \sim \text{NID}(0, \sigma_\kappa^2)$ and $\kappa_t^* \sim \text{NID}(0, \sigma_\kappa^2)$ are mutually uncorrelated, and uncorrelated with the trend disturbances. We set $\rho = 0.9$ and $\lambda_c = \pi/8$ (cycle period is 4 years of quarterly observations).

References

- Clements M. P. and Hendry D. F. (1996) Multi-step Estimation for Forecasting, *Oxford Bulletin of Economics and Statistics*, 58, 4, 657–684.
- Haywood G. and Tunnicliffe Wilson G. (1997) Fitting time series models by minimizing multistep-ahead errors: a frequency domain approach, *Journal of the Royal Statistical Society, Series B*, 59, 237–254.