MLEMVD: A R Package for Maximum Likelihood Estimation of Multivariate Diffusion Models

Matthew Dixon and Tao Wu[1]

Illinois Institute of Technology

May 19th 2017

[1] https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2944341

Time Homogeneous Markovian Diffusions

- (Sahalia2002) consider the multivariate time-homogenous Markovian diffusion of the form

$$d\mathbf{X}_t = \mu(\mathbf{X}_t)dt + \Sigma(\mathbf{X}_t)\mathbf{dW}_t,$$

- $\mu(x) : \mathbb{R}^m \to \mathbb{R}^m, \Sigma(x) : \mathbb{R}^m \to \mathbb{R}^{m \times m}$
- $\mathbf{W}_t \in \mathbf{R}^m$ are independent Wiener processes.
- Key idea: The log of the transition function $f_X(x|x_0, \Delta)$ can be given in closed form under mild restrictions on the form of μ and Σ.

Time Homogeneous Markovian Diffusions

- We observe X_t at times $t_0, t_1, \ldots, t_n$, where Δ denotes the difference between observation times and is assumed independent.
- Under this finite data, the log-likelihood takes the form:

$$\ell_n(\mathbf{p}) := \sum_{i=1}^{n} \ell_X(\Delta, x_{t_i} | x_{t_{i-1}}; \mathbf{p}), \qquad \ell_X := ln\ f_X.$$

- Under a Hermite expansion of ℓ_X and application of a number of transformation we get the following compact closed form expression with K terms:

$$\ell_X^{(K)}(x|x_0) = -\frac{m}{2} ln(2\pi\Delta) - D_v(x) + \frac{C_X^{-1}(x|x_0)}{\Delta} + \sum_{k=0}^{K} C_X^{(k)}(x|x_0)\frac{\Delta^k}{k!},\ D_v := -\frac{1}{2} ln(Det[v(x)]).$$

Example: Geometric Brownian Motion

- Given an initial value $X_0 = x_0$, X_t evolves with drift parameter μ and volatility parameter σ

$$dX_t = \mu X_t dt + \sigma X_t dW_t.$$

- The transition function takes the form

$$f_X(x_t|x_0, t) = \frac{1}{\sqrt{2\pi}\sigma t} \exp - \frac{(ln x_t - ln x_0 - (\mu - \sigma^2/2)t)^2}{2\sigma^2 t}$$

- The exact log likelihood function, evaluated over a uniform time series of n observations of the state variable with spacing Δ is given by

$$\ell_n(\mathbf{p}) = \sum_{i=1}^{n} \ell_X(\Delta, x_{t_i}|x_{t_{i-1}}; \mathbf{p}) = -\frac{1}{2}\sum_{i=1}^{n}(ln(2\pi\Delta\sigma^2 x_{t_i}^2) + (ln[x_{t_i}/x_{t_{i-1}}] - (\mu - \sigma^2/2)\Delta))^2/(\sigma^2\Delta).$$

Example: Heston Model

- The Heston model describes the evolution of the log of the stock price $s_t = ln\ S_t$

$$
\begin{aligned}
ds_t &= (a + bY_t)dt + \sqrt{Y_t}dW_1^Q(t), \qquad (1) \\
dY_t &= \kappa'(\theta' - Y_t)dt + \sigma\sqrt{Y_t}dW_2^Q(t), \qquad (2)
\end{aligned}
$$

- where

$$
a = r - d, \qquad b = -\frac{1}{2}.
$$

Heston Model

- A key characteristic of the model is that the Wiener processes are correlated $dW_1^Q \cdot dW_2^Q = \rho dt =>$ leverage effect.
- There are five parameters in the model
  - κ: mean-reversion rate
  - θ: long-term variance
  - σ: volatility of variance
  - ρ: instantaneous correlation between dW_1^Q and dW_2^Q
  - y_0: initial variance

Likelihood function estimation: full observance

- State vector $X_t := [s_t, Y_t]$ has a transition density function for the conditional density of $X_{t+\Delta} = x$ given $X_t = x_0$ denoted by $f_X(\Delta, x|x_0; \mathbf{p})$.
- The log likelihood function for observations at times $t_0, t_1, \ldots, t_n$ is given by

$$\ell_n(\mathbf{p}) = \frac{1}{n}\sum_{i=1}^{n} \ell_X(\Delta, x_{t_i}|x_{t_{i-1}}; \mathbf{p}),$$

- $\ell_X(\Delta, x|x_0; \mathbf{p}) := \mathit{ln}\ f_X(\Delta, x|x_0; \mathbf{p})$ and is given in closed form.

Likelihood function estimation: partial observance

- $X_t := [s_t; Y_t]'$ is partially observed and hence ℓ_n can not be directly estimated from time series data.
- Approximate the likelihood of the observed state vector $G_t := [s_t; C_t]'$, where C_t is the ATM constant maturity option price.
- The transition density function for the conditional density of $G_{t+\Delta} = g$ given $G_t = g_0$ is now denoted by $f_G(\Delta, g|g_0; \mathbf{p})$ and the log likelihood function is given by

$$\ell_n(\mathbf{p}) := \frac{1}{n}\sum_{i=1}^{n} \ell_G(\Delta, g(t_i)|g(t_{i-1}); \mathbf{p}).$$

Likelihood function estimation: partial observance

- $G_t = f(X_t; \mathbf{p})$ represents the stock and option prices
- Under a change of variables from G_t to X_t, the log of the transition density f_G can be expressed in terms of f_X through a Jacobian J_t to give:

$$\ell_G(\Delta, g|g_0; \mathbf{p}) := -ln\, J_t(\Delta, g|g_0; \mathbf{p}) + \ell_X(\Delta, f^{-1}(g; \mathbf{p})|f^{-1}(g_0; \mathbf{p}); \mathbf{p})$$

- **Goal**: Find a numerical approximation to the inverse map $f^{-1}(G; \mathbf{p})$

Pricing

- The mapping between X_t and G_t is given by

$$f(X_t; \mathbf{p}) = \left[\begin{array}{c} s_t \\ C(S_t, Y_t, K, \Delta; \mathbf{p}) \end{array} \right],$$

- $C(\cdot)$ is the Heston model option price defined above and $\Delta = T - t$ is the constant time to maturity of the option.

Calibration: overview

- Since Y_t, as previously mentioned, is unobserved, we seek the inverse

$$f^{-1}(G_t; \mathbf{p}) = \begin{bmatrix} s_t \\ C^{-1}(S_t, C_t, K, \Delta; \mathbf{p}) \end{bmatrix} = \begin{bmatrix} s_t \\ Y_t \end{bmatrix}$$

- Use the map from G_t to X_t to imply the volatility y_t from the observed state vector $\{g_t\}_{i=1}^{n}$ of underlying prices and corresponding constant maturity, ATM option prices.
- Find the maximum likelihood estimate $\mathbf{p}^*$.

Calibration: Step 1

- Fix $\mathbf{p}$ and find the corresponding implied volatility $y_{t_i}^p$

$$y_{t_i}^p = \arg\min_y |c_{t_i} - C(S_{t_i}, y, K, \Delta; \mathbf{p})|.$$

- This error function is convex and so has a unique solution independent of the initial choice of y in a simple root finding method.

Calibration: Step 2

- Using $(y^p_{t_i}, \mathbf{p})$ we compute the Jacobian of G_t w.r.t. X_t and the log of transition density function $\ell_X(\cdot)$

$$\ell_G(\Delta, g_{t_i}|g_{t_{i-1}}; \mathbf{p}) = -lnJ_{t_i} + \ell_X(\Delta, x^p_{t_i}|x^p_{t_i}\mathbf{p}),\ x^p_{t_i} := [s_{t_i}; y^p_{t_i}]'$$

- The Jacobian is the vega of the option.

Calibration: Step 3

- Steps 1 and 2 are now repeated for all remaining times t_i and the log likelihood function is evaluated for the combination $(y^p, \mathbf{p})$.
- The value of $y^p_{t_{i-1}}$ is used to initialize the solver for the least absolute error problem.

Calibration: Step 4

- A new value of **p** is generated and Steps 1 and 3 are repeated until the likelihood function has been maximized by a numerical solver

$$\mathbf{p}^* \leftarrow \arg\min_{\mathbf{p}\in\mathbf{F}} -\ell_n(\mathbf{p}).$$

Geometric Brownian Motion

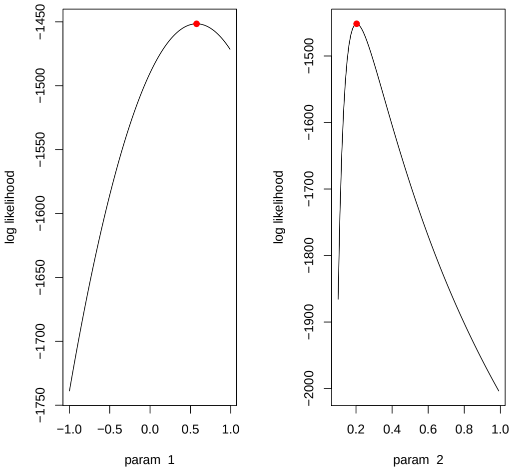

Figure: The marginal log likelihood function with respect to each parameter of the geometric brownian diffusion model. The mean is represented by Parameter 1 and the volatility by Parameter 2.

Geometric Brownian Motion

	μ	σ
Actual	0.5	0.2
Estimated	0.525	0.204
Est. Std. Error	6.415×10^{-2}	5.975×10^{-3}
Est. Huber Sandwich Error	6.420×10^{-2}	6.624×10^{-3}
Error in Est. Std. Error	5.169×10^{-13}	9.339×10^{-12}
Error in Est. Huber Sandwich Error	4.214×10^{-13}	7.931×10^{-12}

Table: The correct parameters, the estimates and the standard error estimates using 500 simulated stock prices.

Geometric Brownian Motion

Actual maximum log likelihood	1326.4464
Est. maximum log likelihood	1326.204
L2 Norm of Score Error	1.659×10^{-7}
L2 Norm of Hessian Error	1.448×10^{-7}
L2 Norm of Information matrix	1.962×10^{-6}

Table: This table lists the numerical error in a selection of estimated values.

Heston Model

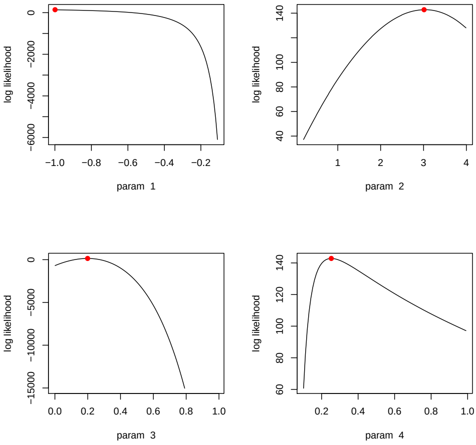

Figure: The marginal log likelihood function with respect to each parameter of the Heston model applied to the simulated underlying and prices. Parameters 1-4 represents ρ, κ, θ and σ.

Heston Model

	ρ	κ	θ	σ
Actual	-0.8	3.0	0.2	0.25
Estimated	-0.81	3.4	0.201	0.2476
Std. Error	6.702×10^{-3}	1.092	1.496×10^{-3}	4.687×10^{-2}
Huber Sandwich Error	4.903×10^{-4}	9.311×10^{-1}	1.117×10^{-2}	3.356×10^{-2}

Table: The correct parameters, the estimates and the lower bounds on the standard error using 50 simulated observations of the stock and the ATM option price.

Heston Model

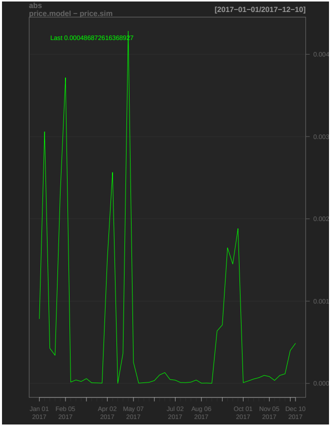

Figure: The absolute error in the calibrated model prices versus the simulated prices.

Heston Model

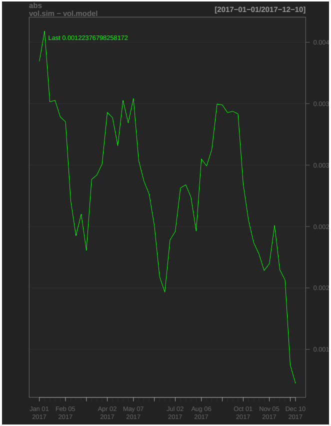

Figure: The absolute error in the implied volatilities versus the simulated volatilities.

MLEMVD Package Installation

```
library("devtools")
install_github("mfrdixon/MLEMVD")
```

MLEVMD Univariate Models

Model	μ	σ	constraints
U1	$x(a+bx)$	$\sigma x^{3/2}$	
U2	$a+bx$	dx	
U3	$b(a-x)$	cx^{d}	
U4	$\kappa(\alpha-x)$	$\sigma x^{1/2}$	
U5	$\sum_{i=0}^{3}\theta_i x^i$	γx^{ρ}	$\rho \geq 1$
U6	$a+bx+cx^2+dx^3$	f	
U7	$\kappa(\alpha-x)$	σ	
U8	$\frac{a_{-1}}{x}+a_0+a_1x+a_2x^2$	σx^{p}	$\rho \geq 1$
U9	$\frac{a_{-1}}{x}+a_0+a_1x+a_2x^2$	$(b_0+b_1x+b_2x^{b_3})^{1/2}$	
U10	$\frac{a_{-1}}{x}+a_0+a_1x+a_2x^2$	$b_0+b_1x+b_2x^{b_3}$	$\rho \geq 1$
U11	$a+bx$	$f+dx$	
U12	$\frac{\beta}{x}-\alpha x^3$	$\gamma x^{1/2}$	
U13	$\frac{a_{-1}}{x}+a_0+a_1x+a_2x^2+a_3x^3$	σx^{ρ}	$\rho \geq 1$

Table: The specification of various univariate diffusion models currently supported by the package.

MLEMVD Bivariate Models

Model	$\mu(x_1, x_2)$	$\Sigma(x_1, x_2)$	constraints
B1	$\begin{pmatrix} a + bx_2 \\ c + dx_2 \end{pmatrix}$	$\begin{pmatrix} \rho\sqrt{(x_2)} & 0 \\ h & \sqrt{(1-\rho^2)x_2} \end{pmatrix}$	
B2	$\begin{pmatrix} a_0 + a_1x_1 + a_2x_2 \\ b_0 + b_1x_1 + b_2x_2 \end{pmatrix}$	$\begin{pmatrix} c_0 + c_1x_1 + c_2x_2 & 0 \\ 0 & d_0 + d_1x_1 + d_2x_2 \end{pmatrix}$	
B3	$\begin{pmatrix} \mu - x_2/2 \\ \alpha + \beta x_2 \end{pmatrix}$	$\begin{pmatrix} \sqrt{x_2} & 0 \\ \sigma\rho x_2^\gamma & \sigma\sqrt{1-\rho^2}x_2^\gamma \end{pmatrix}$	
B4	$\begin{pmatrix} a_0 + a_1x_2 \\ b(a - x_2) + \lambda g x_2^\beta\sqrt{a + f(x_2 - a)} \end{pmatrix}$	$\begin{pmatrix} \sqrt{1-\rho^2}\sqrt{a + f(x_2 - a)} & \rho\sqrt{a + f(x_2 - a)} \\ 0 & gx_2^\beta \end{pmatrix}$	
B5	$\begin{pmatrix} bx_1 \\ c - dx_2 \end{pmatrix}$	$\begin{pmatrix} hx_1\sqrt{x_2} & 0 \\ g\rho\sqrt{x_2} & g\sqrt{1-\rho^2}\sqrt{x_2} \end{pmatrix}$	
B6	$\begin{pmatrix} m - x_2/2 \\ a - bx_2 \end{pmatrix}$	$\begin{pmatrix} \sqrt{x_2} & 0 \\ \sigma\sqrt{1-\rho^2}\sqrt{x_2} & \sigma\rho\sqrt{x_2} \end{pmatrix}$	$2a > \sigma^2$
B7	$\begin{pmatrix} 0 \\ a_1 - a_2x_2 \end{pmatrix}$	$\begin{pmatrix} \frac{2x_1}{\gamma\sqrt{x_2}} & \frac{2\eta x_1}{\gamma} \\ 2\sqrt{x_2} & 0 \end{pmatrix}$	
B8	$\begin{pmatrix} a + bx_1 \\ cx_2 \end{pmatrix}$	$\begin{pmatrix} dx_1^\gamma e^{x_2} & 0 \\ 0 & f \end{pmatrix}$	
B9	$\begin{pmatrix} a + bx_1 \\ cx_2 \end{pmatrix}$	$\begin{pmatrix} dx_1^\gamma e^{x_2} & 0 \\ 0 & f \end{pmatrix}$	
B10	$\begin{pmatrix} b_1(a_1 - x_1) \\ b_2(a_2 - x_2) \end{pmatrix}$	$\begin{pmatrix} g_1 & 0 \\ 0 & g_2\sqrt{x_2} \end{pmatrix}$	
B11	$\begin{pmatrix} k_1 + k_2x_2 \\ \kappa(\theta - x_2) \end{pmatrix}$	$\begin{pmatrix} \sqrt{1-\rho^2}\sqrt{x_2} & \rho\sqrt{x_2} \\ 0 & \sigma x_2 \end{pmatrix}$	
B12	$\begin{pmatrix} ax_1 \\ -bx_2 \end{pmatrix}$	$\begin{pmatrix} cx_1e^{x_2} & 0 \\ dr & d\sqrt{1-r^2} \end{pmatrix}$	
B13	$\begin{pmatrix} b_{11}(a_1 - x_1) + b_{12}(a_2 - x_2) \\ b_{21}(a_1 - x_1) + b_{22}(a_2 - x_2) \end{pmatrix}$	$\begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}$	
B14	$\begin{pmatrix} k_1(x_2 - x_1) \\ k_2(\theta - x_2) \end{pmatrix}$	$\begin{pmatrix} \sigma\sqrt{x_1} & 0 \\ 0 & \sigma_2\sqrt{x_2} \end{pmatrix}$	
B15	$\begin{pmatrix} a + bx_1 \\ fx_1 + dx_2 \end{pmatrix}$	$\begin{pmatrix} \sqrt{x_1} & 0 \\ h & \sqrt{1 + gx_1} \end{pmatrix}$	
B16	$\begin{pmatrix} a + bx_1 + gx_2 \\ d + \eta x_1 + fx_2 \end{pmatrix}$	$\begin{pmatrix} \sqrt{x_1} & 0 \\ h & \sqrt{x_2} \end{pmatrix}$	