Derivation of Kalman Filter Estimates Using Bayesian Theory: Application in Time Varying Beta CAPM Model

Hamed Habibi¹, Reza Habibi² and Hamid Habibi³

Abstract

This paper is concerned with application of Kalman recursive estimates in the capital asset pricing (CAPM) model with time varying beta parameters. Following Kyriazis (2011), Kalman estimates are derived using a Bayesian probability theory. Rate of convergence and sensitivity analysis of estimates are derived. Through five examples, applications of presented estimates are shown. Extension to the non-normal cases and suggestion of Bayes filter is also considered. Comparisons with method of moment estimates are given.

Keywords: Kalman filter; Bayesian method; CAPM model

¹ PhD, Faculty of Science and Engineering, School of Civil and Mechanical Engineering, Curtin University, Perth, Australia. E-mail: hamed.habibi@postgrad.curtin.edu.au
² Iran Banking Institute, Central Bank of Iran, Tehran, Iran. E-mail: r_habibi@ibi.ac.ir
³ Hamid Habibi, Head of Department of Airworthiness, Civil Aviation Organization of Iran. Unfortunately, he passed away before publishing this paper.

Article Info: Received: January 29, 2016. Revised: March 4, 2016. Published online: June 10, 2016.
1 Introduction

There are many asset pricing models such as equilibrium model (CAPM), an econometric model (Fama and French three factors model) or an arbitrage pricing model (APT). In this paper the first model is considered. The CAPM, introduced by Sharpe (1963), implies that the systematic risk is only risk which noticed by investors because diversification cannot eliminate this risk. It states that the expected return of a risky security is decomposed to the sum of riskless rate ($r_f$) and a risk premium of market ($R^m_t - r_f$) which is multiplied by the asset's systematic risk measure beta ($\beta$). There are many extensions to CAPM, for example, the ICAPM (Intertemporal CAPM) or the consumption-based CAPM (CCAPM). For a comprehensive review about the CAPM model see Alexander (2001) and references therein.

In this paper, a CAPM model with time varying betas is considered as follows

$$\begin{cases}
R_t - r_f = \beta_t (R^m_t - r_f) + \varepsilon_t \\
\beta_t = a_t \beta_{t-1} + \xi_t
\end{cases}$$

and a Kalman filter is used to produces estimates mean and variance of $\beta_t$. In the above model, $R_t$ is the return of a risky asset at time $t$. Variables $\varepsilon_t$ and $\xi_t$ are supposed normally distributed with zero means and variances $\sigma^2_t$ and $\delta^2_t$. The Kalman filter is a mathematical power tool that is playing an increasingly important role in finance. It gives optimal recursive estimator of unknown parameters. Since it is in recursive form, new measurements can be processed once newcomer observations arrived. Kalman filter is increasingly used in financial applications. A comprehensive review about the application of Kalman filtering in financial models may be found in Harvey (1989). Nelson and Foster (1994) studied the estimation of ARCH time series using adaptive filtering. Racicot and Theoret (2007) studied the application of Kalman filter in hedge fund problems. Kyriazis (2011) proposed a simplified derivation of scalar Kalman filter using Bayesian setting. Habibi (2013) applied this method to derive adaptive filter
in a regression model in known and unknown variance of residual cases.

The paper is organized as follows. In the rest of Introduction, the Kalman filter equations are derived and the rate of convergence is studied. Sensitivity analysis are studied in Section 2. Some examples and applications are presented in Section 3. Comparisons with method of moment estimates are given in Section 4. Concluding remarks are given in Section 5.

1.1 Derivation

Following Kyriazis (2011), let the posterior distribution of $\beta_{t-1}$ be a normal distribution with $\mu_{t-1}$ and variance $\theta_{t-1}^2$. Thus, the prior density function of $\beta_t$ at time $t$ is proportional to

$$\exp\left\{-\frac{1}{2\tau_t^2}(\beta_t - \theta_t)^2\right\}$$

where

$$\theta_t = a_t\mu_{t-1} \quad \text{and} \quad \tau_t^2 = a_t^2\theta_{t-1}^2 + \delta_t^2.$$ 

Also, the likelihood function of $\beta_t$ given $R_t$ and $R_t^m$ is the probability density function of a normal distribution with mean $\gamma_t$ and variance $u_t^2$. The Kalman gain is $k_t$ where

$$\begin{align*}
\gamma_t &= \frac{R_t - r_f}{R_t^m - r_f} \\
u_t^2 &= \frac{\sigma_t^2}{(R_t^m - r_f)^2} \\
k_t &= \frac{\tau_t^2}{u_t^2 + \tau_t^2}
\end{align*}$$

Therefore, the posterior distribution of beta at state $t$ is normal with mean $\mu_t$ and variance $\theta_t^2$ as follows
The following Kalman equations are derived

\[
\begin{align*}
\mu_t &= a_t \mu_{t-1} + k_t (\gamma_t - a_t \mu_{t-1}) \\
\theta_t^2 &= (1 - k_t) \tau_t^2
\end{align*}
\]

**Remark 1.** Estimates \(\mu_t\) and \(\theta_t^2\) are the conditional mean and variance of \(\beta_t\) given \(R_t^m\). The marginal mean, variance and density function of \(\beta_t\) are given as follows

\[
\begin{align*}
E(\beta_t) &= E(E(\beta_t|R_t^m)) = E(\mu_t) \\
\text{var}(\beta_t) &= E(\text{var}(\beta_t|R_t^m)) + \text{var}(\mu_t) \\
f_{\beta_t}(\beta) &= E_{R_t^m}(f_{\beta_t|R_t^m}(\beta))
\end{align*}
\]

Using a Monte Carlo simulation the marginal density of \(\beta_t\) is computed.

**Remark 2.** Here, we investigate the condition for the CAPM holds. It can be seen that

\[
R_{t+1} = r_f + (a_{t+1} \beta_t + \xi_{t+1})(R_{t+1}^m - r_f) + \epsilon_{t+1}
\]

Then,

\[
E(R_{t+1}) = r_f + a_{t+1} E(\beta_t(R_{t+1}^m - r_f)).
\]

Also, \(E(R_{t+1}) = r_f + E(\beta_{t+1}(R_{t+1}^m - r_f))\). Thus, the condition is

\[
E(\beta_{t+1}(R_{t+1}^m - r_f)) = a_{t+1} E(\beta_t(R_{t+1}^m - r_f)).
\]

Moreover, it is seen that the conditional distribution of \(\beta_{t+1}\) given \(R_t^m\) is normal with mean and variance

\[
\begin{align*}
E(\beta_{t+1}|R_t^m) &= a_{t+1} \gamma_t \\
\text{var}(\beta_{t+1}|R_t^m) &= \frac{a_{t+1}^2}{(R_{t+1}^m - r_f)^2} \sigma_t^2 + \delta_t^2.
\end{align*}
\]
Remark 3. Cheng et al. (2003) considered a discrete time process which is a semi-martingale given as follows

\[ X_t = f_t + \sqrt{\theta_t} \epsilon_t, \]

They defined the adaptive filter as \( \rho f_{t-1} - (1 - \rho)(X_t - f_{t-1}) \). The coefficient \( \rho \) is derived by minimizing a MSE. Here, we want to apply this method. That is,

\[ \mu_t = \rho a_t \beta_{t-1} + (1 - \rho)(\beta_t - a_t \beta_{t-1}) \]

Define \( \varphi = a_t^2 \theta_{t-1}^2 (R_{t+1}^m - r_f)^2 \). By minimizing the variance \( \hat{\mu}_t \) with respect to \( \rho \), it is seen that

\[ \rho^* = \frac{\varphi + \sigma_t^2}{2 \varphi + \sigma_t^2} \]

It is known that \( \beta_t \) is close to the \( \gamma_t \) and \( \beta_{t-1} \) may be approximated by \( \mu_{t-1} \), then

\[ \mu_t = \rho^* a_t \mu_{t-1} + (1 - \rho^*)(\gamma_t - a_t \mu_{t-1}). \]

This is the adaptive filter version of above problem. However, this filter is dropped and hereafter properties of Kalman filter is studied.

1.2 Rate of Convergence

A natural question may arise is that when the value of \( \mu_t \) is close to its previous value \( \mu_{t-1} \). To see this, let \( a_t = \frac{1}{1 - k_t} \) and notice that

\[ |\mu_t - \mu_{t-1}| \leq |\mu_{t-1}| |a_t(1 - k_t) - 1| + |k_t \gamma_t| \leq |k_t \gamma_t|. \]

Since \( \gamma_t \approx \beta_t \), thus

\[ |\mu_t - \mu_{t-1}| \leq \frac{\beta_t |a_t - 1|}{a_t}. \]

The following proposition summaries the above result and proposes the rate of convergence of \( \mu_t \).

Proposition 1. Suppose that
\[
\frac{\beta_t |a_t - 1|}{a_t} = O(c_t) \text{ and } c_t \to 0 \text{ as } t \to \infty.
\]

\[
|\mu_t - \mu_{t-1}| = O(c_t).
\]

Here, we continue as follows. Write \(\mu_t = a_t \mu_{t-1} + k_t(y_t - a_t \mu_{t-1}) = a_t \mu_{t-1} + k_t Z_t\). So

\[
\mu_t - a_t \mu_{t-1} = k_t Z_t = k_t \left( \beta_t + \frac{\varepsilon_t}{(R^m_t - r_f)} - a_t \beta_{t-1} \right) = k_t \left( \frac{\varepsilon_t}{(R^m_t - r_f)} + \xi_t \right)
\]

The variance of error term \(k_t Z_t\) is \(\text{var}(k_t Z_t) = k_t^2 (u_t^2 + \delta^2_t)\). Now, Suppose that

\[
\sqrt{k_t^2 (u_t^2 + \delta^2_t)} = O(d_t), d_t \to 0 \text{ as } t \to \infty,
\]

Then, \(\mu_t\) behaves like the \(a_t \mu_{t-1}\). The following proposition summaries the above result and proposes the rate of convergence of \(\mu_t\).

**Proposition 2.** Suppose that

\[
\sqrt{k_t^2 (u_t^2 + \delta^2_t)} = O(d_t), d_t \to 0 \text{ as } t \to \infty,
\]

Then, \(\mu_t\) behaves like the \(a_t \mu_{t-1}\).

Also, it is seen that

\[
\vartheta_t^2 - \vartheta_{t-1}^2 = ((1 - k_t)a_t^2 - 1)\vartheta_{t-1}^2 + (1 - k_t)\delta_t^2.
\]

Let \(1 - k_t = \frac{1}{a_t^2}\). Therefore, if \(\frac{\delta_t^2}{a_t^2} \to 0\) as \(t \to \infty\), then \(\vartheta_t^2 - \vartheta_{t-1}^2 \to 0\). The following proposition summaries the above result.

**Proposition 3.** Suppose that \(1 - k_t = \frac{1}{a_t^2}\), and \(\frac{\delta_t^2}{a_t^2} \to 0\) as \(t \to \infty\), then

\[
\vartheta_t^2 - \vartheta_{t-1}^2 = O \left( \frac{\delta_t^2}{a_t^2} \right).
\]

Sometimes \(\vartheta_t^2\) behaves like \(a_t^2 \vartheta_{t-1}^2\). The following proposition states this fact.
Proposition 4. Assume that $1 - k_t = \frac{1}{\delta_t}$ and $a_t^2 k_t \to 0$. Then, $\vartheta_t^2$ behaves like $a_t^2 \vartheta_{t-1}^2$.

2 Sensitivity analysis

Here, the sensitivity of Kalman estimates with respect to their parameters is considered.

2.1 Effect of heteroskedasticity

When volatility sequence $\sigma_t^2$ follows a GARCH series, it has too fluctuations. Thus, it may be too large or too small. First, suppose that for some time point $t$, variance term $\sigma_t^2$ gets large, then $u_t^2$ goes to infinity (if $R_t^m - r_f$ is small with respect to $\sigma_t^2$), and $k_t$ goes to zero. Therefore

$$\mu_t = a_t \mu_{t-1}.$$  

Here, $\vartheta_t^2$ behaves like the $\tau_t^2$. Also, suppose that variance $\sigma_t^2$ goes to zero, then $k_t$ closes to unity, therefore $\mu_t = \gamma_t$, where it is the $\beta_t$, approximately. In this case, it is seen that $\vartheta_t^2$ goes to zero. Therefore, considering a GARCH series for $\sigma_t^2$, then

$$\begin{aligned}
\left\{ \begin{array}{ll}
a_t \mu_{t-1} & \text{if } \sigma_t^2 \to \infty \\
\beta_t & \text{if } \sigma_t^2 \to 0
\end{array} \right.
\end{aligned}$$

This phenomena usually happens and is referred as volatility clustering, when the arch or GARCH coefficients of heteroskedasticity model is too large or too small.

2.2 Sensitivity to $a_t$.

The partial derivative of $\mu_t$ with respect to $a_t$ is given by
\[
\frac{\partial \mu_t}{\partial a_t} = (1 - k_t) \mu_{t-1}.
\]
As \( k_t \) goes to one, the effect of \( a_t \) on \( \mu_t \) is small. To study the effect of \( a_t \) on \( \vartheta_t^2 \) let \( b_t = a_t^2 \). Then,
\[
\frac{\partial \vartheta_t^2}{\partial b_t} = \frac{\partial \vartheta_t^2}{\partial \tau_t^2} \times \frac{\partial \tau_t^2}{\partial b_t} = \frac{u_t^4 \vartheta_{t-1}^2}{(\vartheta_t^2 + u_t^4)^2}
\]
As the \( u_t \) goes to zero, then \( a_t \) has no effect on \( \vartheta_t^2 \). However, as \( u_t \) goes to infinity, then
\[
\frac{\partial \vartheta_t^2}{\partial b_t} = \frac{\vartheta_{t-1}^2}{(1 + k_t)^2}.
\]

**Remark 4.** This method may be applied in hedging a portfolio. Suppose that a portfolio contains a stock where its dynamic of return follows the model of Introduction. Then, the mean of \( \beta_t \) may change as

\[
\frac{u_t^4 \vartheta_{t-1}^2}{(\vartheta_t^2 + u_t^4)^2}
\]
because of variation in \( b_t \). Therefore, to hedge this risk, it is possible to modify the \( \mu_t \) to

\[
\mu_t - b_t \frac{u_t^4 \vartheta_{t-1}^2}{(\vartheta_t^2 + u_t^4)^2}.
\]
The other sensitivity may be applied in the same way.

### 2.3 Effect of Kalman gain

Again, it is seen that
\[
\frac{\partial \mu_t}{\partial k_t} = \gamma_t - a_t \mu_{t-1} \quad \text{and} \quad \frac{\partial \vartheta_t^2}{\partial k_t} = -\tau_t^2.
\]
As \( u_t \) is large (small), then the sensitivity of \( \vartheta_t^2 \) to \( k_t \) is small (large), conversely.
2.4 Effect of normality

The proposed results lies heavily on the normality of $\varepsilon_t$. Here, it is assumed that $\varepsilon_t$ has a $\alpha - \text{percent pollution normal distribution given by}$

$$aN(0, \sigma_t^*^2) + (1 - \alpha)N(0, \sigma_t^2).$$

Therefore,

$$\begin{cases}
    u_t^*^2 = \frac{\alpha \sigma_t^*^2 + (1 - \alpha)\sigma_t^2}{(R_t^m - r_f)^2} \\
    k_t^* = \frac{\tau_t^2}{u_t^*^2 + \tau_t^2}.
\end{cases}
$$

One can see that

$$\frac{\partial \mu_t}{\partial \alpha} = \frac{\partial \mu_t}{\partial k_t^*} \times \frac{\partial k_t^*}{\partial \alpha} = \frac{-(\gamma_t - a_t \mu_{t-1})\tau_t^2}{(u_t^*^2 + \tau_t^2)^2}.$$

As $\alpha$ goes to one, then $\sigma_t^2$ is replaced by $\sigma_t^*^2$, more. In the following Remark, the Bayes filtration is suggested for general non-normal cases.

**Remark 5.** Kalman filtering uses the normality assumption for $\varepsilon_t$. However, it is not a realistic assumption, in practice. Historical data analysis shows that fat tail distributions are usually suitable for $\varepsilon_t$. Thus, in the case of heavy tail distribution, Kalman filter fails and some extensions like the particle filters or generally the Bayes filter are needed (Arulampalam *et al.*). Using the Chapman-Kolmogorov equation, the Bayes prediction step is given by

$$f(\beta_t|R_{t-1}, ..., R_1) = \int f(\beta_t|\beta_{t-1}) f(\beta_{t-1}|R_{t-1}, ..., R_1) d\beta_{t-1},$$

and the Bayes update equation is

$$f(\beta_t|R_t, ..., R_1) \propto f(R_t|\beta_t) f(\beta_t|R_{t-1}, ..., R_1).$$

In order to initialize the recurrence algorithm, it is assumed that the initial return $R_0$ has known probability distribution $f(R_0)$. Using the Bayes filter, the probability distribution $f(\beta_t|R_t, ..., R_1)$ and $f(\beta_t|R_{t-1}, ..., R_1)$ are not computed. Only, the expectations $E(\beta_t|R_t, ..., R_1)$ and $E(\beta_t|R_{t-1}, ..., R_1)$ are calculated.
Remark 6. Distribution of $\beta_{t+1}$. Suppose that the mean of market risk of a special stock at time $t$ is $\mu_t$. Using the Kalman estimate $\mu_{t+1}$, it is possible to forecast the future market risk distribution $\beta_{t+1}$. Notice that

$$f(\beta_{t+1} | \mu_t = \mu^*) = \int f(\beta_{t+1} | \mu_{t+1}) f(\mu_{t+1} | \mu_t = \mu^*) d\mu_{t+1},$$

where $f(\beta_{t+1} | \mu_{t+1})$ comes from normal distribution with mean $\mu_{t+1}$ and $\theta_{t+1}^2$ and $f(\mu_{t+1} | \mu_t = \mu^*)$ is a normal with mean

$$\frac{k_{t+1} y_{t+1}}{1 - a_{t+1} (1 - k_{t+1})}$$

and variance $a_{t+1}^2 (1 - k_{t+1})^2 \theta_{t+1}^2$. Term $\theta_{t+1}^2$ is obtained by updating the variance term of Kalman equation.

Remark 7. Application in VaR. One of the main factors which exists in each financial activity is risk. The risk induces the uncertainty to the financial problems and therefore decision making is too difficult in such conditions. Indeed, after famous financial disasters, it is advised to estimate the market risk (see, Alexander (2001)). One of these risk measures is VaR. VaR calculations attempt to provide a risk assessment of the form: we are $(1 - \alpha)$% certain that we will not lose more than $Y$ dollars in the next $N$ days. The variable $Y$ is function of two parameters. The first is $N$, the time horizon, and the second is $(1 - \alpha)$, the confidence level.

For a portfolio with initial value $V_0$, the $(1 - \alpha)$ % VaR is $-V_0 q_{1-\alpha}$, where $q_{1-\alpha}$ is the $\alpha$-th quantile of return process. Here, suppose that $\text{VaR}_{t+1}$ is the value of VaR at time point $t + 1$. Thus,

$$P\left( R_{t+1} \leq \frac{\text{VaR}_{t+1}}{-V_0} | \mu_t = \mu^* \right) = \int P\left( R_{t+1} \leq \frac{\text{VaR}_{t+1}}{-V_0} | \mu_{t+1}, \mu_t = \mu^* \right) f(\mu_{t+1} | \mu_t = \mu^*) d\mu_{t+1} =$$
where density \( f(\mu_{t+1}|\mu_t = \mu^*) \) proposed in Example 4 and \( \Phi \) is the CDF of standard normal distribution.

\[
E_{\mu_{t+1}|\mu_t = \mu^*} \left( \Phi \left( \frac{-V_0^{-1} \text{VaR}_{t+1} - \mu_{t+1}}{\nu_{t+1}} \right) \right) = \alpha
\]

### 3 Comparisons

Here, we compare the Kalman filter estimate with the method of moment estimate. One can see that

\[
R_t = r_f + a_t \beta_{t-1} (R^m_t - r_f) + (R^m_t - r_f) \xi_t + \epsilon_{t+1}.
\]

Thus, the method of moment (mm) estimate of \( \hat{\mu}_t \) is given

\[
\hat{\mu}_t = E(R_t|R^m_t) = E(E(R_t|\beta_{t-1}, R^m_t)) = r_f + a_t \mu_{t-1} (R^m_t - r_f).
\]

It is seen that

\[
\mu_t - E(\hat{\mu}_t) = (k_t \gamma_t - r_f) - a_t \mu_{t-1} (k_t - r_f + E(R^m_t) - 1).
\]

Suppose that \( \gamma_t \) is close to one and \( E(R^m_t) \) is close to unity. Then,

\[
\mu_t - E(\hat{\mu}_t) = (k_t - r_f)(1 - a_t \mu_{t-1}).
\]

When \( k_t \) is close to \( r_f \) then \( \mu_t \) is close to \( E(\hat{\mu}_t) \). The following proposition summarizes the above discussion.

**Proposition 5.** Suppose that \( \gamma_t \) is close to one and \( E(R^m_t) \) is close to unity and \( k_t \) is close to \( r_f \). Then, the Kalman estimate \( \mu_t \) is close to the expectation of mm estimate of \( \hat{\mu}_t \). Next, the proportion of variance of \( R_t \) to is given by

\[
\frac{\sigma^2}{\sigma^2_{\hat{\mu}}} = \frac{a^2_t}{\sigma^2_{\hat{\mu}}} (\mu^2_{t-1} + \theta^2_{t-1} + \delta^2_t) var(R^m_t) + \frac{\sigma^2_t}{\sigma^2_{\hat{\mu}}}. \]

Suppose that \( \frac{a^2_t}{\sigma^2_{\hat{\mu}}} \to 0 \) then \( \frac{\sigma^2}{\sigma^2_{\hat{\mu}}} = \frac{\sigma^2_t}{\sigma^2_{\hat{\mu}}}. \) That is as soon as \( \sigma^2_t < \delta^2_t \), then the Kalman filter method works better that the method of moment estimate.
Proposition 6. If $\frac{a_t^2}{\sigma_t^2} \to 0$, then $\frac{\sigma_t^2}{\sigma_t^2} = \frac{a_t^2}{\sigma_t^2}$.

4 Some examples

In this section, some examples are proposed.

Example 1: Simulation. For first Example, let $\sigma_t^2 = \delta_t^2 = 1$ and the monthly risk free rate if $r_f = 0.05$. It is assumed that $R_t^m$ is normally distributed with mean 0.6 and standard deviation 0.1. Coefficients $\beta_t$ are sampled from a first order autoregressive model with mean 1, $a_t = 0.2$ and $R_t$'s are computed. The actual beta is given by 0.083. The following plot again proposes the accuracy of presented method.

Table 1: Descriptive statistics of $\mu_t$ and $\vartheta_t$

<table>
<thead>
<tr>
<th>Kalman est.</th>
<th>Min</th>
<th>1st Qu</th>
<th>Median</th>
<th>Mean</th>
<th>3rd Qu</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_t$</td>
<td>0.011</td>
<td>0.057</td>
<td>0.0846</td>
<td>0.0826</td>
<td>0.106</td>
<td>0.13</td>
</tr>
<tr>
<td>$\vartheta_t$</td>
<td>0</td>
<td>0.529</td>
<td>0.739</td>
<td>0.716</td>
<td>0.933</td>
<td>1.01</td>
</tr>
</tbody>
</table>
Example 2: Real data set. Here, a real data set is studied. The S&P 500 stock market index comprises 505 common stocks where one of them is Affiliated Managers Group (AMG) stock. It is a global asset management company. The monthly and daily risk free rates are 0.13, 0.00433, respectively. To test the truth of CAPM, the model

$$R_t - r_f = \alpha + \beta(R^m_t - r_f) + \epsilon_t$$

is fitted and it is tested to check if $H_0: \alpha = 0$ isn't rejected or not. The data set is chosen such that the null hypothesis is retained. The following Table gives the values of $R^2$, the p-value for $H_0: \alpha = 0$ and the estimated beta for various choice of sample size $n$.

<table>
<thead>
<tr>
<th>Sample size selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
</tr>
<tr>
<td>p-value</td>
</tr>
<tr>
<td>$R^2$</td>
</tr>
<tr>
<td>$\beta$</td>
</tr>
</tbody>
</table>
Using this strategy, the data set (taken from Google-finance) contains 150 daily returns for time period January 2, 2015 to August 7, 2015. During this period, the best fit of CAPM model is observed (see Table 2). The sequential least square estimates are computed as the time varying betas. Using the Modelrisk Vose software (Adds-in Excel) an first order Autoregressive AR(1) model with mean 0.182 is fitted to betas and it is seen that $a_t = 0.8436$ and $\delta_t = 0.1068$. Also, a GARCH(1,1) is fitted for $\sigma_t^2$. It is seen that

$$\sigma_t^2 = 0.000018 + 0.584\varepsilon_t^2 + 0.416\sigma_{t-1}^2.$$ 

The following Table gives the descriptive characteristics Kalman estimates $\mu_t$ and $\sigma_t^2$ of Example 1.

<table>
<thead>
<tr>
<th>Kalman est.</th>
<th>Min</th>
<th>1st Qu</th>
<th>Median</th>
<th>Mean</th>
<th>3rd Qu</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_t$</td>
<td>0</td>
<td>1.33</td>
<td>1.35</td>
<td>1.34</td>
<td>1.38</td>
<td>1.39</td>
</tr>
<tr>
<td>$\sigma_t$</td>
<td>0</td>
<td>0.0239</td>
<td>0.0239</td>
<td>0.0237</td>
<td>0.0239</td>
<td>0.0239</td>
</tr>
</tbody>
</table>

The mean of $\mu_t$ corresponds to the actual beta which is 1.37. The standard deviation of $\mu_t$ also shows the accuracy of results. The following plot shows the convergence of Kalman estimates of $\mu_t$ on actual estimate of beta. In the above example, Kalman filter works well because in spite of existence of a GARCH series, it does not generate too large or too small variance values.

The stability of $\mu_t$ can be checked by drawing the CUSUM plot. It is presented as follows

$$\text{cusum}_t = \sum_{i=1}^{t} (\beta_i - \bar{\beta}), t = 1,2, ... 149.$$ 

If there is no change in betas then this plot fluctuates around zero. The below figure shows that the stability of betas in mean is failed and a time varying beta CAPM model is a suitable selection.
5 Concluding remarks

In this paper, the Bayesian approach is applied to provide recursive Kalman estimation of the time varying beta of the CAPM model. Applications of method are presented and rate of convergence are derived. Comparisons with method of moment estimates are given.
References


