Maximum likelihood estimation of mean reverting processes

José Carlos García Franco
Onward, Inc.
jcpollo@onwardinc.com

Abstract

Mean reverting processes are frequently used models in real options. For instance, some commodity prices (or their logarithms) are frequently believed to revert to some level associated with marginal production costs. Fundamental parameter knowledge, based on economic analysis of the forces at play, is perhaps the most meaningful channel for model calibration. Nevertheless, data based methods are often needed to complement and validate a particular model. The Ornstein-Uhlenbeck mean reverting (OUMR) model is a Gaussian model well suited for maximum likelihood (ML) methods. Alternative methods include least squares (LS) regression of discrete autoregressive versions of the OUMR model and methods of moments (MM). Each method has advantages and disadvantages. For instance, LS methods may not always yield a reasonable parameter set (see Chapter 3 of Dixit and Pindyck[2]) and methods of moments lack the desirable optimality properties of ML or LS estimation .

This note develops a maximum-likelihood (ML) methodology for parameter estimation of 1-dimensional Ornstein-Uhlenbeck (OR) mean reverting processes. Furthermore, our methodology ultimately relies on a one-dimensional search which greatly facilitates estimation and easily accommodtes missing or unevenly spaced (time-wise) observations.

The simple Ornstein-Uhlenbeck mean reverting (OUMR) process given by the stochastic differential equation (SDE)

$$d\,x(t) = \eta(\bar{x} - x(t))\,dt + \sigma\,dB(t); \qquad x(0) = x_0 \tag{1}$$

for constants $\bar{x}$, η and x_0 and where $B(t)$ is standard Brownian motion.

In this model the process $x(t)$ fluctuates randomly, but tends to revert to some fundamental level $\bar{x}$. The behavior of this 'reversion' depends on both the short term standard deviation σ and the speed of reversion parameter η.

An example. Figure 1 shows a sample path for 120 months of a mean reverting process starting at a level $x(0) = 12$, that tends to revert to a level $\bar{x} = 15$, with a speed of reversion $\eta = 4$ and a short term standard deviation $\sigma = 5$ (one third of the level of reversion). The solid line shows the level of reversion. One characteristic that may be evident is that, as opposed to random walks (with drift), the process does not exhibit an explosive behavior, but rather tends to fluctuate around the reversion level. Furthermore, it may be shown that the long-term variance of the process has a limit. This behavior is often desirable for the analysis of economic variables that have a fundamental reason to fluctuate around a given level. For example, the price of some commodities or the marginal cost curve for the production of some good. However, fitting or calibration of such models is not easy to come by. While all the parameters may have some intuitive meaning to the analyst, measuring them is quite another story. In the best of cases there is some fundamental knowledge that leads to fixing a parameter, this is hopefully the case for the reversion level $\bar{x}$, yet, it is unlikely to have expert knowledge of all parameters and we are forced to rely on data driven

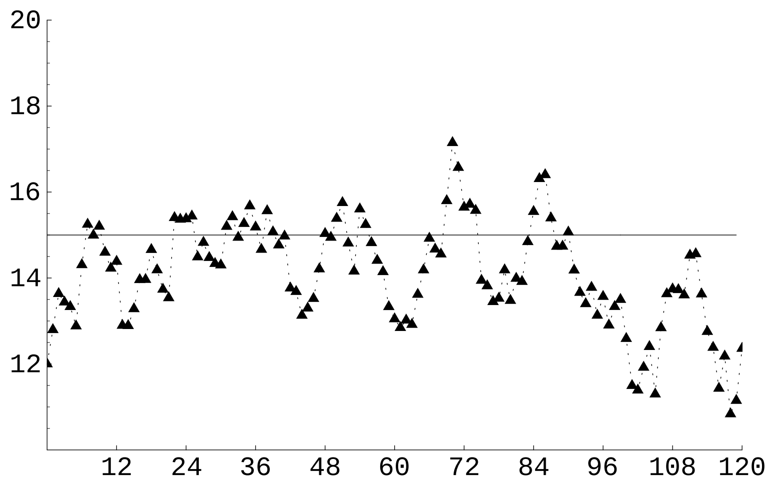

Figure 1: *OUMR sample path.*

estimation methods. Assuming of course that such data is available.

We will illustrate a maximum likelihood (ML) estimation procedure for finding the parameters of the mean-reverting process. However, in order to do this, we must first determine the distribution of the process $x(t)$. The process $x(t)$ is a gaussian process which is well suited for maximum likelihood estimation. In the section that follows we will derive the distribution of $x(t)$ by solving the SDE (1).

1 The distribution of the OR process

The OU mean reverting model described in (1) is a gaussian model in the sense that, given X_0, the time t value of the process $X(t)$ is normally distributed with

$$\mathrm{E}[x(t)|x_0] = \bar{x} + (x_0 - \bar{x})\exp[-\eta\, t] \quad \text{and} \quad \mathrm{Var}[x(t)|x_0] = \frac{\sigma^2}{2\,\eta}\,(1 - \exp[-2\eta\, t])\,.$$

Appendix A explains this based on the solution of the SDE (1).

Figure 2 shows a forecast in the form of 10-50-90 confidence intervals corresponding to the process we previously used as an example. The fact that long-term variance tends to a constant, is demonstrated by the flatness of the confidence intervals as we forecast farther into the future. We can also see from this forecast that the long-term expected value (which is equal to the median in the OU mean reverting model) tends to the level of reversion.

2 Maximum likelihood estimation

For $t_{i-1} < t_i$, the $x_{t_{i-1}}$ conditional density f_i of x_{t_i} is given by

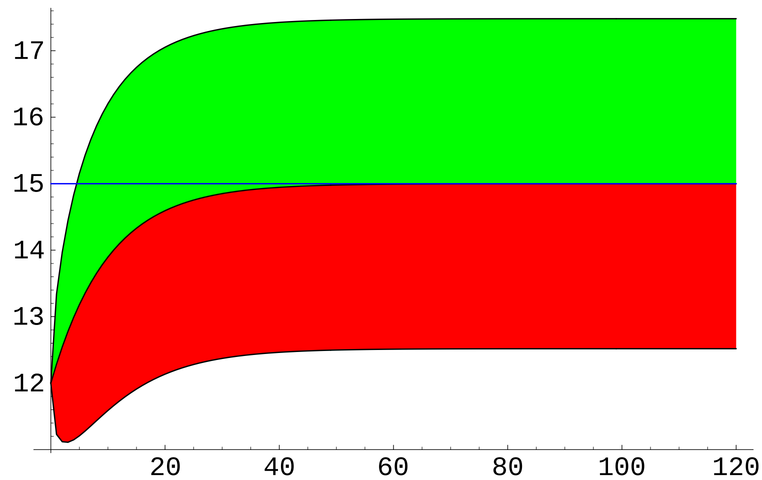

Figure 2: *OUMR 10-50-90 confidence interval forecast.*

$$f_i(x_{t_i};\bar{x},\eta,\sigma) =$$

$$(2\,\pi)^{-\frac{1}{2}}\left(\frac{\sigma^2}{2\eta}\left(1-e^{-2\eta(t_i-t_{i-1})}\right)\right)^{-\frac{1}{2}}\times \tag{2}$$

$$\exp\left[-\frac{\left(x_{t_i}-\bar{x}-(x_{t_{i-1}}-\bar{x})e^{-\eta(t_i-t_{i-1})}\right)^2}{2\,\frac{\sigma^2}{2\eta}\left(1-e^{-2\,\eta\,(t_i-t_{i-1})}\right)}\right]. \tag{3}$$

Given $n+1$ observations $\mathbf{x}=\{x_{t_0},\ldots,x_{t_n}\}$ of the process x, the log-likelihood function[1] corresponding to (**??**) is given by

$$\begin{aligned}\mathcal{L}(\mathbf{x};\bar{x},\eta,\sigma) &= -\frac{n}{2}\log\left[\frac{\sigma^2}{2\eta}\right]-\frac{1}{2}\sum_{i=1}^{n}\log\left[1-e^{-2\,\eta(t_i-t_{i-1})}\right]\\ &\quad -\frac{\eta}{\sigma^2}\sum_{i=1}^{n}\frac{\left(x_{t_i}-\bar{x}-(x_{t_{i-1}}-\bar{x})\,e^{-\eta(t_i-t_{i-1})}\right)^2}{1-e^{-2\eta(t_i-t_{i-1})}}.\end{aligned} \tag{4}$$

The maximum likelihood estimates (MLE) $\hat{\bar{x}}$, $\hat{\eta}$ and $\hat{\sigma}$ maximize the log-likelihood function and can be found by Quasi-Newton optimization methods. Another alternative is to rely on the first order conditions which requires the solution of a non-linear system of equations. Quasi-Newton methods are also applicable in this case. However, optimization or equation solving techniques require considerable computation time. In the sections that follow we will attempt to obtain an analytic alternative for ML estimation, based on the first order conditions. This approach is based on the approach found in Barz [1]. However, we do allow for arbitrarily spaced observations (timewise) and avoid some simplifying assumptions made in that work (for practical purposes).

2.1 First order conditions

The first order conditions for maximum likelihood estimation require the gradient of the log-likelihood to be equal to zero. In other words, the maximum likelihood estimators $\hat{\bar{x}}$, $\hat{\eta}$ and $\hat{\sigma}$ satisfy the first order conditions:

[1] With constant terms omitted.

$$\left.\frac{\partial\mathcal{L}(\mathbf{x};\bar{x},\eta,\sigma)}{\partial\bar{x}}\right|_{\hat{\bar{x}}} = 0$$
$$\left.\frac{\partial\mathcal{L}(\mathbf{x};\bar{x},\eta,\sigma)}{\partial\eta}\right|_{\hat{\eta}} = 0$$
$$\left.\frac{\partial\mathcal{L}(\mathbf{x};\bar{x},\eta,\sigma)}{\partial\sigma}\right|_{\hat{\sigma}} = 0$$

The solution to this non-linear system of equations may be found using a variety of numerical methods. However, in the next section we will illustrate an approach that simplifies the numerical search by exploiting some convenient analytical manipulations of the first order conditions.

2.2 A hybrid approach

We first turn our attention to the first element of the gradient. We have that

$$\frac{\partial\mathcal{L}(\mathbf{x};\bar{x},\eta,\sigma)}{\partial\bar{x}} = -\frac{\eta}{\sigma^2}\sum_{i=1}^{n}\frac{x_{t_i}-\bar{x}-\left(x_{t_{i-1}}-\bar{x}\right)\,e^{-\eta(t_i-t_{i-1})}}{1+e^{-\eta(t_i-t_{i-1})}}$$

Under the assumption that η and σ are non-zero, the first order conditions imply

$$\hat{\bar{x}} = f(\hat{\eta}) = \sum_{i=1}^{n}\frac{x_{t_i}-x_{t_{i-1}}\,e^{-\hat{\eta}(t_i-t_{i-1})}}{1+e^{-\hat{\eta}(t_i-t_{i-1})}}\left(\sum_{i=1}^{n}\frac{1-e^{-\hat{\eta}(t_i-t_{i-1})}}{1+e^{-\hat{\eta}(t_i-t_{i-1})}}\right)^{-1}. \tag{5}$$

The derivative of the log-likelihood function with respect to σ is

$$\frac{\partial\mathcal{L}(\mathbf{x};\bar{x},\eta,\sigma)}{\partial\sigma} = -\frac{n}{\sigma}+\frac{2\eta}{\sigma^3}\sum_{i=1}^{n}\frac{\left(x_{t_i}-\bar{x}-(x_{t_{i-1}}-\bar{x})\,e^{-\eta(t_i-t_{i-1})}\right)^2}{1-e^{-2\eta(t_i-t_{i-1})}},$$

which together with the first order conditions implies

$$\hat{\sigma} = g(\hat{\bar{x}},\hat{\eta}) = \sqrt{\frac{2\hat{\eta}}{n}\sum_{i=1}^{n}\frac{\left(x_{t_i}-\hat{\bar{x}}-(x_{t_{i-1}}-\hat{\bar{x}})\,e^{-\hat{\eta}(t_i-t_{i-1})}\right)^2}{1-e^{-2\hat{\eta}(t_i-t_{i-1})}}}. \tag{6}$$

Expressions (5) and (6) define functions that relate the maximum likelihood estimates. Specifically we have $\hat{\bar{x}}$ as a function f of $\hat{\eta}$ and $\hat{\sigma}$ as a function g of $\hat{\eta}$ and $\hat{\bar{x}}$.

In order to solve for the maximum likelihood estimates, we could solve the system of non-linear equations given by $\hat{\bar{x}} = f(\hat{\eta})$, $\hat{\sigma} = g(\hat{\bar{x}},\hat{\eta})$ and the first order condition $\partial\mathcal{L}(\mathbf{x};\bar{x},\eta,\sigma)/\partial\sigma|_{\hat{\sigma}} = 0$. However, the expression for $\partial\mathcal{L}(\mathbf{x};\bar{x},\eta,\sigma)/\partial\sigma$ is algebraically complex and would not lead to a closed form solution, requiring a numerical solution.

A simpler approach is to substitute the functions $\hat{\bar{x}} = f(\hat{\eta})$ and $\hat{\sigma} = g(\hat{\bar{x}},\hat{\eta})$ directly into the likelihood function and maximize with respect to η. So our problem becomes

$$\min_{\eta} V(\eta) \tag{7}$$

where

$$\begin{aligned} V(\eta) &= -\frac{n}{2}\log\left[\frac{g(f(\eta),\eta)^2}{2\eta}\right]-\frac{1}{2}\sum_{i=1}^{n}\log\left[1-e^{-2\,\eta(t_i-t_{i-1})}\right] \\ &\quad -\frac{\eta}{g(f(\eta),\eta)^2}\sum_{i=1}^{n}\frac{\left(x_{t_i}-f(\eta)-(x_{t_{i-1}}-f(\eta))\,e^{-\eta(t_i-t_{i-1})}\right)^2}{1-e^{-2\eta(t_i-t_{i-1})}}. \end{aligned}$$

Parameter $\eta = 1.2$				
Number of samples	Mean bias	Bias standard deviation	Mean relative bias	Relative bias std. dev.
104	2.381	2.519	198.4%	209.9%
260	0.962	1.145	80.2%	95.5%
520	0.447	0.675	37.3%	56.2%
1040	0.223	0.426	18.6%	35.5%
2080	0.102	0.290	8.5%	24.1%

Parameter $\bar{x} = 16$				
Number of samples	Mean bias	Bias standard deviation	Mean relative bias	Relative bias std. dev.
104	-0.119	4.654	-9.9%	29.1%
260	-0.121	2.055	-10.1%	12.8%
520	-0.042	1.081	-3.5%	6.8%
1040	-0.006	0.722	-0.5%	4.5%
2080	0.005	0.539	-0.4%	3.4%

Parameter $\sigma = 4$				
Number of samples	Mean bias	Bias standard deviation	Mean relative bias	Relative bias std. dev.
104	0.039	0.289	3.2%	7.2%
260	0.017	0.182	1.4%	4.5%
520	0.013	0.125	1.1%	3.1%
1040	0.004	0.089	0.3%	2.2%
2080	0.004	0.062	0.3%	1.6%

Table 1: An example of MLE performance.

It is not hard to show that the solution to the problem (7) yields the maximum likelihood estimator $\hat{\eta}$. Once we have obtained $\hat{\eta}$ we can easily find $\hat{\bar{x}} = f(\hat{\eta})$ and $\hat{\sigma} = g(\hat{\bar{x}}, \hat{\eta})$. The advantage of this approach is that the problem (7) requires a one dimensional search and requires the evaluation of less complex expressions than solving for all three first order conditions.

3 Example

Consider a family of weekly observations (samples) from an Ornstein-Uhlenbeck mean reverting process with parameters $\bar{x} = 16$, $\eta = 1.2$ and $\sigma = 4$ starting at $X(0) = 12$. It is known (1) that the MLE's converge to the true parameter as the sample size increases and (2) that the MLE's are asymptotically normally distributed. However, in practice we do not enjoy the convergence benefits given by the MLE large sample properties. In order to get an idea of how the MLE's behave under different sample sizes, a simulation experiment was conducted where we estimated the mean and variance of the estimation bias.

Table 3 summarizes the simulation results. From these results we can begin to appreciate the accuracy of the method as well as the asymptotic behavior of the maximum likelihood estimation.

4 Conclusion

In this note we developed a practical solution for the maximum likelihood estimation of an Ornstein-Uhlenbeck mean reverting process. The main advantage of our approach is that by leveraging on some manipulation of the first order conditions, we can reduce ML estimation to a one dimensional optimization problem which can generally be solved in a matter of seconds. The reduction of the problem to one dimension also facilitates the localization of a global maximum for the likelihood function. In addition, it is worth mentioning that the method rivals alternative methods such as regression of a discrete version of the OU mean reverting model or moment matching methods. Finally, we note that the method presented in this note trivially accommodates fundamental knowledge of any of the process parameters by simply substituting the known parameter(s) into the corresponding equations. For instance, if $\bar{x}$ is known, we forget about the function $\hat{\bar{x}} = f(\hat{\eta})$ and simply plug in the known $\bar{x}$ into the other equations as the MLE $\hat{\bar{x}}$.

References

[1] Barz, G. (1999) *Stochastic Financial Models for Electricity Derivatives*, Ph.D. dissertation, Department of Engineering-Economic Systems and Operations Research, Stanford University, Stanford, CA.

[2] Dixit, A.K. and Pindyck R.S. (1994) *Investment Under Uncertainty*, Princeton University Press, Princeton, NJ.

[3] Greene, William H. (1997) *Econometric Analysis*, 3rd. Edition, Prentice Hall, Upper Saddle River, NJ.

[4] Luenberger, D.G. (1998) *Investment Science*, Oxford University Press, New York, NY.

[5] Øksendal, B. (1995) *Stochastic Differential Equations: An Introduction with Applications*, 4th ed., Springer-Verlag, New York, NY.

A Solving the Ornstein-Uhlenbeck SDE

Consider a mean reverting Ornstein-Uhlenbeck process which is described by the following stochastic differential equation (SDE)

$$d\,x(t) = \eta\,(\bar{x} - x(t))\,dt + \sigma\,d\,B(t); \quad x(0) = x_0 \tag{8}$$

The solution of the OR SDE is standard in the literature. First note that

$$d\left(e^{\eta t}\,x(t)\right) = x(t)\,\eta e^{\eta t}\,dt + e^{\eta t}\,dx(t),$$

and therefore we have

$$e^{\eta t}\,dx(t) = d\left(e^{\eta t}x(t)\right) - x(t)\,\eta e^{\eta t}\,dt. \tag{9}$$

Multiplying both sides of (8) by $e^{\eta t}$, we get

$$e^{\eta t}dx(t) = e^{\eta t}\,\eta(\bar{x} - x(t))\,dt + e^{\eta t}\sigma\,dB(t), \tag{10}$$

which together with (9) implies

$$d\left(e^{\eta t}x(t)\right) = \eta e^{\eta t}\bar{x}\,dt + e^{\eta t}\sigma\,dB(t). \tag{11}$$

Therefore, we can now solve for (11) as

$$e^{\eta t}x(t) = x_0 + \int_0^t \eta e^{\eta s}\bar{x}\,ds + \int_0^t e^{\eta s}\sigma\,dB_s, \tag{12}$$

or equivalently

$$x(t) = x_0\,e^{-\eta\,t} + \int_0^t \eta\,e^{-\eta(t-s)}\,\bar{x}\,ds + \int_0^t e^{-\eta(t-s)}\sigma\,dB_s. \tag{13}$$

The first integral on the right hand side evaluates to $\bar{x}(1 - e^{-\eta t})$ and since B_t is Brownian motion, the second integral is normally distributed with mean zero and variance $\mathrm{E}\left[\left(\int_0^t e^{-\eta(t-s)}\sigma\,dB_s\right)^2\right]$. By Ito isometry[2] we have

$$\begin{aligned}
\mathrm{E}\left[\left(\int_0^t e^{-\eta(t-s)}\sigma\,dB_s\right)^2\right] &= \int_0^t \left(e^{-\eta(t-s)}\sigma\right)^2\,ds \\
&= \int_0^t e^{-2\eta(t-s)}\sigma^2\,ds \\
&= \frac{\sigma^2}{2\eta}\left(1 - e^{-2\eta t}\right).
\end{aligned}$$

Hence, X_t is normally distributed with $\mathrm{E}[X_t|X_0] = \bar{x} + (X_0 - \bar{x})e^{-\eta t}$ and $\mathrm{Var}[X_t|X_0] = \frac{\sigma^2}{2\eta}\left(1 - e^{-2\eta t}\right)$.

[2]See Øksendal [5] for details.