1-1-2011

Initial values in estimation procedures for State Space Models (SSMs)

Raed Alzghool
Al-Balqa Applied University

Yan-Xia Lin
University of Wollongong, yanxia@uow.edu.au

Initial values in estimation procedures for State Space Models (SSMs)

Abstract

In this paper, we will focus on State Space Models(SSMs), especially the stochastic volatility model, and lookfor a standard approach for assigning initial values in theQuasi-Likelihood (QL) and Asymptotic Quasi-Likelhood (AQL)estimation procedures.

Keywords

Initial, values, estimation, procedures, for, State, Space, Models, SSMs

Disciplines

Physical Sciences and Mathematics

Publication Details

Alzghool, R. & Lin, Y. (2011). Initial values in estimation procedures for State Space Models (SSMs). Proceedings of the World Congress on Engineering 2011, WCE 2011 (pp. 307-313). Hong Kong: IAENG.

Proceedings of the World Congress on Engineering 2011 Vol I
WCE 2011, July 6 - 8, 2011, London, U.K.

Initial Values in Estimation Procedures for State Space Models (SSMs)

Raed Alzghool and Yan-Xia Lin

***Abstract*—In this paper, we will focus on State Space Models (SSMs), especially the stochastic volatility model, and look for a standard approach for assigning initial values in the Quasi-Likelihood (QL) and Asymptotic Quasi-Likelihood (AQL) estimation procedures.**

***Index Terms*—State Space Models (SSMs), Quasi-Likelihood (QL), Asymptotic Quasi-Likelhood (AQL), Kalman filter, Non-linear and/or Non-Gaussian SSMs.**

I. Introduction

THE class of state space models (SSM) provides a flexible framework for describing a wide range of time series in a variety of disciplines. For extensive discussion on SSM and their applications see Harvey [16] and Durbin and Koopman [13]. A state-space model can be written as

$$y_t = f_1(\alpha_t, \theta) + h_1(y_{t-1}, \theta)\epsilon_t, \quad t = 1, 2, \cdots, T \qquad (1)$$

where $y_1, \ldots, y_T$ represent the time series of observations; θ is an unknown parameter that needs to be estimated; $f_1(.)$ is a known function of state variable α_t and θ; and $\{\epsilon_t\}$ are uncorrelated disturbances with $E_{t-1}(\epsilon_t) = 0$, $Var_{t-1}(\epsilon_t) = \sigma^2_\epsilon$; in which E_{t-1}, and Var_{t-1} denote conditional mean and conditional variance associated with past information updated to time $t-1$ respectively. State variables $\alpha_1, \ldots, \alpha_T$ are unobserved and satisfy the following model

$$\alpha_t = f_2(\alpha_{t-1}, \theta) + h_2(\alpha_{t-1}, \theta)\eta_t, \quad t = 1, 2 \cdots, T, \qquad (2)$$

where $f_2(.)$ is a function of past state variables and θ; $\{\eta_t\}$ are uncorrelated disturbances with $E_{t-1}(\eta_t) = 0$, $Var_{t-1}(\eta_t) = \sigma^2_\eta$. $h_1(.)$ and $h_2(.)$ are unknown functions.

One special application that we will consider in detail is the Stochastic Volatility Model (SVM), a frequently used model for returns of financial assets. Applications, together with estimation for SVM, can be found in Jacquier, et al [22]; Briedt and Carriquiry [8]; Harvey and Streible [19]; Sandmann and Koopman [27]; Pitt and Shepard [25].

There are several approaches in the literature for estimating the parameters in SSMs by using the maximum likelihood method when the probability structure of underlying model is normal or conditional normal. Durbin and Koopman ([14], [13]) obtained accurate approximation of the log-likelihood for Non-Gaussian state space models by using Monte Carlo simulation. The log-likelihood function is maximised numerically to obtain estimates of unknown parameters. Kuk [23] suggested an alternative class of estimate models based on conjugate latent process and applied it to approximate the likelihood of a time series model for count data. To overcome the complex likelihoods of a time series model with count data, Chan and Ledolter [10] proposed the Monte Carlo EM algorithm that uses a Markov chain sampling technique in the calculation of the expectation in the the E-step of the EM algorithm. Davis and Rodriguez-Yam [12] proposed an alternative estimation procedure which is based on an approximation to the likelihood function. Alzghool and Lin [2] proposed quasi-likelihood (QL) approach for estimation of state space models without full knowledge on the probability structure of relevant state-space system. The QL method relaxes the distributional assumptions and only assumes the knowledge on the first two conditional moments of y_t and α_t associated past information. This weaker assumption makes the QL method widely applicable and become a popular method of estimation. A comprehensive review on the QL method is available in Heyde [21]. A limitation of the QL is that in practice, the conditional second moments of of y_t and α_t might not available. The AQL approach provides an alternative method of parameter estimation when unknown form of heteroscedasticity is presented.

The estimation procedure for SSMs consists of two parts. The first part is, given observations $\{y_1, \ldots, y_T\}$, to estimate state variables α_t. The second part is to combine the information of $\{y_t\}$ and $\{\hat{\alpha}_t\}$ to estimate unknown parameter θ in the model. The Kalman filter and the smoother methods are widely used to estimate an unobservable series, state variables, in SSMs (Anderson and Moore [7], Harvey [17]).

In summary, the QL and AQL estimating procedures discussed in Alzghool and Lin ([2],[3], [5]), Alzghool [4], and Alzghool, et al [6]. consist of the following steps:

(i) Assign initial values to α_0, θ_0 and $\mathbf{\Sigma}_0 = \mathbf{I}$.

(ii) Obtain the QL/AQL estimates $\hat{\alpha_t}$ of α_t for $t = 1, 2, \ldots, T$.

(iii) For the AQL estimating procedure, obtain $\hat{\Sigma}_{t,n}$ by using the kernel method.

(iv) Obtain the QL/AQL estimate $\hat{\theta}$ of θ.

(v) Steps (ii), (iii) and (iv) will be alternatively repeated until estimates converge.

The final estimation results for SSMs might be jointly affected by the initial values α_0 and θ_0 which initially assigned to the underlying model during the inference procedure.

In this paper, following two issues are investigated.

(1) How sensitive are the final estimates to the initial values assigned to the state variable α_0 and θ_0?

(2) If the estimation results are sensitive to the choice of the initial values, what should initial value of the state variable α_0 be and how is the final estimate of θ determined?

This paper is structured as follows. In Section II, the sensitivity of the QL and AQL estimation procedures to the initial values assigned to state variable α_0 is investigated

Manuscript received March 04, 2011; revised XXXX XX, 2011.

R. Alzghool is with the Department of Applied Science, Faculty of Prince Abdullah Ben Ghazi for Science and Information Technology, Al-Balqa' Applied University, Al-Salt, Jordan e-mail: raedalzghool@bau.edu.jo .

Y. Lin is with School of Mathematics and Applied Statistics, University of Wollongong, Wollongong, NSW 2500, Australia e-mail: yanxia@uow.edu.au .

via simulation studies. In Section III, a new suggestion for choosing the initialisation of the state variable α_0 is given. In Section IV, the impact of the starting values of system parameters θ_0 in the estimation results is investigated via simulation studies. In Section V, a standard procedure to improve the grid search procedure for obtaining a better estimation of θ is established. In Section VI applications of the QL and AQL methods to real data modelled by SSMs are given. In Section VII, a conclusion is provided.

II. Effect of Initialisation of α_0

The impact of the initial value of the state variable α_0 on the final inference result is illustrated via simulation studies in this section. Simulation study based on stochastic volatility model (SVM) is presented below.

A. Stochastic Volatility Models (SVM)

Consider the stochastic volatility model,

$$\ln(y_t^2) = \alpha_t + \ln {\xi_t}^2, \quad t = 1, 2, \cdots, T, \tag{3}$$

$$\alpha_t = \gamma + \phi\alpha_{t-1} + \eta_t, \quad t = 1, 2, \cdots, T, \tag{4}$$

where both ξ_t and η_t are i.i.d. r.v.'s; η_t has mean 0 and variance σ_η^2.

In order to show how the initial value α_0 effects the final estimation in the SVM when the QL and AQL approaches are applied, we carried out a simulation study on SVM Model defined by (3) and (4). The simulation was conducted as follows. First, 1,000 independent samples of size 500 are generated from (3) and (4) based on a true parameter $\theta = (\gamma, \phi)$, where $\eta_t \sim N(0, \sigma_\eta^2)$, $\xi_t \sim N(0, 1)$, and the initial value for α_0 in the true model is $\alpha_0 = 0$. Once $\{y_t\}$ and $\{\alpha_t\}$ are generated, pretend that $\{\alpha_t\}$ is unobserved and γ, and ϕ are unknown. Then apply the QL and AQL estimation procedures to $\{y_t\}$ only to obtain the estimate of α_t, γ, and ϕ. Different parameter settings for $(\gamma, \phi, \sigma_\eta^2)$ are considered in the simulation. The mean and root mean squared errors for $\hat{\gamma}$ and $\hat{\phi}$ based on 1,000 independent samples are calculated.

Let $\hat{\alpha}_0$ be the initial state used in the inference procedure. In Table I, different values of $\hat{\alpha}_0$, mean and root mean squared errors for $\hat{\gamma}$, and $\hat{\phi}$ given by the QL and AQL methods are reported.

We can see from Table I that the RMSE of QL and AQL estimates are increased when $\hat{\alpha}_0$ is chosen farther from the true value α_0. Since the increase in the RMSE for QL is less than for AQL, this indicates that the QL approach is less sensitive to the initial value of state variable than the AQL approach.

TABLE I
QL and AQL estimates, based on 1,000 replications. The root mean square error of each estimate is reported below that estimate, based on different initial values for α_0 (T = 500).

		$\sigma_\eta = 0.675$		$\sigma_\eta = 0.260$		$\sigma_\eta = 0.061$	
	α_0	γ	ϕ	γ	ϕ	γ	ϕ
true	0	-0.821	0.90	-0.368	0.95	-0.141	0.98
$\hat{\alpha}_0$=1	AQL	-0.873	0.915	-0.411	0.924	-0.349	0.954
		0.138	0.020	0.234	0.047	0.255	0.036
	QL	-0.843	0.931	-0.431	0.927	-0.228	0.964
		0.141	0.033	0.098	0.025	0.091	0.017
$\hat{\alpha}_0$=2	AQL	-0.860	0.916	-0.328	0.934	-0.230	0.970
		0.136	0.022	0.210	0.046	0.157	0.021
	QL	-0.893	0.927	-0.482	0.920	-0.250	0.970
		0.159	0.029	0.134	0.032	0.120	0.022
$\hat{\alpha}_0$=3	AQL	-0.817	0.916	-0.255	0.933	-0.157	0.982
		0.169	0.032	0.307	0.076	0.112	0.021
	QL	-0.935	0.923	-0.527	0.913	-0.286	0.954
		0.179	0.026	0.175	0.039	0.149	0.027
$\hat{\alpha}_0$=4	AQL	-0.770	0.912	-0.144	0.921	-0.089	0.982
		0.240	0.045	0.442	0.084	0.237	0.042
	QL	-0.965	0.921	-0.574	0.905	-0.318	0.949
		0.198	0.024	0.219	0.049	0.178	0.032

III. Determination of $\hat{\alpha}_0$

Consider the univariate time series y_t satisfying

$$y_t = \alpha_t + \epsilon_t, \quad t = 1, 2, \cdots, T \tag{5}$$

$$\alpha_t = \alpha_{t-1} + \eta_t, \quad t = 1, 2, \cdots, T \tag{6}$$

where $\epsilon_t \sim N(0, \sigma_\epsilon^2)$, $\eta_t \sim N(0, \sigma_\eta^2)$, and $\alpha_0 \sim N(a_0, P_0)$. $\{\epsilon_t\}$ and $\{\eta_t\}$ are two independent Gaussian white noise series. The initial value α_0 is independent of $\{\epsilon_t\}$ and $\{\eta_t\}$ for $t > 0$. In literature, α_t is referred to as the *trend* of the series, which is not directly observable, and y_t is observable. The model is called a *local level model* in Durbin and Koopman ([13], Chapter 2), which is a simple case of the *structural time series model* of Harvey [17].

When nothing is known about the initial value α_0, the initialisation of α_0 is usually given by a diffuse prior approach that fixes a_0 at an arbitrary value and let $P_0 \to \infty$ (Zivot *et al.* [30], Durbin and Koopman [13], Harvey [16]). However, some researchers consider that the diffuse approach is not realistic because they regard that the assumption of infinite variance is unnatural, given that all observed time series have finite values. From this point of view an alternative approach is suggested, which assumes that α_0 is an unknown constant and needs to be estimated from the data. In Harvey [18], it is suggested that the initial value of α_0 can be taken as y_1. This is the same value as that obtained by assuming that α_0 is diffuse. More details about the intitialisation of the Kalman filter under the normality assumption for SSM are provided in Durbin and Koopman ([13], Chapter 5 and references therein). Several other suggestions on initialisation for the state variable in SSM under normality assumption are given in a recent survey by Casals and Sotoca [9]. They derived an exact expression for the conditional mean and variance of the initial state of SSM.

In this paper, we follow the QL method to derive a simple method for determining $\hat{\alpha}_0$ without assigning any probability distribution to α_0.

Consider the following state-space model:

$$y_t = f(\alpha_t, \theta) + \epsilon_t, \quad t = 1, 2, \cdots, T, \tag{7}$$

$$\alpha_t = g(\alpha_{t-1}, \theta) + \eta_t, \quad t = 1, 2 \cdots, T. \tag{8}$$

For $t = 1$, we have

$$y_1 = f(\alpha_1, \theta) + \epsilon_1, \tag{9}$$

$$\alpha_1 = g(\alpha_0, \theta) + \eta_1. \tag{10}$$

ISBN: 978-988-18210-6-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

In models (9), and (10), α_1, α_0, ϵ_1, and η_1 are unobserved. Assume θ is known or determined by empirical knowledge.

The rule used to determine $\hat{\alpha}_0$ should meet the condition that given observation y_1, $\hat{\alpha}_0$ is able to ensure that $f(\hat{\alpha}_1, \theta)$ is an optimal estimation of $E(y_1)$.

From (9), consider

$$\epsilon_1 = y_1 - f_1(\alpha_1, \theta)$$

Let α_1 be an unknown parameter and consider estimating function space

$$\mathcal{G}_T^{(1)} = \{a_1(y_1 - f_1(\alpha_1, \theta)) \ \mid \ a_1 \in R\}.$$

A standardised optimal estimating function in $\mathcal{G}_T^{(1)}$ is

$$G^*_{(1)}(\alpha_1) = -E(\frac{\partial f}{\partial \alpha_1})[Var(\epsilon_1)]^{-1}(y_1 - f(\alpha_1, \theta)).$$

If $E(\frac{\partial f}{\partial \alpha_1}) \neq 0$, and f^{-1} exists, the optimal estimator of α_1 will be given by $G^*_{(1)}(\alpha_1) = 0$, that is,

$$\hat{\alpha}_1 = f^{-1}(y_1, \theta). \quad (11)$$

Using (10), consider

$$\eta_1 = \alpha_1 - g(\alpha_0, \theta).$$

Let α_0 be an unknown parameter and consider estimating function space

$$\mathcal{G}_T^{(0)} = \{a_0(\alpha_1 - g(\alpha_0, \theta)) \ \mid a_0 \in R\}.$$

A standardised optimal estimating function in $\mathcal{G}_T^{(0)}$ is

$$G^*_{(0)}(\alpha_0) = -E(\frac{\partial g}{\partial \alpha_0})[Var(\eta_1)]^{-1}(\alpha_1 - f(\alpha_0, \theta)).$$

If $E(\frac{\partial g}{\partial \alpha_0}) \neq 0$, and g^{-1} exists, the optimal estimator of α_0 will given by $G^*_{(0)}(\alpha_0) = 0$, that is,

$$\hat{\alpha}_0 = g^{-1}(\alpha_1, \theta). \quad (12)$$

Therefore, we make the following suggestion for determining the initial state $\hat{\alpha}_0$ in inference process.

Suggestion: *For a SSM*

$$y_t = f(\alpha_t, \theta) + \epsilon_t, \quad t = 1, 2, \cdots, T$$

$$\alpha_t = g(\alpha_{t-1}, \theta) + \eta_t, \quad t = 1, 2 \cdots, T.$$

If $E(\frac{\partial f}{\partial \alpha_1}) \neq 0$, $E(\frac{\partial g}{\partial \alpha_0}) \neq 0$, f^{-1} and g^{-1} exist, the optimal decision on $\hat{\alpha}_0$ is

$$\hat{\alpha}_0 = g^{-1}(f^{-1}(y_1)). \quad (13)$$

For convenience, denote this $\hat{\alpha}_0$ as $\hat{\alpha}_0^*$.

As an example for (5) and (6), the optimal value for $\hat{\alpha}_0$ is y_1, which is the same as the one given under diffuse conditions.

In the following, we apply the Suggestion to stochastic volatility model, and use simulation to investigate whether the Suggestion is practicable or not.

TABLE II
QL AND AQL ESTIMATES BASED ON 1,000 REPLICATION. THE ROOT MEAN SQUARE ERROR OF EACH ESTIMATE IS REPORTED BELOW THAT ESTIMATE. $\hat{\alpha}_0^*$ IS DIFFERENT FROM SAMPLE TO SAMPLE. (T = 500).

		$\sigma_\eta = 0.675$		$\sigma_\eta = 0.260$		$\sigma_\eta = 0.061$	
	α_0	γ	ϕ	γ	ϕ	γ	ϕ
true	0	-0.821	0.90	-0.368	0.95	-0.141	0.98
α_0=0	AQL	-0.878	0.92	-0.499	0.91	-0.437	0.94
		0.136	0.019	0.229	0.049	0.354	0.052
	QL	-0.788	0.94	-0.391	0.94	-0.198	0.97
		0.140	0.037	0.071	0.019	0.063	0.013
α_0=$\hat{\alpha}_0^*$	AQL	-0.857	0.92	-0.499	0.91	-0.440	0.94
		0.163	0.024	0.243	0.051	0.402	0.060
	QL	-0.830	0.93	-0.378	0.94	-0.194	0.97
		0.142	0.034	0.082	.019	0.071	.014

A. Stochastic Volatility Model

Consider stochastic volatility process defined by (3) and (4), i.e.

$$\ln(y_t^2) = \alpha_t + \ln {\xi_t}^2, \quad t = 1, 2, \cdots, T.$$

$$\alpha_t = \gamma + \phi\alpha_{t-1} + \eta_t, \quad t = 1, 2, \cdots, T,$$

where both ξ_t and η_t are i.i.d. r.v.'s; η_t has mean 0 and variance σ_η^2, $\phi \neq 0$.

Let

$$\epsilon_1 = \ln {\xi_1}^2 - E(\ln {\xi_1}^2).$$

Using (3) and (4), it follows that

$$\begin{aligned} \epsilon_1 &= \ln(y_1^2) - \alpha_1 - E(\ln {\xi_1}^2) \\ &= \ln(y_1^2) - f(\alpha_1, \theta), \end{aligned}$$

and

$$\begin{aligned} \eta_1 &= \alpha_1 - (\gamma + \phi\alpha_0) \\ &= \alpha_1 - g(\alpha_0, \theta), \end{aligned}$$

where $\theta = (\gamma, \phi)'$, $f(\alpha_1, \theta) = \alpha_1 + E(\ln {\xi_1}^2)$, and $g(\alpha_0, \theta) = \gamma + \phi\alpha_0$.

Since $E(\frac{\partial f}{\partial \alpha_1}) = 1 \neq 0$, $E(\frac{\partial g}{\partial \alpha_0}) = \phi \neq 0$, and f^{-1}, g^{-1} exist, therefore,

$$\hat{\alpha}_0^* = g^{-1}(f^{-1}(y_1)) = \frac{\ln(y_1^2) - E(\ln {\xi_1}^2) - \gamma}{\phi}. \quad (14)$$

If ξ_t has standard normal distribution, then $E(\ln \xi_t^2) = -1.2704$ and $Var(\ln \xi_t^2) = \pi^2/2$ (see Abramowitz and Stegun [1], p. 943). Then, substituting in (14)

$$\hat{\alpha}_0^* = g^{-1}(f^{-1}(y_1)) = \frac{\ln(y_1^2) + 1.2704 - \gamma}{\phi}. \quad (15)$$

To show how the optimal initial value $\hat{\alpha}_0^*$ effects the final estimation when the QL and AQL approaches are applied, we carried out a simulation study on SVM model defined by (3) and (4). We camper the estimation of (, ϕ) given by the true α_0 and $\hat{\alpha}_0^*$. Results are presented by Table II.

Table II shows that, compared to results in Table I, the estimation given by $\hat{\alpha}_0^*$ are close related to those given by the true $\alpha_0 = 0$.

ISBN: 978-988-18210-6-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

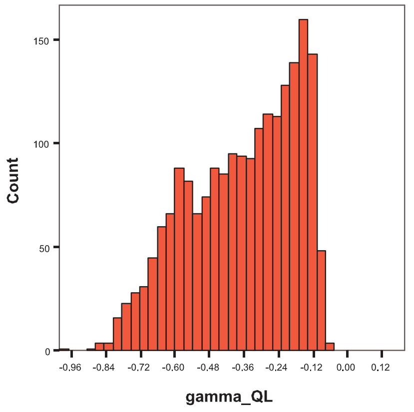

Fig. 1. Histogram of QL estimation of γ in SVM, based on 2,000 different starting values.

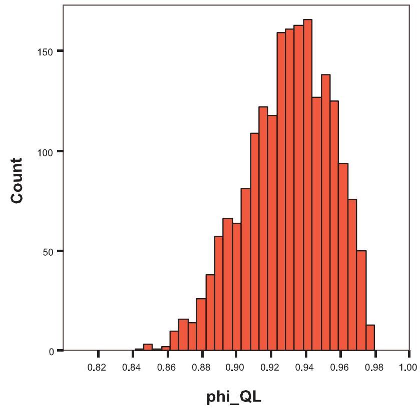

Fig. 2. Histogram of QL estimation of ϕ in SVM, based on 2,000 different starting values.

IV. The Starting Values for System Parameter θ_0

In this section, we consider the starting value for system parameters θ_0. As described in literature, the outputs of non-linear inference procedures rely strongly on the appropriate value of the initial parameter θ_0. It is usually suggested that θ_0 should be chosen from a close neighbourhood of its true value (Zivot *et al*. [30]). Since the true value of θ_0 is unknown, it is an issue how to identify the close neighbourhood of θ_0.

The impact of the starting values of system parameters θ_0 is illustrated via simulation studies below.

A. Stochastic Volatility Models

Consider SVM as given in (3) and (4) where $\eta_t \sim N(0, 0.675^2)$, $\xi_t \sim N(0,1)$, and the initial value for α_0 in the true model is given by $\alpha_0 = 0$. In this example, the state space model is involved with the parameter $\theta = (\gamma, \phi)$. Let $\theta = (-0.368, 0.95)$, a sequence of observations $y_1, \cdots, y_{1000}$ from the state space model were generated. Then we pretend θ is unknown. Consider a two-dimensional range (-0.868,0.132; 0.80,0.99) for $\theta = (\gamma, \phi)$, which covers the true parameter (-0.368,0.95). Then we apply a two-dimensional grid search to (-0.868,0.132; 0.80,0.99) with increasment of 0.01. For each starting value of θ from the grid area, we apply the QL and AQL estimating procedures to the realisation $y_1, \cdots, y_{1000}$ and obtain the QL and AQL estimation of θ where $\hat{\alpha}_0 = \alpha_0$ are used. In Figure 1 - 4, we show the histograms of QL and AQL estimation of γ and ϕ based on 2000 different starting values.

Like others estimation procedures described in literature, the QL and AQL estimations of θ rely strongly on the value of the initial parameter θ_0.

We note an interesting phenomenon in the histograms illustrated in Figures 1 - 4. The true value of a parameter is not always allocated in the low frequency area. Obviously, the size of the low frequency area relies on the nature of the true model. This suggests that, although it is not appropriate to quantitatively identify an optimal estimation on system parameters utilising the information provided by a histogram diagram indirectly through the grid search approach, it is possible to narrow down and obtain a potential range covering the true value of parameters in underlying model by using the information provided by the histogram diagrams.

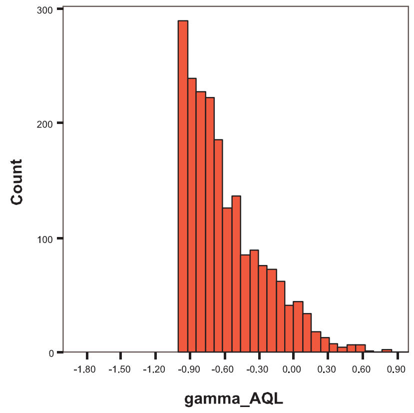

Fig. 3. Histogram of AQL estimation of γ in SVM, based on 2,000 different starting values.

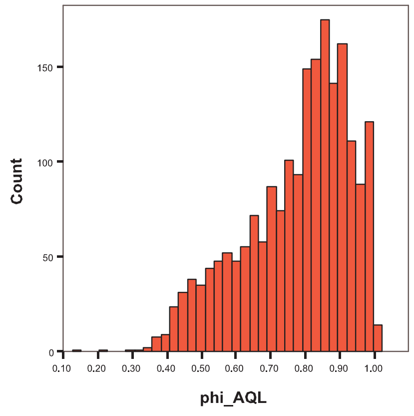

Fig. 4. Histogram of AQL estimations of ϕ in SVM, based on 2,000 different starting values.

ISBN: 978-988-18210-6-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

V. Determination of the Estimation of the System Parameter θ

In their survey article, Zivot *et al*. [30] suggested choosing a starting value θ_0 close to the true value of θ. The estimation of θ using a Monte Carlo approximation for count data given by Kuk [23] is only good when the initial value of θ is assigned around the true value of θ. Other approaches to decide θ_0 are also suggested in literatuer. For example, Durbin and Koopman [14] numerically maximised the approximate likelihood for non-Gaussian SSMs to obtain the starting value for θ_0; Sandmann and Koopman [27] used a two-dimensional grid search procedure which searches for an appropriate starting value for θ_0 across the surface of a Gaussian log-likelihood function; Geweke and Tanizaki [15] and Tanizaki and Mariano ([28], [29]) used a simple grid search for θ_0 where the expected log-likelihood function is maximised.

The ML method is a popular method for estimating the parameters of SSMs. The ML method works if the probability structure of the underlying state space system is known. In practice, it is not realistic to assume that the system's probability structure is known. Then, the maximum likelihood method becomes impracticable. Therefore, searching θ_0 based on maximising the log-likelihood function cannot be applied. Without knowledge of the log-likelihood, a distribution free procedure can be considered. It is implemented by a grid search over a feasible region of the parameter space, and the parameter estimation will be the one giving the minimum residual sum of squares (RSS)(see Coakley *et al*. [11] and Naik-nimbalkar and Rajarshi [24]).

In this paper, we adapt grid search procedure but with some improvements. It is sensible to obtain the estimate of θ by utilising a the grid search, and the residual sum of squares. However, if the grid search area is relatively large, the smallest sum of residuals might not lead to the best estimation of θ. One example can be fond from the simulation study discussed below. To improve the outcomes of the grid search procedure and sum of residuals, we need to reduce the area of the grid search into a reasonable size.

We suggest the following steps in determining the estimation of θ for SSMs: (in the following, we used a two-dimensional parameter as an example.)

Step 1. First determine a reasonable range. Based on experience, this range should cover the true parameter θ. For example, for PM and SVM, decide a two-dimensional area [a,b; c,d], covering the true parameter θ.

Step 2. Following the two-dimensional grid search procedure, we assign θ_0 with a different starting value, and obtain the QL or AQL estimation of the parameter.

Step 3. Draw the histogram of the QL or AQL estimates obtained from step 2.

Step 4. Consider the region with the highest frequency estimation values in the histogram as a potential region to cover the true value of the parameter. Obviously this potential region tends to be smaller than the range in Step 1.

Step 5. Let $\hat{y}_t(\hat{\theta})$ be the predicted value of y_t based on the observation equation. Find $\hat{\theta}$, which minimises $RSS_y(\theta) = \sum_{t=1}^{T}(y_t - y_t(\hat{\theta}))^2$ in the potential region.

The above steps used to determine the estimate of θ for SSMs are illustrated by the following example.

TABLE III
QL, AQL, QL^*, and AQL^* estimates, and RSS_y are reported below each estimate.

		SVM $\sigma_\eta = 0.675$ γ ϕ
true	0	-0.363 0.95
	AQL	-0.30 0.95
	RSS_y	323.53
	QL	-0.45 0.93
	RSS_y	660.62
	AQL^*	-0.31 0.94
	RSS_y	457.65
	QL^*	-0.32 0.95
	RSS_y	725.03

Example : Consider SVM as given in (**??**) and (**??**), where $\eta_t \sim N(0, 0.675^2)$, $\xi_t \sim N(0,1)$, and the initial value for α_0 in the true model is given by $\alpha_0 = 0$. In this example, the state space model is involved with the parameter $\theta = (\gamma, \phi)$. Let $\theta = (-0.368, 0.95)$, a sequence of observations $y_1, \cdots, y_{1000}$ and $\alpha_1, \cdots, \alpha_{1000}$ from the SVM were generated. Then we pretend $\{\alpha_t\}$ and θ are unknown.

Step 1. Consider a two-dimensional range (-0.868,0.122; 0.80,0.99) for $\theta = (\gamma, \phi)$, which covers the true parameter (-0.368,0.95).

Step 2. Apply a two-dimensional grid search to (-0.868,0.122; 0.80,0.99) with increases of 0.01. For each starting value of θ from the grid area, we apply the QL/AQL estimating procedures and obtain the QL/AQL estimate of θ.

Step 3. In Figures 1-4, we show the histograms of the QL and AQL estimates of γ and ϕ, based on 2,000 different starting values.

Step 4. From the histograms of the QL estimates of γ and ϕ given in Figures 1 and 2, the potential region for parameter (γ, ϕ) is chosen as [-0.36,-0.12; 0.91,0.95]. By using the histogram of the AQL estimates of and ϕ given in Figures 3 and 4, the potential region for parameter (γ, ϕ) is chosen as [-1.0,-0.30; 0.80,0.95].

Step 5. Find the estimate of γ and ϕ by minimising the residual sum of squares ($RSS_y(\theta)$) in the potential region and give the QL estimate of θ (-0.32,0.95), and the AQL estimate of θ (-0.31,0.94).

In Table III, the QL and AQL denote the estimation of θ, which gives the smallest RSS_y based on the region given in Step1, and the QL^* and AQL^* denote the estimates of θ, which gives the smallest RSS_y based on the potential region determined in Step 4. We can see from Table III, that the estimate of θ has improved in all cases after using the potential region determined by the information provided by histogram diagram. The above examples indicate that using the potential region is able to significantly improve the performance of RSS_y.

VI. Real data application

In this section, we consider log returns of Pound/Dollar exchange rates. The data are the daily observation of weekdays' closing pound to dollar exchange rates x_t from 1/10/81 to 28/6/85 and have been taken from the site:www//staff.feweb.vu.nl/ koopman/sv/. This data set has

ISBN: 978-988-18210-6-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

been studied and analysed by Harvey *et al.* [20], Davis and Rodriguez-Yam [12]; Rodriguez-Yam [26]; Durbin and Koopman [13] and Alzghool and Lin [2].

Let $y_t = log(x_t/x_{t-1}), t = 1, 2, \cdots, 945$. To model y_t, we adopt the same SVM used by Davis and Rodriguez-Yam [12].

$$y_t = \sigma_t \xi_t = e^{\alpha_t/2} \xi_t, \quad t = 1, 2, \cdots, 945, \tag{16}$$

$$\alpha_t = \gamma + \phi \alpha_{t-1} + \eta_t, \quad t = 1, 2, \cdots, T, \tag{17}$$

where both ξ_t and η_t are i.i.d. r.v.'s; η_t has mean 0 and variance σ_η^2. Therefore,

$$\ln(y_t^2) = \alpha_t + \ln {\xi_t}^2, \quad t = 1, 2, \cdots, T. \tag{18}$$

If ξ_t were standard normal, then $E(\ln \xi_t^2) = -1.2704$ and $Var(\ln \xi_t^2) = \pi^2/2$ (see Abramowitz and Stegun [1], p. 943). Let $\epsilon_t = \ln \xi_t^2 + 1.2704$, and $\delta_t = (\epsilon_t, \eta_t)'$.

We apply the QL method to the data model under the assumption that the conditional covariance matrix is known as follows:

$$Var_{t-1}(\delta_t) = \mathbf{\Sigma}_t = \begin{pmatrix} \frac{\pi^2}{2} & 0 \\ 0 & \sigma_\eta^2 \end{pmatrix}.$$

The AQL method is applied to the data by assuming no knowledge of the conditional covariance matrix. In the QL approach, σ_η will estimate from the residuals, but in AQL approach it is estimated by the Kernel estimator.

Following steps are for obtaining the estimate of $\theta = (\phi, \gamma)$ for the Pound/ Dollar exchange rate data:

Step 1. Decide a grid search area, based on previous studies: (-0.813,0.177; 0.80,0.99).

Step 2. Apply a two-dimensional grid search to (-0.813,0.177; 0.80,0.99) with increases of 0.01. For each starting value of θ from the grid area, we apply the QL/AQL estimating procedures and obtain the QL/AQL estimate of θ.

Step 3. In Figures 5-8, we show the histograms of the QL and AQL estimates of γ and ϕ, based on 2,000 different starting values.

Step 4. From the histograms of the QL estimates of γ and ϕ given in Figures 5 and 6, the potential region for parameter (γ, ϕ) is chosen as (-0.17,-0.04; 0.86,0.95). By using the histogram of the AQL estimates of γ and ϕ given in Figures 7 and 8, the potential region for parameter (γ, ϕ) is chosen as (-0.45,0.1; 0.825,0.99).

Step 5. Find the estimate of γ and ϕ by minimising the residual sum of squares $(RSS_y(\theta))$ in the potential region and the QL estimate of θ is (-0.048,0.949), and the AQL estimate of θ is (-0.082,0.971).

Table IV shows estimations of $\theta = (\phi, \gamma)$ obtained by different methods. AQL denotes the asymptotic quasi-likelihood estimate, QL the estimate obtained by quasi-likelihood approach, AL the estimate obtained by maximising the approximate likelihood proposed by Davis and Rodriguez-Yam [12] and MCL the estimate obtained by maximising the estimate of the likelihood proposed by Durbin and Koopman [14]. AL and MCL outputs are taken from Rodriguez-Yam [26].

In Table IV, the estimate of γ and ϕ by QL, AL and MCL are close to each other. These three methods are carried out under the same assumption where ξ_t and η_t are independent. This might indicate that the performance of QL, AL and MCL will be similar. However, the AQL

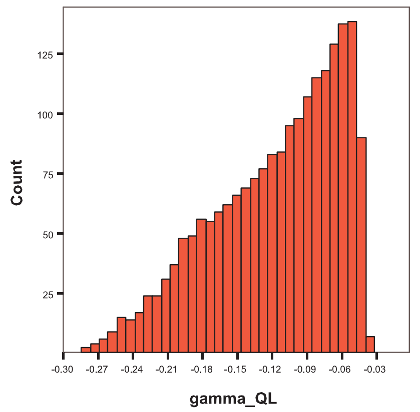

Fig. 5. Histogram of QL estimates of γ in SVM, based on 2,000 different starting values.

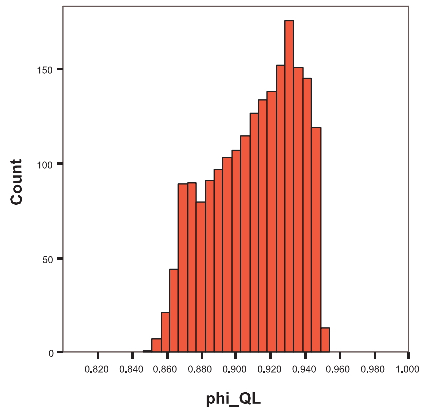

Fig. 6. Histogram of QL estimates of ϕ in SVM, based on 2,000 different starting values.

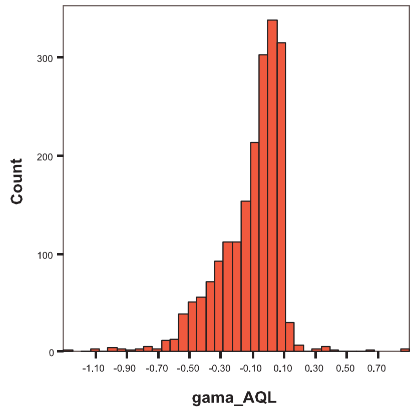

Fig. 7. Histogram of AQL estimates of γ in SVM, based on 2,000 different starting values.

estimates are slightly different from those of QL, as well as the estimates of AL and MCL.

The estimates of AQL and QL are obtained based on different model settings. The main difference between their models is that one assumes that $cov(\eta_t, \xi_t) = 0$ and the other

ISBN: 978-988-18210-6-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

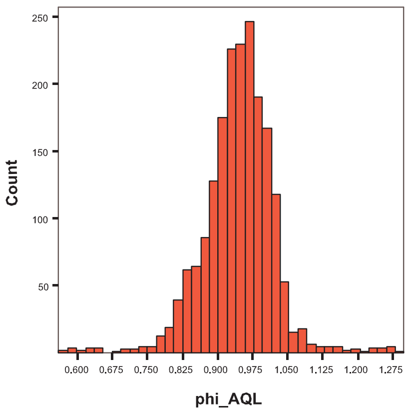

Fig. 8. Histogram of AQL estimates of ϕ in SVM, based on 2,000 different starting values.

TABLE IV
ESTIMATES OF γ, ϕ AND σ_η^2 FOR POUND/DOLLAR EXCHANGE RATE DATA.

	$\hat{\gamma}$	$\hat{\phi}$	$\hat{\sigma}_\eta^2$
AQL	-0.082	0.971	0.239
QL	-0.048	0.949	0.025
AL	-0.023	0.957	0.026
MCL	-0.023	0.975	0.027

does not. To understand which model setting is appropriate, it requires checking whether we can accept $cov(\eta_t, \xi_t) = 0$. We consider $\hat{\epsilon}_t$ and $\hat{\eta}_t$ given by QL and find that $\hat{\epsilon}_t$ and $\hat{\eta}_t$ are highly correlated with $r = 0.91$ and significant at level 0.01. So, the assumption of ϵ_t and η_t uncorrelated is not valid. Therefore, it is not appropriate to apply the QL method to the data. Thus, we rather accept the estimations given by the AQL method than those given by the QL method.

VII. CONCLUSION

In this paper, we investigated the sensitivity of the QL and AQL estimation procedures to initial values assigned to state variable α_0 and θ_0 via simulation studies. A suggestion on choosing the initial value of state variable α_0, without knowing the system's probability structure has been given. Simulation studies indicate that it is relatively reliable to follow the suggestion in determining the initialisation of the state variable α_0 during inference procedure. Apart from the impact of α_0, the QL and AQL estimates of θ also sensitive to the value of the starting parameter θ_0. In literature, it always suggestes that θ_0 has to be chosen from a neighbourhood close to the true value of θ. But, it dose not mention how to determine the close neighbourhood given the location of the true θ is unknown. In this paper, we established a standard procedure for determing the "close neighbourhood" and the estimation of θ in terms of minimizing RSS_y.

REFERENCES

[1] M. Abramovitz, and N. Stegun, *Handbook of Mathematical Functions*, Dover Publication, New York, 1970.

[2] R. Alzghool and Y.-X. Lin, Asymptotic Quasilikelihood Based on Kernel Smoothing for Nonlinear and Non-Gaussian State-Space Models, ICCSDE, London, UK, pp. 926-932, WCE 2007.

[3] R. Alzghool and Y.-X. Lin, Parameters Estimation for SSMs: QL and AQL Approaches, *IAENG International Journal of Applied Mathematics*, **38**, pp. 34-43, 2008.

[4] R. Alzghool, Estimation for state space models: quasi-likelihood and asymptotic quasi-likelihood approaches, PhD thesis, School of Mathematics and Applied Statistics, University of Wollongong, Australia, 2008.

[5] R. Alzghool and Y.-X. Lin, Estimation for State-Space Models: Quasi-likelihood, *Proceedings of the Tenth Islamic Countries Conference on Statistical Sciences (ICCS-X)*, **Volume I**, The Islamic Countries Society of Statistical Sciences, Lahore: Pakistan, pp. 409-423, 2010.

[6] R. Alzghool and Y.-X. Lin, and S. X., Chen, Asymptotic Quasi-likelihood Based on Kernel Smoothing for Multivariate Heteroskedastic Models with Correlation, *American Journal of Mathematical and Management Sciences*,(accepted), 2010.

[7] B. D. O. Anderson and J. B. Moore. *Optimal Filtering*, Prentice-Hall, New York, 1979.

[8] F.J. Breidt, and A.L. Carriquiry, Improved quasi-maximum likelihood estimation for stochastic volatility models. In: Zellner, A. and Lee, J.S. (Eds.), *Modelling and Prediction: Honouring Seymour Geisser*, Springer, New York, 228-247, 1996.

[9] J. Casals and S. Sotoca, Exact initial condition for maximum likelihood estimation of state space models with stochastic inputs, *Economics Letters*, **57**, 261-267, 1997.

[10] K. S. Chan and J. Ledolter, Monte Carlo EM estimation for time series models involving counts, *J. Amer. Statist. Assoc.*, **90**, 242-252, 1995.

[11] J. Coakley, A. Fuertes, and Perez, Numerical issues in threshold autoregressive modeling of time series, *Journal of Economic Dynamics Control*, **27**, 2219-2242, 2003.

[12] R. A. Davis, and G. Rodriguez-Yam, Estimation for State-Space Models: an approximate likelihood approach, *Statistica Sinica*, **15**, 381-406, 2005.

[13] J. Durbin, and S. J. Koopman, *Time Series Analysis by State Space Methods*, Oxford, New York, 2001.

[14] J. Durbin, and S. J. Koopman, Monte Carlo maximum likelihood estimation for non-gaussian state space models, *Biometrika*, **84**, 669-684, 1997.

[15] J. Geweke and H. Tanizaki, On Markov Chain Monte Carlo methods for nonlinear and non-Gassian state-space models, *Comm. Statist. Simulation Comput.*, **28**, 867-894, 1999.

[16] A. C. Harvey, *Forecasting, Structural Time Series Moodels and the Kalman Filter*, Cambridge University Press, 1989.

[17] A. C. Harvey, *Forecasting, Structural Time Series Moodels and the Kalman Filter*, Cambridge University Press, 1994.

[18] A. C. Harvey, Forecasting with unobserved components time series models, *volume 1 of Handbook of Economic Forecasting*, chapter 7, pages 328-408. ELSEVIER B. V., 2006.

[19] A. C. Harvey, and M. Streible, Testing for a slowly changing level with special reference to stochastic volatility, *J. Econometrics*, **87**, 167-189, 1998.

[20] A. Harvey, E. Ruiz, and N. Shephard, Multivariate stochastic variance models, *Review of Economic Studies*, **61**, 247-264, 1994.

[21] C. Heyde, *Quasi-likelihood and its Application: a general Approach to Optimal Parameter Estimation*, Springer, New York, 1997.

[22] E. Jacquire, and N. G. Polson, and P. E. Rossi, Bayesian analysis of stochastic volatility models (with discussion), *J. Bus. Econom. Statist*, **12**, 371-417, 1994.

[23] A. Y. Kuk, The use of approximating models in Monte Carlo maximum likelihood estimation, *Statist. Probab. Letter*, **45**, 325-333, 1999.

[24] U. V. Naik-nimbalkar and M. B. Rajarshi, Filtering and smoothing via estimating functions, *Journal of American statistical Association*, **90,(429)**,301-306, 1995.

[25] M. K. Pitt, and N. Shepard, Filtering via simulation: auxiliary particle filters, *J.Amer. Statist. Assoc.*, **94**, 590-599, 1999.

[26] G. Rodriguez-Yam, *Estimation for State-Space Models and Baysian regression analysis with parameter constraints*, Ph.D. Thesis, Colorado State University, 2003.

[27] G. Sandmann, and S. J. Koopman, Estimation of stochastic volatility models via Monte Carlo maximum likelihood, *J. Econometrics*, **87**, 271-301. 1998.

[28] H. Tanizaki and R. S. Mariano, Prediction, fltering and smoothing in nonlinear and nonnormal cases using Monte-Carlo integration, *J. of App. Econo- metrics*, **9(2)**, 163-179, 1994.

[29] H. Tanizaki and R. S. Mariano, Nonlinear and non-Gaussian state-space modeling with Monte-Carlo simulation, *J. of Econometrics*, **83**, 263-290, 1998.

[30] E. Zivot, J. Wang, and S. J. Koopman, State Space and Unobserved Com- ponent Models: Theory and Application, *chapter State space modelling in macroeconomics and finance using SsfPack*, 284-335, Cambridge University Press, 2004.

ISBN: 978-988-18210-6-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)