Particle filters for continuous likelihood evaluation and maximisation

Shasha CUI, Wensi DING, Quan SUN

ENSAE

19 Jan 2018

Plan

- Introduction to algorithm
- Application
  - AR(1) plus noise model
  - Stochastic Volatility model with leverage
  - GARCH plus error model
  - Continuous time model
- Conclusion

Algorithm

Task: **Continuous** estimation of the likelihood with respect to parameters θ, $logL(\theta) = logp(y_1, \cdots, y_T|\theta)$

The boostrap filter: algorithm

All operations to be performed for all $n \in 1:N$.

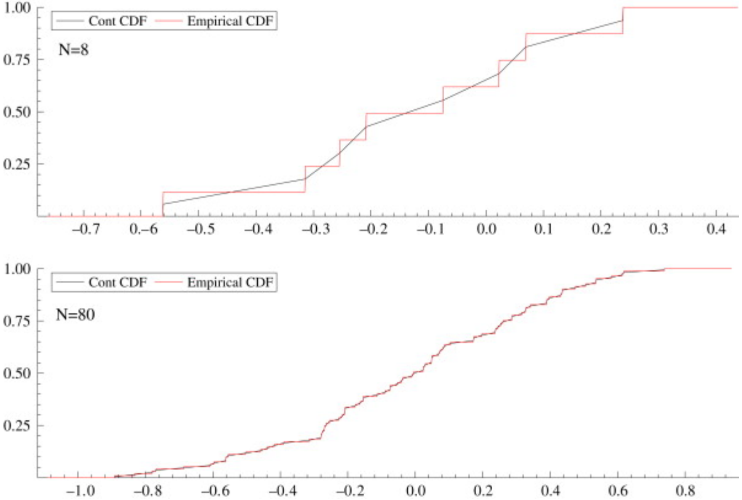
At time 0:

(a) Generate $X_0^n \sim P_0(\mathrm{d}x_0)$.
(b) Compute $w_0^n = f_0(y_0|X_0^n)$, $W_0^n = w_0^n / \sum_{m=1}^N w_0^m$, and $L_0^N = N^{-1} \sum_{n=1}^N w_0^n$.

Recursively, for $t = 1, \ldots, T$:

(a) Generate ancestor variables $A_t^n \in 1:N$ independently from $\mathcal{M}(W_{t-1}^{1:N})$.
(b) Generate $X_t^n \sim P_t(X_{t-1}^{A_t^n}, \mathrm{d}x_t)$.
(c) Compute $w_t^n = f_t(y_t|X_t^n)$, $W_t^n = w_t^n / \sum_{m=1}^N w_t^m$, and $L_t^N = L_{t-1}^N \{N^{-1} \sum_{n=1}^N w_t^n\}$.

Algorithm


1.00
0.75
0.50
0.25
Cont CDF
Empirical CDF
N=8
–0.7
0.–6
–0.5
–0.4
–0.3
–0.2
–0.1
0.0
0.1
0.2
0.3
0.4
1.00
0.75
0.50
0.25
Cont CDF
Empirical CDF
N=80
–1.0
–0.8
–0.6
–0.4
–0.2
0.0
0.2
0.4
0.6
0.8

Algorithm: Convergence Results

- Particle filter delivers a consistent and an unbiased estimator for the true likelihood function $L(\theta)$
- The resulting simulated maximum likelihood (SML) estimator is consistent if T and $N \to \infty$
- For the corresponding estimator of the log-likelihood, $\sqrt{N}\{log\hat{L}_N(\theta) - logL(\theta)\} \to N(-\frac{\sigma^2_{SMC,T}}{2\sqrt{N}}, \sigma^2_{SMC,T})$
- The bias introduced by using the approximation of the empirical distribution function is only of order $1/N$

Application: AR(1) plus noise model

AR(1) plus noise model

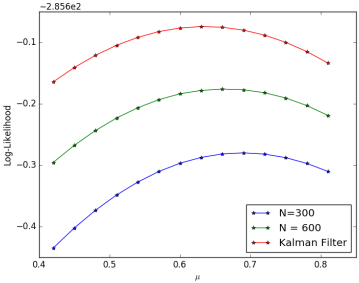
- Model:
  - $y_t = x_t + \epsilon_t, \epsilon_t \sim N(0, \sigma_\epsilon^2)$
  - $x_{t+1} = \mu + \phi(x_t - \mu) + \eta_t, \eta_t \sim N(0, \sigma_\epsilon^2)$

Application: AR(1) plus noise model

AR(1) plus noise model

- Model:
  - $y_t = x_t + \epsilon_t, \epsilon_t \sim N(0, \sigma_\epsilon^2)$
  - $x_{t+1} = \mu + \phi(x_t - \mu) + \eta_t, \eta_t \sim N(0, \sigma_\epsilon^2)$
- Implementation:
  - CSIR Method
  - Kalman Filter

Application: AR(1) plus noise model



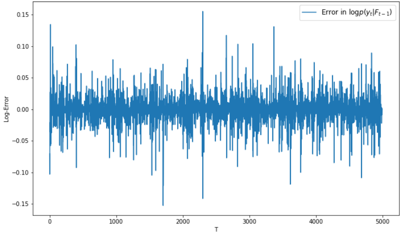


Figure: Fixed dataset, boxplot of 50 runs with different random seeds, T = 150, N = 300 (left), N = 600 (right)

Application: AR(1) plus noise model

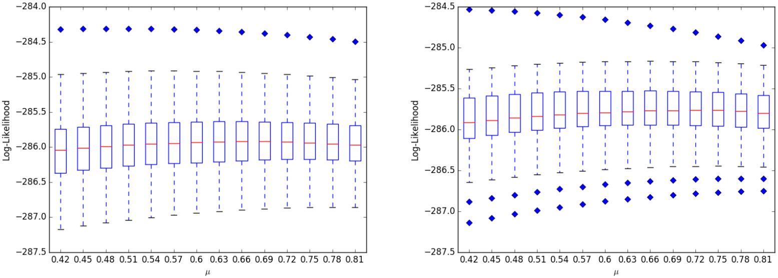


Figure: Fixed dataset, boxplot of 50 runs with different random seeds, T = 150, N = 300 (left), N = 600 (right)

Application: AR(1) plus noise model

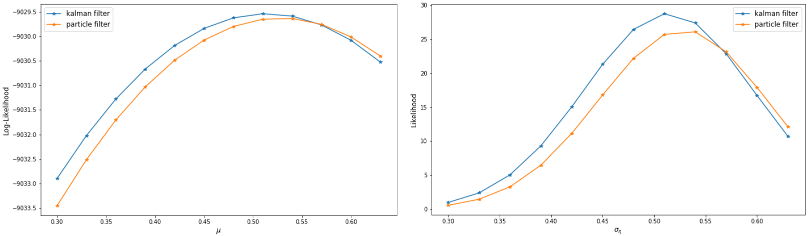


Figure: The true and estimated log-likelihood (left) and likelihood (right) profiles for the AR(1) model plus noise for (from top to bottom) σ_η, ϕ, μ. Length of series $T = 5000$. True parameters: $\sigma_\eta = \sqrt{0.02}, \mu = 0.5, \phi = 0.975$ and $N = 300$

Application: AR(1) plus noise model

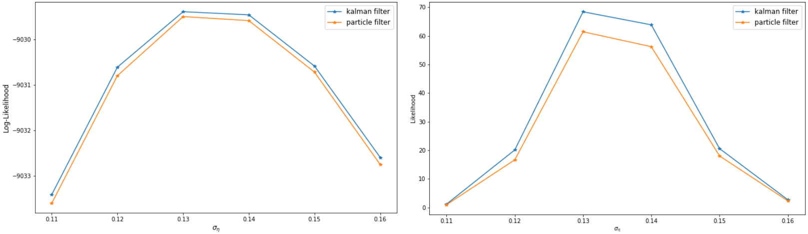


Figure: The true and estimated log-likelihood (left) and likelihood (right) profiles for the AR(1) model plus noise for (from top to bottom) σ_η, ϕ, μ. Length of series $T = 5000$. True parameters: $\sigma_\eta = \sqrt{0.02}, \mu = 0.5, \phi = 0.975$ and $N = 300$

Application: AR(1) plus noise model

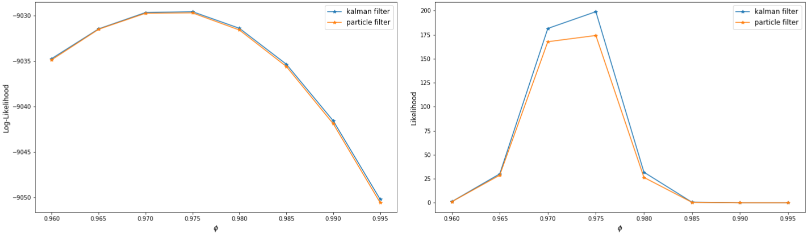


Figure: The true and estimated log-likelihood (left) and likelihood (right) profiles for the AR(1) model plus noise for (from top to bottom) σ_η, ϕ, μ. Length of series $T = 5000$. True parameters: $\sigma_\eta = \sqrt{0.02}, \mu = 0.5, \phi = 0.975$ and $N = 300$

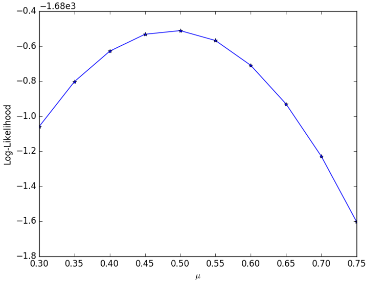
Applicaton: Stochastic Volatility model with leverage

Model:

- $y_t = \epsilon_t \exp(x_t/2), t = 1, ..., T$
- $x_{t+1} = \mu(1-\phi) + \phi x_t + \sigma_\eta \eta_t$
- $\begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} \sim N(0, \Sigma), \Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$

Applicaton: Stochastic Volatility model with leverage

$$\mu = 0.5, \phi = 0.975, \sigma_\eta^2 = 0.02, \rho = -0.8$$



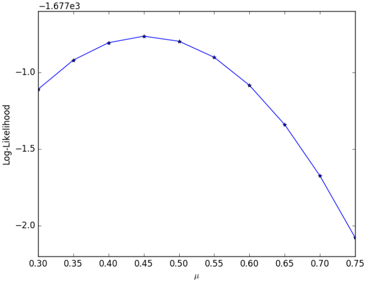


Figure: Fixed dataset, mean value of 50 runs with different random seeds, T = 1000, N = 300 (left), N = 600 (right)

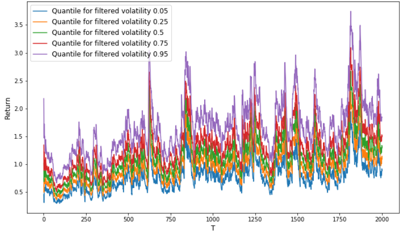
Applicaton: Stochastic Volatility model with leverage

μ	0.3	0.4	0.5	0.6	0.65	0.7	0.75
N = 300	11.226	10.766	10.390	10.246	10.097	9.970	9.845
N = 600	3.468	3.356	3.282	3.270	3.253	3.236	3.226

Table: Empirical variance comparaison with N=300 and N=500

Applicaton: Stochastic Volatility model with leverage

2000 daily returns of SP 500 from 16 May 1995 to 24 April 2003



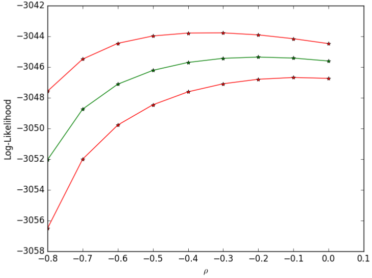


Figure: Results for SP 500 index returns. SV with leverage model. (left):.The quantiles for filtered volatility (0.05, 0.25, 0.5, 0.75, 0.95).

Application: Garch plus noise model

Model:

- $y_t|_t \sim N(x_t, \sigma^2)$
- $x_t|\sigma_t^2 \sim N(0, \sigma_t^2)$

Model:

- $y_t|\sigma_t^2 \sim N(0, \sigma^2 + \sigma_t^2)$
- $x_t|\sigma_t^2, y_t \sim N(\frac{b^2 y_t}{\sigma^2}, b^2)$

$\sigma_{t+1}^2 = \beta_0 + \beta_1 x_t^2 + \beta_2 \sigma_t^2$

with $b^2 = \frac{\sigma^2 \sigma_t^2}{\sigma^2 + \sigma_t^2}$

Application: Garch plus noise model

Result of data generated from
$\beta_0 = 0.01, \beta_1 = 0.2, \beta_2 = 0.75, \sigma = 0.1$

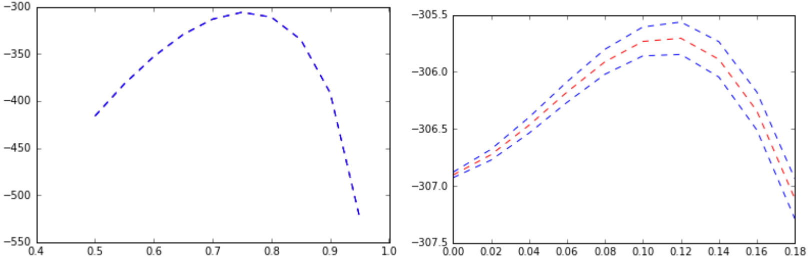


Figure: The sliced estimated log-likelihood for the GARCH model plus noise for β_2 (left),σ (right), $T = 500, N = 500$

Application: Garch plus noise model

Result of continuous compounded daily return on UK pounds vs US dollars from 2 January 1982 to 29 December 1982:

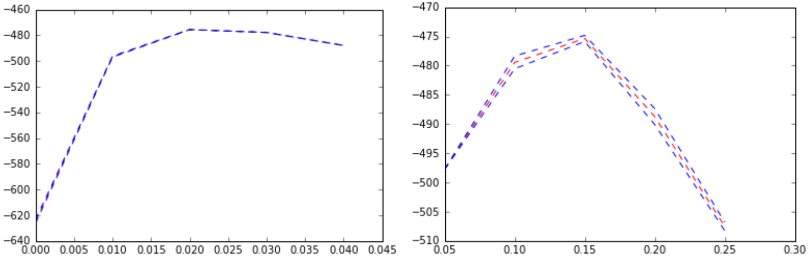


Figure: estimated likelihood of β_1 of us vs uk dataset,left: $\sigma = 0$, right: $\sigma = 0.55$, $T = 500$

Log Likelihood with noise -473.9, without noise -476.02.
Likelihood ratio test: reject $\sigma = 0$ with 1% level of significance

Application: Continuous time model

General Case: Model

- $dy(t) = \{\mu + \beta\sigma^2(t)\}dt + \sigma(t)dW_1(t)$
- $d\sigma^2(t) = a(\sigma^2(t))dt + b(\sigma^2(t))dW_2(t)$
- $Corr(W_1(t), W_2(t)) = \rho$
- with observations at time: $\tau_1 < \tau_2 < \cdots < \tau_n < \tau_{n+1}$

- $r_s = y(\tau_{s+1}) - y(\tau_s)$
- $r_s \sim N(\mu\Delta_s + \beta\sigma_s^{2*} + \rho Z_s, (1-\rho^2)\sigma_s^{2*})$
- for $s = 1, \cdots, n$
- $\Delta_s = \tau_{s+1} - \tau_s$, $\sigma_s^{2*} = \int_{\tau_s}^{\tau_{s+1}} \sigma^2(u)du$, $Z_s = \int_{\tau_s}^{\tau_{s+1}} \sigma(u)dW_2(u)$

Application: Continuous time model

General Case: Euler approximation

Place $M_s - 1$ evenly spaced latent points between $\sigma^2(\tau_s)$ and $\sigma^2(\tau_{s+1})$, noted by $\sigma^2_{s,1}, \cdots, \sigma^2_{s,M_s-1}$, with time interval between latent points $\delta = \Delta_s / M_s$

- $\sigma^2_{s,m+1} = \sigma^2_{s,m} + a(\sigma^2_{s,m})\delta + b(\sigma^2_{s,m})\sqrt{\delta} u_m$
- for $m = 0, \cdots, M_s - 1$, where $u_m \sim NID(0,1)$
- $r_s \sim N(\mu\Delta_s + \beta\hat{\sigma}^{2*}_s + \rho\hat{Z}_s, (1-\rho^2)\hat{\sigma}^{2*}_s)$
- $\hat{\sigma}^{2*}_s = \delta \sum_{m=0}^{M_s-1} \sigma^2_{s,m}$, $\hat{Z}_s = \sqrt{\delta} \sum_{m=0}^{M_s-1} \sigma_{s,m} u_m$

Application: Continuous time model

Adjustment for Resampling

Before: $\omega_j = p(r_s | \hat{\sigma}_s^{2*j}; \hat{Z}_s^{\,j}),\ \pi_j = \frac{\omega_j}{\sum_{i=1}^{N} \omega_i},\ j = 1, \cdots, N$

After: $\omega_j^* = \frac{\sum_{i=1}^{N} \omega_i \phi((x_j - x_i)/h)}{\sum_{i=1}^{N} \phi((x_j - x_i)/h},\ \pi_j^* = \frac{\omega_j^*}{\sum_{i=1}^{N} \omega_i^*}$

Nelson volatility process

$$dy(t) = \sigma(t) dW_1(t)$$

$$d\sigma^2(t) = k(\theta - \sigma^2) dt + \sqrt{\xi} \sigma^2 dW_2(t)$$

Application: Continuous time model

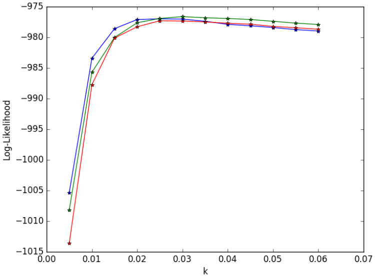


Figure: The log-likelihood for Nelson volatility model. Three runs of a fixed sample, n=1000, N=600, true parameters k=0.02, θ=0.5, ξ=0.0178

Application: Continuous time model

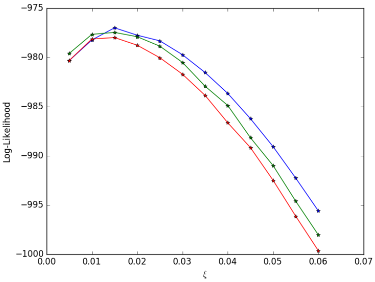


Figure: The log-likelihood for Nelson volatility model. Three runs of a fixed sample, n=1000, N=600, true parameters k=0.02, θ=0.5, ξ=0.0178

Conclusion

- Investigation and implementation of CSIR method for different models
- Good performance for both simulated data and real data