Uppsala universitet, 2015
1TD184 Optimisation
Daniel Noreland

Assignment 1 – basic option
2015-10-26

Assignment 1: Black-box optimisation–parameter estimation in financial mathematics

The following assignment is a highly practical exercise in black-box optimisation. From an existing (more or less obscure) simulation code, an optimisation problem is formulated and solved using library software. The problem and the simulation code have been contributed by Josef Höök at the Division of Scientific Computing.

Background

Financial mathematics deals with the determination of the most likely development of a stochastic process in order to suggest financial actions. The market value of financial assets vary in time more or less erratically, but a financial actor who has exclusive information about probable future changes in the market value is in a position where he or she can speculate in the asset. An example of such an asset are the treasury bills emitted by the Swedish National Debt Office. Fig. 1 shows the variation of the treasury bill interest rate over a period of three years. It is a stochastic process, which is not totally random (white noise) but there is some systematic in the variation. A stochastic model for interest rates that can capture the some of the underlying variations is the so-called *Cox–Ingersoll–Ross* model:

$$dX = a(b - x) + \sigma\sqrt{x}dW. \tag{1}$$

This is a *stochastic differential equation*, where dX is the change in the interest rate x over a short time interval dt. a, b and σ are model parameters and dW is a random variable known as the Wiener process. Eq. (1) can be simulated numerically using a forward Euler type scheme,

$$x(t_{i+1}) \approx x_{i+1} = x_i + a(b - x_i)\Delta t + \sigma\sqrt{x_i}\sqrt{\Delta t}\xi, \tag{2}$$

where Δt is the time step and $\xi \in N(0, 1)$ is a normally distributed random variable. Starting from some initial condition (t_0, x_0), integration of Eq. (2) results in one *realisation* of the process. Running the simulation again results in a different solution owing to the appearance of the stochastic term.

If one were to simulate the process a large number of times, a set of curves like in Fig. 2 would be the result. For any time t_i, it would be possible to draw a histogram over the corresponding values x_i that in the limit would approach the probability density function $f(x_i)$ for the interest rate at that time. Before an accurate set of density functions (one for each t_i) were obtained, a very high

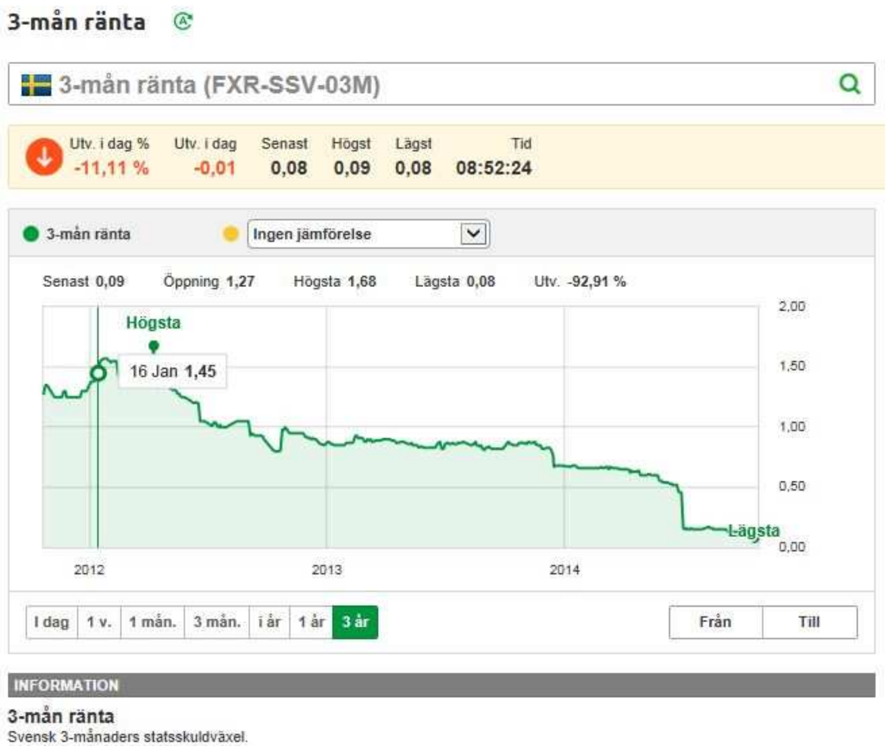

Figure 1: Swedish treasury bill rate over a three-year period.

number of simulations would be needed. Fortunately, it is possible to derive a partial differential equation directly in the probably density function f,

$$\frac{\partial f(x,t)}{\partial t} = \frac{\partial}{\partial x}(a(b-x)f(x,t)) + \frac{1}{2}\sigma^2 \frac{\partial^2}{\partial x^2}(xf(x,t)),$$

whose solution $f(x,t)$ gives the probability density for interest rate x at time t given the initial condition $f(x,t_0) = \delta(x-x_0)$[1] where x_0 is a (observed) starting point. It is a Fokker–Planck diffusion equation for the probability density of a type that appears also in statistical mechanics. The Dirac distribution at the beginning simply states that we have the interest rate x_0 at t_0 with 100% certainty, which is the case when we have made an observation.

Knowing the values of a, b and σ one would, insofar as Eq. (1) is representative of reality, be able to predict the probability of different interest rate scenarios in the future. The goal of this assignment is to identify the best set of parameters values by confrontation of the model with historical interest rate data.

The Maximum likelihood approach

The solution to the Fokker-Planck depends on the three parameters a, b and σ in the model. We have one set of historical data, namely the interest rate over three years, and we want to estimate the parameters for which the model mimics the data as closely as possible. So how do we interpret correspondence

[1] $\delta(x-x_0)$ denotes the *Dirac measure*, which has the property that $\delta(x) = 0, x \neq 0$, $\int_{-\infty}^{\infty} \delta(x)dx = 1$.

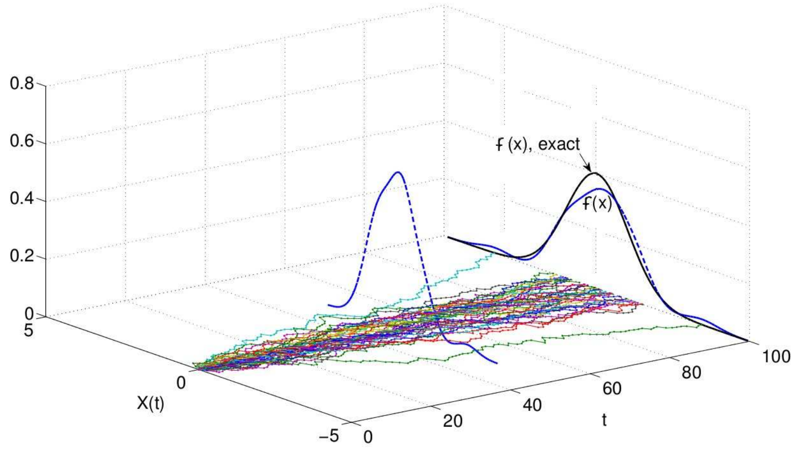

Figure 2: Multiple realisations of the Cox–Ingersoll–Ross model.

for a stochastic process? The Maximum likelihood (ML) approach is based on the idea that the "best" parameter set is the one that maximises the probability of observing the given data. For a single time t_i, this would correspond to the peak in the probability density function. When there are many data points, the correspondence is interpreted in the least squares sense.

Parameter estimation for the interest rate model

At each time t_{i-1}, the rate $x(t_{i-1})$ is observed. Counting from here, the rate at t_i is a stochastic variable that depends only on the state at t_{i-1}–it does not matter what happened earlier, since the process has no "memory". Let $f(x_i, t_i; x_{i-1}, t_{i-1}; a, b, \sigma)$ denote the probability density function at time t_i given the *observed* value x_{i-1} at t_{i-1}. The likelihood function

$$\mathcal{L}(a, b, \sigma) = \prod_{i=1}^{n} f(x_i, t_i; x_{i-1}, t_{i-1}; a, b, \sigma)$$

can then be defined. The ML interpretation of best fit is when $\mathcal{L}$ has its maximum. In order not to have to deal with a product (probably with a very small value), one can equally well maximise the function

$$\ell(a, b, \sigma) = \ln \mathcal{L} = \sum_{i=1}^{n} \ln f(x_i, t_i; x_{i-1}, t_{i-1}; a, b, \sigma) \tag{3}$$

What we will deal with in the following is to solve the parameter identification problem using a supplied data file `CIRDataSet` with a three-year time series for the interest rate along with an existing function `cirpdf.m` (the "black box") for the solution of the partial differential equation. (The provided function `cirpdf.m` is really not so complex, since the Fokker–Planck equation above happens to have an analytical solutions. A more general model, however, would lack a simple solution an require a considerably larger and more obscure code.) The function `cirpdf.m` is very simple to use. It returns the value of the probability density function f in Eq. (3) and has the same arguments list.

Tasks

1. It might seem natural to solve the Fokker–Planck equation from t_0 and three years ahead in one go and then compare the result directly with the data to calculate the objective function. Why would this not work?

2. Implement an objective function for maximising Eq. (3) with a suitable optimisation routine from the optimisation toolbox of Matlab in mind.

3. Run the optimisation routine with your objective function and the given data as input. What values of (a, b, σ) do you find?