EUROPEAN CENTRAL BANK

EUROSYSTEM

Working Paper Series

Marek Jarocinski

A note on implementing the Durbin and Koopman simulation smoother

No 1867 / November 2015

Abstract

The correct implementation of the Durbin and Koopman simulation smoother is explained. A possible misunderstanding is pointed out and clarified for both the basic state space model with a non-zero mean of the initial state and with time-varying intercepts (mean adjustments).

Keywords: state space model; simulation smoother; trend output (*JEL:* C3; C15)

Non-technical summary

Economic models often rely on unobservable, but policy relevant, time-varying variables, such as potential output, output gap, non-accelerating inflation rate of unemployment (NAIRU), and many others. This paper is useful for economists who want to infer the values of such unobservable time-varying variables from observable data.

The paper discusses the technical details of an algorithm, called simulation smoother, that is programmed on the computer and used to characterize the likely values of unobservable variables. By definition, we can never *know* the values of the unobservable variables *precisely*, we can only infer which values of such variables at each point of time are more, and which are less likely in light of the available observable data. This is precisely what this algorithm achieves.

Economists working with unobservable variables use three algorithms that build on one another: (i) the Kalman filter, (ii) the Kalman smoother and (iii) the simulation smoother. Suppose that the unobservable variables are related to some observed variables via a linear model with Gaussian disturbances, and consider a sample of observed variables covering T periods. The Kalman filter returns the T means and variances of the unobservable variables, one at each point of time in the sample, and each conditional on the information in the part of the sample from the beginning up to that point of time. The Kalman smoother returns the T means and variances of the unobservable variables, one at each point of time in the sample, and each conditional on the information *in the whole sample*. Hence, the Kalman smoother characterizes the unobservable variables more reliably. Finally, the simulation smoother generates draws of the unobservable variables from their distribution, taking into account their means and variances at each point of time, which are available from the Kalman smoother, but also their covariances across time, which are not available from the Kalman smoother or filter. The simulation smoother is needed to answer questions involving the covariances of the unobservable variables across time. It is also used for Bayesian inference about the unobservable variables based on the Gibbs sampler. Finally, it is used as a building block for both classical and Bayesian inference in the cases where the assumptions of linearity and Gaussianity are not applicable.

There exist several alternative simulation smoothers and they all, of course, produce the same results, while using different steps. The Durbin and Koopman simulation smoother discussed in this paper is among the fastest and most convenient to implement on the computer.

1 Introduction

Consider the state space model

$$y_t = Z_t\alpha_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, H_t), \tag{1a}$$

$$\alpha_{t+1} = T_t\alpha_t + R_t\eta_t, \quad \eta_t \sim N(0, Q_t), \quad t = 1, ..., T, \quad \text{and} \tag{1b}$$

$$\alpha_1 \sim N(a_1, P_1), \tag{1c}$$

where y_t is the observation vector, α_t is the unobserved state vector, and ε_t and η_t are vectors of disturbances uncorrelated at all lags. The matrices $Z_t, H_t, T_t, R_t, Q_t, P_1$ and vector a_1 are assumed to be known. For further details and illustrations of this model see, e.g., Durbin and Koopman (2012).

A simulation smoother is an algorithm for drawing the states $\alpha = (\alpha_1', ..., \alpha_T')'$, or the disturbances $(\epsilon_1', \eta_1', ..., \epsilon_T', \eta_T)'$, from their distribution conditional on the observables $y = (y_1', ...y_T')'$. This note explains the implementation of the Durbin and Koopman (2002) simulation smoother for this model, pointing out a possible misunderstanding. The misunderstanding may arise when drawing the states. It does not arise when drawing the disturbances.

2 The correct implementation

This section explains how to implement Durbin and Koopman's approach to drawing α conditional on y in the model (1a-1c). Let us call this algorithm 'Algorithm 2a' to differentiate it from their Algorithm 2.

Algorithm 2a. *(modified from Durbin and Koopman (2002) Algorithm 2, p.607)*

Step 1. Draw α^+ and y^+ by means of recursion (1a-1b), where the recursion is initialized *by a draw* $\alpha_1^+ \sim N(0, P_1)$.

Step 2. Construct the artificial series $y^* = y - y^+$ *and compute* $\hat{\alpha}^* = E(\alpha|y^*)$ *by putting* y^* *through the Kalman filter and smoother.*

Step 3. Take $\tilde{\alpha} = \hat{\alpha}^* + \alpha^+$. $\tilde{\alpha}$ *is a draw from the distribution of* α *conditional on* y.

An alternative implementation of this algorithm, which is also correct, uses (1a-1c) for the simulation of y^+, α^+ in Step 1 but then uses the model with $\alpha_1 \sim N(0, P_1)$ to compute the conditional expectation $\hat{\alpha}^* = E(\alpha|y^*)$ in Step 2.

The value added of this note lies in stating the above algorithm explicitly and in particular, in pointing out that a_1 needs to be reset to 0, i.e., the initial condition $\alpha_1 \sim N(a_1, P_1)$ in (1c) needs to be replaced by $\alpha_1 \sim N(0, P_1)$ either in Step 1 or in Step 2. Durbin and Koopman (2002) state Algorithm 2, which is slower, and only suggest Algorithm 2a informally without stating it explicitly. In particular, they do not warn the reader that a_1 should be reset to 0 either in Step 1 or in Step 2, which gives rise to a possible misunderstanding that the unmodified model (1a-1c) can be used both in Step 1 and in Step 2.

Two conditions have the potential to render the above misunderstanding immaterial.

1. Diffuse initialization. Durbin and Koopman (2002) prove in their Appendix 2 that the diffuse elements of α_1^+ can be set equal to arbitrary quantities, hence the values of a_1 corresponding to these elements do not matter.

2. Zero mean. For the elements of α_1^+ that have a zero mean the correction obviously does not matter, since the corresponding values of a_1 equal 0 anyway.

Therefore, the misunderstanding is immaterial when all the elements of α_1 are either diffuse or have a zero mean.

In a model with intercepts another modification of Algorithm 2 is needed. Suppose the model is given by (1c),

$$y_t = d_t + Z_t\alpha_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, H_t) \quad \text{and} \tag{2a}$$

$$\alpha_{t+1} = c_t + T_t\alpha_t + R_t\eta_t, \quad \eta_t \sim N(0, Q_t), \tag{2b}$$

where d_t and c_t are intercepts that are known and may change over time. The remaining quantities are defined under equations (1a-1c). Algorithm 2a can also be used with this model, but the intercepts d_t and c_t should be reset to 0 for all t either in Step 1 or in Step 2.

3 A formal justification

I now provide a formal justification of Algorithm 2a. This algorithm assumes

$$\begin{pmatrix} \alpha \\ y \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_\alpha \\ \mu_y \end{pmatrix}, \begin{pmatrix} \Sigma_{\alpha\alpha} & \Sigma_{\alpha y} \\ \Sigma_{\alpha y} & \Sigma_{yy} \end{pmatrix}\right) \quad \text{and} \quad \begin{pmatrix} \alpha^+ \\ y^+ \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma_{\alpha\alpha} & \Sigma_{\alpha y} \\ \Sigma_{\alpha y} & \Sigma_{yy} \end{pmatrix}\right), \tag{3}$$

where the unconditional moments μ_α, μ_y, $\Sigma_{\alpha\alpha}$, $\Sigma_{\alpha y}$ and Σ_{yy} are functions of Z_t, H_t, T_t, R_t, Q_t, P_1, a_1 (c_t, d_t if applicable) implied by (1a-1c) or by (2a,2b,1c). Note, in particular, that resetting μ_α and μ_y to 0 is achieved by resetting a_1 and, if applicable, c_t and d_t to 0.

A draw $\tilde{\alpha}$ is generated as

$$\tilde{\alpha} = E(\alpha|y^*) + \alpha^+ = \mu_\alpha + \Sigma_{\alpha y}\Sigma_{yy}^{-1}(y - y^+ - \mu_y) + \alpha^+.$$

The first and second moments of $\tilde{\alpha}$ conditional on y are

$$E(\tilde{\alpha}|y) = \mu_\alpha + \Sigma_{\alpha y}\Sigma_{yy}^{-1}(y - \mu_y) = E(\alpha|y) \quad \text{and}$$

$$V(\tilde{\alpha}|y) = \Sigma_{\alpha y}\Sigma_{yy}^{-1}\Sigma_{yy}\Sigma_{yy}^{-1}\Sigma_{\alpha y}' - 2\Sigma_{\alpha y}\Sigma_{yy}^{-1}\Sigma_{\alpha y}' + \Sigma_{\alpha\alpha} = \Sigma_{\alpha\alpha} - \Sigma_{\alpha y}\Sigma_{yy}^{-1}\Sigma_{\alpha y}' = V(\alpha|y).$$

Hence, the first and second moments of $\tilde{\alpha}$ are correct and $\tilde{\alpha}$ is indeed a draw from $p(\alpha|y)$. Note, however, that setting the mean of (α^+, y^+) to (μ_α, μ_y) due to the discussed misunderstanding would have changed the value of $E(\tilde{\alpha}|y)$ and hence would have produced a draw from an incorrect density.

4 Numerical example

I illustrate the effect of the possible misunderstanding using the Watson (1986) model as an example. Watson fits the following model for the real Gross National Product (GNP) of the United States, y_t, observed quarterly from 1949 to 1984.

$$y_t = \tau_t + \varsigma_t, \tag{4a}$$

$$\tau_t = 0.008 + \tau_{t-1} + \eta_t^\tau, \quad \eta_t^\tau \sim N(0, 0.0057^2) \quad \text{and} \tag{4b}$$

$$\varsigma_t = 1.501\varsigma_{t-1} - 0.577\varsigma_{t-2} + \eta_t^\varsigma, \quad \eta_t^\varsigma \sim N(0, 0.0076^2), \tag{4c}$$

where τ_t is a trend and ς_t is a cycle, both unobservable.

Table 1 reports the mean and standard deviation of 10,000 draws of trend GNP, generated with several setups. First, I assume that ς_1 comes from the ergodic distribution of ς_t and that τ_1 is centered at the last value of GNP before the start of the sample, with the ergodic variance of ς_t. This is a natural assumption exploiting the stationarity of ς_t. I generate 10,000 draws using Algorithm 2a. I then generate 10,000 draws with an incorrect variation of this algorithm, where I never reset a_1 and c_t to 0 (neither in Step 1 nor in Step 2). It is clear from Table 1 that the misunderstanding seriously distorts the simulation smoother: the mean of the trend GNP in the first period, τ_1, is 6.24 with the correct algorithm (column 1) and 6.14 with the incorrect variation (column 2). After 50 quarters the initialization matters less and the means of trend GNP in period 50, τ_{50}, obtained with Algorithm 2a and its incorrect variation are then more similar, 6.66 vs 6.65.

Next, I use the diffuse initialization of τ, while keeping the ergodic distribution for ς. The diffuse initialization of τ changes the numerical results so little that they are again the same as those in Table 1 up to the reported precision. Let me stress the finding that the results obtained with Algorithm 2a continue to differ from the results obtained with the incorrect variation, and hence the misunderstanding matters in model (4a-4c) even with the diffuse initialization of τ. This is because when this model is cast in form (1a-1b) the constant term of equation (4b) is a state with a non-zero and non-diffuse initialization and the failure to reset a_1 to 0 distorts the simulation smoother. Equivalently, when this model is cast in form (2a-2b) all the states are zero-mean or diffuse, but the failure to reset c_t to zero distorts the simulation smoother and yields the same numerical results.

Table 1: Trend GNP in Watson's model based on simulation smoothers. Mean, standard deviation in parenthesis.

	Algorithm 2a	No resetting of a_1, c_t
τ_1	6.24 (0.02)	6.14 (0.02)
τ_{50}	6.66 (0.02)	6.65 (0.02)

5 Conclusion

This note discusses the implementation of the Durbin and Koopman algorithm for drawing the states conditionally on the observables in a state space model while pointing out a possible misunderstanding. The misunderstanding matters when the initial state vector is not all zero-mean or diffuse, or when a nonzero intercept is present, and leads to incorrect draws of the states, especially in the beginning of a sample. By clarifying the possible misunderstanding, this note will hopefully encourage an even wider use of the Durbin and Koopman algorithm by practitioners.

References

Durbin, J. and Koopman, S. J. (2002). A simple and efficient simulation smoother for state space time series analysis. *Biometrika*, 89(3):603–615.

Durbin, J. and Koopman, S. J. (2012). *Time Series Analysis by State Space Methods: Second Edition*. Oxford Statistical Science Series. OUP Oxford.

Watson, M. W. (1986). Univariate detrending methods with stochastic trends. *Journal of Monetary Economics*, 18(1):49–75.

Acknowledgements
I thank Gianni Amisano, Siem Jan Koopman and anonymous referees for their very useful comments. The views expressed in this paper are solely those of the author and do not necessarily reflect the views of the European Central Bank. The Matlab programs used for this note can be downloaded from www.runmycode.org.

Marek Jarocinski
European Central Bank, Frankfurt am Main, Germany;
e-mail: marek.jarocinski@ecb.int

Postal address	60640 Frankfurt am Main, Germany
Telephone	+49 69 1344 0
Website	www.ecb.europa.eu

This paper can be downloaded without charge from www.ecb.europa.eu, from the Social Science Research Network electronic library at http://ssrn.com or from RePEc: Research Papers in Economics at https://ideas.repec.org/s/ecb/ecbwps.html.

Information on all of the papers published in the ECB Working Paper Series can be found on the ECB's website, http://www.ecb.europa.eu/pub/scientific/wps/date/html/index.en.html.

ISSN	1725-2806 (online)
ISBN	978-92-899-1680-6
DOI	10.2866/516204
EU catalogue No	QB-AR-15-107-EN-N