MLEMV: A R Package for Maximum Likelihood Estimation of Multi-variate Diffusion Models

Matthew Dixon
Stuart School of Business
Illinois Institute of Technology
mdixon7@stuart.iit.edu

February 17, 2017

Abstract

Continuous-time Markov processes are typically defined by stochastic differential equations, describing the evolution of one or more state variables. Maximum likelihood estimation of the model parameters to historical observations is only possible when at least one of the state variables is observable. In these cases, the form of the transition function corresponding to the stochastic differential equations must be known to assess the efficacy of fitting a continuous model to discrete samples. This paper describes a R package `MLEVD` for calibrating multi-variate diffusions models.

1 Introduction

Continuous-time Markov processes are typically defined by stochastic differential equations, describing the evolution of one or more state variables. Maximum likelihood estimation of the model parameters to historical observations is only possible when at least one of the state variables is observable. In these cases, the form of the transition function corresponding to the stochastic differential equations must be known to assess the efficacy of fitting a continuous model to discrete samples.

[1] provide closed form expansions for the likelihood function of a general class of univariate diffusion models. The same author lated extended the approach to multi-variate diffusion models [2] and to, in particular, stochastic volatility models [3]. The later work describes an approach when only of the state variables is observed in financial times series and the other state variable is estimated from both the observed state variable and the corresponding at-the-money constant maturity option prices. The approach is applied to the Heston model [6], a model which has received considerable attention in the context of calibration owing to the many practical challenges and material defects.

[To do: describe other MLE approaches for Heston, e.g. [7]]

[8] address the problem of calibrating Heston's stochastic volatility model by providing guidance on how to calibrate the model to a chain of vanilla call option quotes at one instance of time.

[Tao: can you describe why calibrating to the surface of prices is important?] The authors draw attention to the fact that the calibration procedure is non-trivial – it is a non-linear programming problem with a non-linear constraint and non-convex objective function. Since multiple local-minima may exist, [8] propose using a combination of global search and local optimizers. The authors further note that the use of common stochastic algorithms for global search, such as simulated annealing, generally renders the calibration problem more computationally burdensome. The global optimizers that the authors consider include the differential evolution (DE) algorithm and simulated annealing (SA), both of which have been employed elsewhere in the quantitative finance literature [4].

The work of A"it-Sahalia provides a more rigorous alternative to calibrating by least squares, replacing a non-smooth, non-convex of non-concave objective function with a smooth convex or concave function. The calibration of the Heston model to at-the-money option prices is not without its own share of numerical stability challenges, in regions where one or more components of the Jacobian vanish. We provide a R package to compute the log likelihood functions of a general class of multi-variate diffusions and the proceed to perform a numerical study of the estimation of the Heston model parameters first applied to simulated option prices and then applied to high frequency time series data. Finally we comment on the extension of this approach to calibrating to the history of the option chain.

1.1 Overview of Package

[To do: Describe the purpose of this package.]

1.2 Maximum Likelihood Estimation

The principle of maximum likelihood estimation (MLE), originally developed by R.A. Fisher in the 1920s, states that the desired parametric probability distribution is the one that renders the observed data most probable. The maximum likelihood estimator (MLE) is the parameter vector value that maximizes the likelihood function.

Let i denote index observations whose values are x_i. Let $\theta \in R^p$ be a p parameter vector. Let $y \to f_i(x|\theta)$ be a smooth positive density. Let X_i be independent with density $f_i(\cdot|\theta)$ which are not independent.

The data is modeled as observed values of X_i for $i \in 1, 2, \ldots, n$. The likelihood function is

$$\mathcal{L}(\theta) = \sum_{i=0}^{n} log f_i(X_i|\theta). \tag{1}$$

The first and second partial derivatives of $\mathcal{L}$ with respect to θ are referred to as the score and the Hessian and are given by

$$\mathcal{D}(\theta) = \frac{\partial \mathcal{L}}{\partial(\theta)} \tag{2}$$

and

$$\mathcal{H}(\theta)_{ij} = \frac{\partial^2 \mathcal{L}}{\partial\theta_i \partial\theta_j}. \tag{3}$$

In the absence of model specification error, we first consider the curvature of the log likelihood function at the stationary point. A large curvature represents more confidence in the MLE and hence a lower standard error. The curvature is represented by the Information matrix - the negative of the expected value of the Hessian matrix:

$$[\mathcal{I}(\theta))] = -E[\mathcal{H}(\theta)] \tag{4}$$

The variance-covariance matrix of the parameter is

$$var(\theta) = [\mathcal{I}(\theta)]^{-1}. \tag{5}$$

The standard errors of the estimator are just the square roots of the diagonal terms in the variance-covariance matrix.

By the Cramer-Rao Theorem, under certain regularity conditions on the distribution, the variance of any unbiased estimator of a parameter θ must be at least as large as

$$var(\theta) \geq [-E[\mathcal{H}(\theta)]^{-1}. \tag{6}$$

This theorem implies that the maximum likelihood estimator is efficient but are our assumption is that the data is generated from the model is too strong.

1.3 Huber Sandwich Estimator

If the model is not well-specified but the mean function is correctly specified and the variance function is reasonably specified, then maximum likelihood is asymptotically normal with the following variance-covariance matrix

$$var(\hat{\theta}) = [\mathcal{I}(\hat{\theta})]^{-1} E[\mathcal{D}(\hat{\theta})\mathcal{D}(\hat{\theta})^T][\mathcal{I}(\hat{\theta}]^{-1}. \tag{7}$$

This is the variance-covariance matrix that provides whose square root of the diagonals provide the robust standard error estimates that are asymptotically correct even when the model is mis-specified. This is the maximum likelihood analogue of White?s consistent standard errors.

1.4 Diffusion Models

Following [1], consider the multivariate time-homogenous Markovian diffusion of the form

$$d\mathbf{X}_t = \mu(\mathbf{X}_t)dt + \Sigma(\mathbf{X}_t)d\mathbf{W}_t \tag{8}$$

where $\mathbf{X}_t, \mu \in R^m$, $\Sigma(\mathbf{X}_t) \in R^{m\times m}$ and $\mathbf{W}_t \in \mathbf{R}^m$ are independent Wiener processes.

Prior to the pioneering work of [1], the log of the transition function $f_X(x|x_0, \Delta)$ was only given in closed form under severe restrictions on the form of μ and Σ. We shall refer the variance-covariance matrix $v(x) := \Sigma\Sigma^T$. [1] constructs closed form expansions for the log-transition function for a large class of multivariate Markovian diffusions. The primary use of such closed form expansions is to permit the computation of the MLE rather than rely on less desirable approaches to inferring the log transition function numerically by solving a partial differential equation, simulating the process to Monte Carlo integrate the transition density or approximating the process with binomial trees.

We observe X 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:

$$l_n(\theta, \Delta) := \sum_{i=1}^{n} l_X(x_{i+1}|x_i, \Delta), \tag{9}$$

where the log of the transition density $l_X := lnf_X$. Under a Hermite expansion of l_X and application of a number of transformations, [1] eventually arrive at the following compact closed form expression with K terms.

$$l^(K)_X(x|x_0) = -\frac{m}{2}ln(2\pi\Delta) - D_v(x) + \frac{C_X^{-1}(x|x_0)}{\Delta} + \sum_{k=0}^{K} C_X^{(k)}(x|x_0)\frac{Delta^k}{k!}, \tag{10}$$

where

$$D_v := -\frac{1}{2}ln(Det[v(x)]). \tag{11}$$

1.5 Geometric Brownian motion

$$dX_t = \mu X_d t + \sigma X_t dW_t \tag{12}$$

$$f_X(x|x_0, t) = \frac{1}{\sqrt{2\pi}\sigma t} \exp - \frac{(lnX_t - lnX_0 - (\mu - \sigma^2/2)t)^2}{2\sigma^2 t} \tag{13}$$

$$l_n(\theta, \Delta) := \sum_{i=1}^{n} l_X(x_{i+1}|x_i, \Delta) = -\frac{1}{2}\sum_{i=1}^{n-1}(ln(2\pi\Delta\sigma^2 x_{i+1}^2) + (ln[X_{i+1}/X_i] - (\mu - \sigma^2/2)\Delta))^2/(\sigma^2\Delta) \tag{14}$$

1.6 Results

Maximum log likelihood	-1309.06362308552
Standard Error Estimate	0.0641498144626433
Standard Error Estimate	0.00597536567361609
Huber Sandwich Error Estimate	0.064200121510126
Huber Sandwich Error Estimate	0.00662406799506829
Exact Standard Error	0.0641498144626433
Exact Standard Error	0.00597536567361609
Exact Huber Sandwich Error	0.064200121510126
Exact Huber Sandwich Error	0.00662406799506829
Error in S.E. Estimate	5.16892952051595e-13
Error in H.S.E. Estimate	9.33926894153281e-12
L2 Norm of Score Error	1.65901651231863e-07
L2 Norm of Hessian Error	1.44866483253736e-0
L2 Norm of Information matrix	1.96150837439291e-06

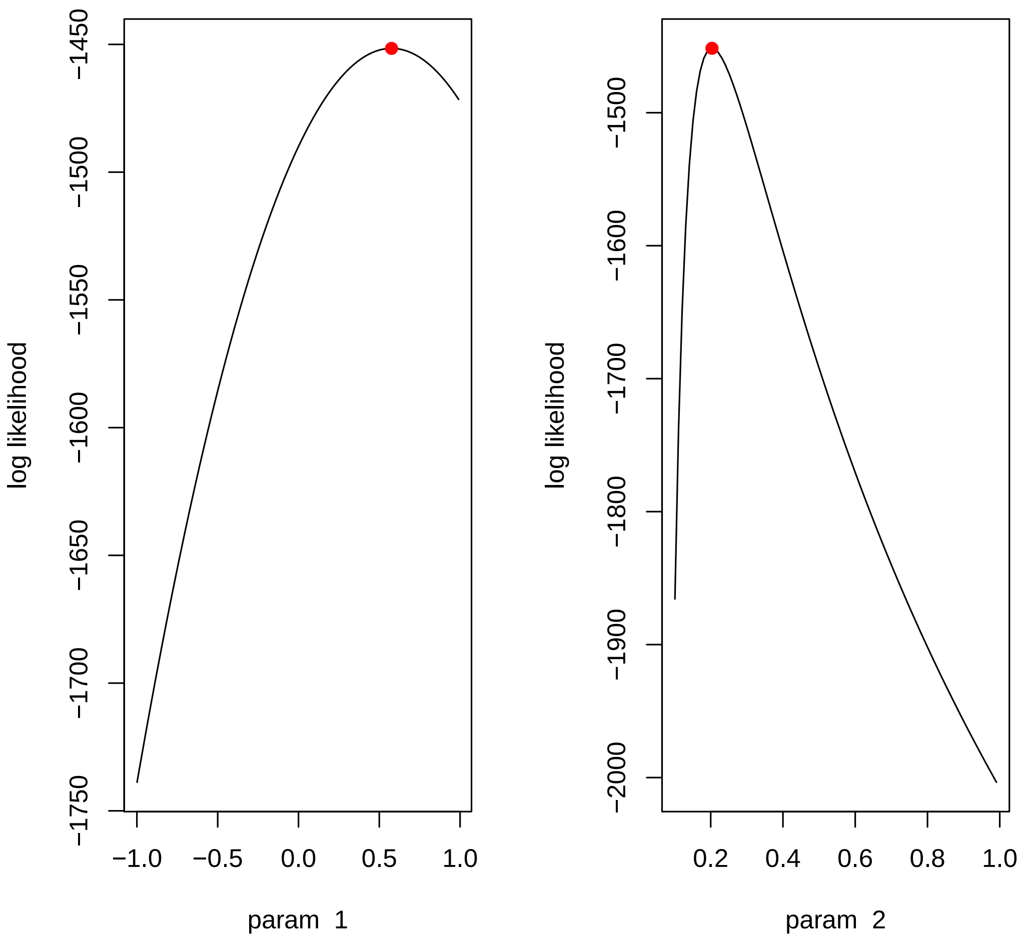

Figure 1: Example plot showing the marginal log likelihood function with respect to each parameter of the geometric brownian diffusion model.

1.7 Heston Model

Under the pricing measure Q, the Heston model describes the evolution of the log of stock price $s_t = ln\ S_t$ whose variance Y_t is given by a mean reverting square root process:

$$
\begin{aligned}
ds_t &= (a + bY_t)dt + \sqrt{Y_t} dW_1 Q(t), & (15)\\
dY_t &= \kappa'(\theta' - Y_t)dt + \sigma \sqrt{Y_t} dW_2^Q(t), & (16)
\end{aligned}
$$

where

$$a = r - d, \qquad b = -\frac{1}{2}, \tag{17}$$

A key characteristic of the model is that the Wiener processes are correlated $dW_1^Q \cdot dW_2^Q = \rho dt$. This feature enables the model to exhibit the '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
- v_0: initial variance

The parameter set $\mathbf{p} := [\kappa, \theta, \sigma, \rho]$ and the additional non-linear constraint (the Feller condition) $2\kappa\theta - \sigma^2 > 0$ is imposed during the calibration to ensure that Y_t is positive.

1.8 Likelihood function estimation

Given a set of observed underlying and ATM constant maturity option prices $g_t := [S_t; C_t]$ sampled at dates $t_0, t_1, \ldots, t_n$, the likelihood function takes the form:

$$l_n(\mathbf{p}) := \frac{1}{n} \sum_{i=1}^{n} l_G(\Delta t_i, g(t_i)|g(t_{i-1})); \mathbf{p}) \tag{18}$$

where

$$l_G(\Delta, g|g_0; \mathbf{p}) := ln f_G(\Delta, g|g_0; \mathbf{p}) = -ln J_t(\Delta, g|g_0; \mathbf{p}) + l_X(\Delta, f^{-1}(g; \mathbf{p})|f^{-1}(g_0; \mathbf{p}); \mathbf{p}) \tag{19}$$

and l_X denotes the likelihood function of the partially observed state vector $x_t := [ln\ S_t, Y_t]$ evaluated at each date $t_0, t_1, \ldots, t_n$. Here $\Delta t_i := t_i - t_{i-1}$ denotes the time step between observations. J_t denotes the Jacobian of the option price with respect to Y_t, which is equivalently to vega.

1.9 Pricing

With marginal loss of generality, we will restrict the scope of this section to European equity options. The Heston stochastic volatility model permits closed-form solutions for computing risk neutral European option prices. The price can be represented as a weighted sum of the delta of the European call option P_1 and P_2 - the probability that the asset price will exceed the strike price at maturity. Adopting standard option pricing notation, the call price of a vanilla European option is

$$C(S_0, K, \tau; \mathbf{z}_0) = S_0 P_1 - K e^{-(r-q)\tau} P_2, \tag{20}$$

P_1 and P_2 can be expressed as:

$$P_j = \frac{1}{2} + \frac{1}{\pi} \int_o^{\infty} \Re \left[\frac{\phi_j(S_0, \tau, u; \mathbf{z}_0) e^{-iu \ln K}}{iu} \right] du, j = 1, 2. \tag{21}$$

where ϕ_j are Heston analytic characteristic functions and are given in a convenient form in [**?**], and $\mathbf{z}_0$ is the vector of Heston model parameters. Following Fang and Oosterlee [5], the entire inverse Fourier integral in Equation (21) is reconstructed from Fourier-cosine series expansion of the integrand to give the following approximation of the call price

$$C(S_0, K, \tau; \mathbf{z}_0) \approx K e^{-r\tau} \Re \left[\sum_{k=0}^{N-1} \phi \left(\frac{k\pi}{b-a}; \mathbf{z}_0 \right) e^{ik\pi \frac{x-a}{b-a}} U_k \right], \tag{22}$$

where $x := \ln(S_0/K)$ and $\phi(w; \mathbf{z}_0)$ denotes the Heston characteristic function of the log-asset price, U_k the payoff series coefficients and N denotes the number of terms in the cosine series expansion (typically 128 will suffice).

1.10 Calibration

Given a sequence of observed underlying and corresponding near expiry, constant maturity, ATM options, we follow the these steps

- Initialize the unknown parameter vector to the model
- For each new parameter set $\mathbf{p}$ generated by the numerical optimization routine, compute the value of Y_t which satisfies the option price. Note this requires solving a nested one dimensional convex optimization, with linear bound constraints, so that $C_t \to \hat{Y}_t$.
- With $\hat{Y}_t$ and $\mathbf{p}$ compute ν of the option and thus the Jacobian term in Equation 19.
- Using $\hat{x}_t = [ln\ S_t, \hat{Y}_t]$ and $\mathbf{p}$ compute $l_X(\Delta, f^{-1}(g;\mathbf{p})|f^{-1}(g_0;\mathbf{p});\mathbf{p})$
- Minimize the log likelihood l_G over the parameters subject to the Feller condition.

We implemented the above Algorithm together with the Fourier-Cosine method in R and C++. The Heston price and vega are called by non-linear optimization R packages `NLoptr` and `DEOptim`[10, 9]. More specifically, we combine the `DEOptim` global optimizer with one of three constrained local optimization solvers provided in the `NLopt` package. These optimizers are (i) the Sequential Least SQuares Programming (SLSQP) method; (ii) the L-BFGS-B algorithm; and (iii) the Truncated Newton (TNC) method. Each method exploits the smoothness of the error function over the feasible region by approximating the Jacobian with first order forward differences under perturbations of each parameter. A small number of Hessian vectors are also computed at each main iteration in the L-BGFS-B algorithm. The NLoptr methods described above incorporate the non-linear inequality constraint to enforce the Feller condition.

The number of function evaluations per iteration is thus dependent on the number of model parameters. The global optimizer is terminated if either the objective function is below a threshold or the number of iterations exceeds a limit. The specifiable stopping criterion varies for each of the local optimizers. However, for ease of comparison of convergence properties between each, it is possible to terminate if either the absolute difference in function values between successive iterations is within a tolerance or the number of function evaluations exceeds a limit. In practice, a tolerance on the absolute difference of the function value is neither intuitive or ideally suited to calibration. In further experiments, not reported here, we find that specifying the tolerance on the norm of the difference in solution iterates leads to more stable parameters over successive calibrations. Of the three aforementioned local solvers, only the TNC method permits a tolerance of this form.

1.11 Results

[To do: add results calibrating to intra-day atm historical prices]

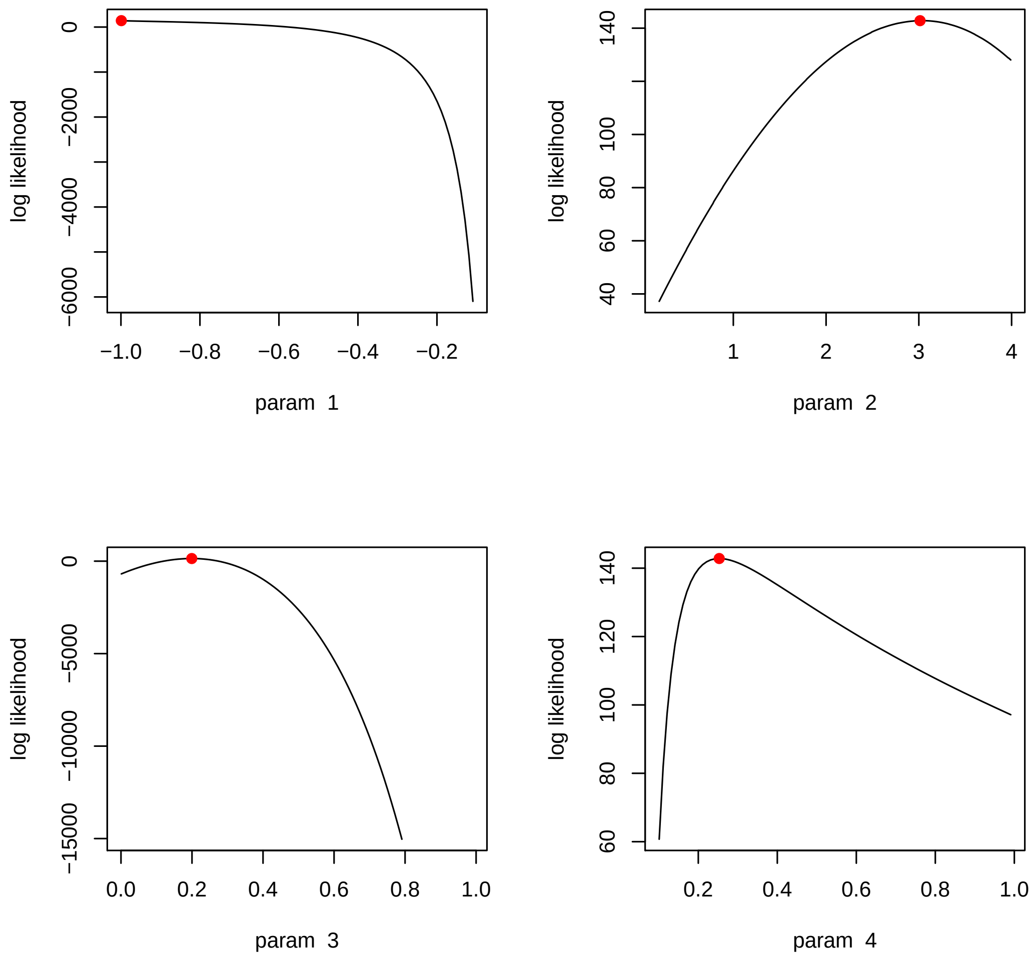

Figure 2: Example plot showing the marginal log likelihood function with respect to each parameter of the Heston model applied to the simulated underlying and prices.

A Model Reference

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	$\sum_{i=0}^{3} \theta_i x^0$	γx^ρ	$\rho \geq 1$
U6	$a+bx+cx^2+dx^3$	f	
U7	$\kappa(\alpha - x)$	σ	
U8	$\frac{a_{-1}}{x} + a_0 + a_1 x + a_2 x^2$	σx^p	$\rho \geq 1$
U9	$\frac{a_{-1}}{x} + a_0 + a_1 x + a_2 x^2$	$(b_0 + b_1 x + b_2 x^{b_3})^{1/2}$	
U10	$\frac{a_{-1}}{x} + a_0 + a_1 x + a_2 x^2$	$b_0 + b_1 x + b_2 x^{b_3}$	$\rho \geq 1$
U11	$a+bx$	$f+dx$	
U12	$\frac{\beta}{x} - \alpha x^3$	$\gamma x^{1/2}$	
U13	$\frac{a_{-1}}{x} + a_0 + a_1 x + a_2 x^2 + a_3 x^3$	σx^ρ	$\rho \geq 1$

Model	$\mu(x_1, x_2)$	$\Sigma(x_1, x_2)$
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	$\begin{pmatrix} m - x_2/2 \\ a - bx_2 \end{pmatrix}$	$\begin{pmatrix} \sqrt{x_2} & 0 \\ \sigma\sqrt{1-\rho^2}\sqrt{x_2} & \sigma\rho\sqrt{x_2} \end{pmatrix}$
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	$\begin{pmatrix} a + bx_1 + gx_2 \\ d + \eta x_1 + fx_2 \end{pmatrix}$	$\begin{pmatrix} \sqrt{x_1} & 0 \\ h & \sqrt{x_2} \end{pmatrix}$
B17	$\begin{pmatrix} a_{00} - (a_1 + a_2x_2)/2 + (n_0\sqrt{1-g_1^2} + nu_1g_1)(\sqrt{a_1 + a_2x_2}^{b+d} \\ a_{01} + a_{11}x_2 + (nu_1g_{11})(\sqrt{a_1 + a_2x_2}^{b+d} \end{pmatrix}$	$\begin{pmatrix} \sqrt{1-g_1^2}\sqrt{a_1 + a_2x_2} & g_1\sqrt{a_1 + a_2x_2} \\ 0 & g_{11}(\sqrt{a_1 + a_2x_2})^b \end{pmatrix}$
B18	$\begin{pmatrix} b_1x_1 \\ a_2 + b_2x_2 \end{pmatrix}$	$\begin{pmatrix} g_{11}e^{x_1} & 0 \\ g_{22}r & g_{22}\sqrt{1-r^2} \end{pmatrix}$
B19	$\begin{pmatrix} b_1x_1 \\ a_2 + b_2x_2 \end{pmatrix}$	$\begin{pmatrix} e^{x_2} & 0 \\ g_{22}r & g_{22}\sqrt{1-r^2} \end{pmatrix}$
B20	$\begin{pmatrix} a_1 + b_1x_1 \\ a_2 + b_2x_2 \end{pmatrix}$	$\begin{pmatrix} \sqrt{x_2} & 0 \\ gr\sqrt{x_2} & g\sqrt{1-r^2}\sqrt{x_2} \end{pmatrix}$
B21	$\begin{pmatrix} a_1(b_1 - x_1) \\ a_{21}(b_1 - x_1) + a_2(b_2 - x_2) \end{pmatrix}$	$\begin{pmatrix} \sqrt{x_1} & 0 \\ g_{21}\sqrt{x_1} & g_{22}\sqrt{x_1} \end{pmatrix}$
B22	$\begin{pmatrix} k_1 + k_2x_2 \\ k(a - x_2) \end{pmatrix}$	$\begin{pmatrix} \sqrt{1-r^2}\sqrt{x_2} & r\sqrt{x_2} \\ 0 & sx_2^b \end{pmatrix}$

References

[1] Y. Ait-Sahalia. Maximum likelihood estimation of discretely sampled diffusions: A closed-form approximation approach. *Econometrica*, 70(1):223–262, 2002.

[2] Y. Ait-Sahalia. Closed-form likelihood expansions for multivariate diffusions. *The Annals of Statistics*, 36(2):906–937, 2008.

[3] Y. Ait-Sahalia and R. Kimmel. Maximum likelihood estimation of stochastic volatility models. *Journal of Financial Economics*, 83(2):413 – 452, 2007.

[4] D. Ardia, J. David, O. Arango, and N. Gomez. Jump-Diffusion Calibration using Differential Evolution. *Wilmott Magazine*, 55:76–79, Sept. 2011.

[5] F. Fang and C. W. Oosterlee. A Novel Pricing Method for European Options based on Fourier-Cosine Series Expansions. *SIAM Journal on Scientific Computing*, 31:826–848, 2008.

[6] S. Heston. A Closed-form Solution for Options with Stochastic Volatility. *Review of Financial Studies*, 6:327–343, 1993.

[7] F. Mariani, G. Pacelli, and F. Zirilli. Maximum likelihood estimation of the heston stochastic volatility model using asset and option prices: an application of nonlinear filtering theory. *Optimization Letters*, 2(2):177–222, 2008.

[8] S. Mikhailov and U. Nögel. Heston's Stochastic Volatility Model Implementation, calibration and some extensions. *Wilmott Magazine*, 4:74–79, July 2003.

[9] K. Mullen, D. Ardia, D. Gil, D. Windover, and J. Cline. DEoptim: An R package for global optimization by differential evolution. *Journal of Statistical Software*, 40(6):1–26, 2011.

[10] K. V. Price, R. M. Storn, and J. A. Lampinen. *Differential Evolution - A Practical Approach to Global Optimization*. Natural Computing. Springer-Verlag, January 2006. ISBN 540209506.