Computer exercise 1 - Maximum Likelihood

Aim: You will be introduced to some of the tools that are used on the financial markets. Given historical financial data the parameters of a stochastic process will be determined by Maximum Likelihood estimation and evaluated.
Background: You will be working in Matlab and the Excel environment of Thomson Reuters Eikon. In the Excel document historical prices are loaded from ThomsonReuters. All the functions are run from the Matlab script runML.m.
Preparation: Read through the laboration and ensure that you are familiar with each concept, and know how they are computed.

Download ml.zip from Lisam, which includes a Matlab program (runMLGARCH.m) for reading historical data from the file labML.xls and for calling the optimization algorithm which use the objective function value computed in likelihoodModGARCH.m.

The GARCH(1,1) process

$$\begin{aligned} \ln \frac{S_{i+1}}{S_i} &= \nu \Delta t + \sigma_i \xi_i \sqrt{\Delta t} \\ \sigma_{i+1}^2 &= \beta_0 + \beta_1 \sigma_i^2 + \beta_2 \frac{1}{\Delta t} \left(\ln \frac{S_{i+1}}{S_i} \right)^2 \end{aligned} \tag{1}$$

where $\xi_i \sim N(0,1)$, models a long run variance level, $\bar{\sigma}^2 = \frac{\beta_0}{1-\beta_1-\beta_2}$, and can also describe volatility clustering. Note that the expected return and volatility are expressed in years. To ensure that the variance is always positive and finite it is necessary to satisfy $\beta_0, \beta_1, \beta_2 \geq 0$ and $\beta_1 + \beta_2 \leq 1$. The maximum likelihood estimation can be formulated as

$$\begin{aligned} \max_{\nu,\beta,v} \quad & l = -\sum_{i=0}^{T-1} \frac{1}{2} \left(\ln 2\pi v_i + \frac{\left(\ln \frac{S_{i+1}}{S_i} - \nu \Delta t \right)^2}{v_i \Delta t} \right), \\ \text{s.t.} \quad & v_{i+1} = \beta_0 + \beta_1 v_i + \beta_2 \frac{1}{\Delta t} \left(\ln \frac{S_{i+1}}{S_i} \right)^2, \end{aligned} \tag{2}$$

where $v_i = \sigma_i^2$ has been used to get a constraint which is linear in v_i.

Exercise: Complete the function likelihoodNormal which computes the vector of log likelihood values of each realized return for a normally distributed variable (including the constant). Determine the optimal parameters (write in table 2) using the interior point solver in fmincon with mlGARCH! (Note that this function should be called from the function runML.)

Exercise: Determine the yearly average return and volatility for the historical data as well as the implied values from the GARCH(1,1) parameters (in table 1). Motivate if the GARCH(1,1) parameters are plausible. (RiskMetrics use an EWMA model for daily returns which implies $\beta_0 = 0$, $\beta_1 = 0.95$ and $\beta_2 = 0.05$.)

Använder nu och long term volatility enligt ovan. Beräknar aritmetiskt medelvärde och varians för historiska värden.

Tabell 2 verkar rimlig om man jämför med riskmetrics parametrar.

Exercise: Which improvements can you see in the QQ-plot?

Färre standardavvikelser för normaliserade returns. Dvs fördelningen fångar variationen bättre. Smalare svansar.

Exercise: Determine the log likelihood, AIC and BIC (in table 3).

Table 1: Model properties

	Average return	Volatility
Historical	8.551902%	32.538457%
Implied from GARCH(1,1)	5.517748%	31.521263%

To improve the model to also consider the aspect that volatility usually increase more when asset prices decrease, and an increase in asset prices usually lead to decreasing volatility, a modification is proposed. The stochastic process for the share price, S_i, $i = 0, \ldots, T$, is given by

$$\begin{aligned} \ln \frac{S_{i+1}}{S_i} &= \nu \Delta t + \sigma_i \xi_i \sqrt{\Delta t} \\ \sigma_{i+1}^2 &= \beta_0 + \beta_1 \sigma_i^2 + \beta_2 \frac{1}{\Delta t} \left(\ln \frac{S_{i+1}}{S_i} - \alpha_0 \Delta t \right)^2 \quad \text{if } \ln \frac{S_{i+1}}{S_i} < 0 \\ \sigma_{i+1}^2 &= \beta_0 + \beta_1 \sigma_i^2 + \beta_2 \frac{1}{\Delta t} \left(\ln \frac{S_{i+1}}{S_i} - \alpha_1 \Delta t \right)^2 \quad \text{if } \ln \frac{S_{i+1}}{S_i} \geq 0 \end{aligned} \tag{3}$$

where

$$\xi_i \sim N(0, 1). \tag{4}$$

By selecting appropriate values for α_0 and α_1 a larger increase in volatility when asset prices decrease, and a smaller increase in volatility when asset prices increase can be obtained. Maximum Likelihood estimation of the parameters gives the following optimization problem

$$
\begin{aligned}
\max_{\nu,\beta,\alpha,v} \quad & l = \sum_{i=0}^{T-1} -\frac{1}{2}\ln v_i - \frac{1}{2}\frac{\left(\ln\frac{S_{i+1}}{S_i} - \nu\Delta t\right)^2}{v_i\Delta t} \\
\text{s.t.} \quad & v_{i+1} = \beta_0 + \beta_1 v_i + \beta_2 \tfrac{1}{\Delta t}\left(\ln\tfrac{S_{i+1}}{S_i} - \alpha_0\Delta t\right)^2 \quad \text{if } \ln\tfrac{S_{i+1}}{S_i} < 0 \\
& v_{i+1} = \beta_0 + \beta_1 v_i + \beta_2 \tfrac{1}{\Delta t}\left(\ln\tfrac{S_{i+1}}{S_i} - \alpha_1\Delta t\right)^2 \quad \text{if } \ln\tfrac{S_{i+1}}{S_i} \geq 0
\end{aligned}
\tag{5}
$$

Note that in this particular instance the if-clause does not cause any problem for the optimization algorithm, since it is determined by $\ln\frac{S_{i+1}}{S_i}$ which is independent of the variables in the ML estimation. (The constraints are still continuously differentiable.)

Exercise: Write the function varModGARCH which computes the variance for the modified GARCH. Determine the optimal parameters (write in table 2) using the interior point solver in fmincon with mlModGARCH!

Table 2: Model parameters

	GARCH(1,1)	mod GARCH(1,1)
ν	0.0552	0.0538
β_0	8.3460e-04	5.1617e-04
β_1	0.9342	0.9314
β_2	0.0574	0.0590
α_0	0	1.5306
α_1	0	1.8494

Exercise: Which improvements can you see in the QQ-plot?

Inget uppenbart, men aningen smalare svansar på Mod garch QQ.
. . . .

. . . .

Exercise: Determine the log likelihood, AIC and BIC (in table 3).

Exercise: Split the data into an in-sample and out-of-sample data set. Estimate both the GARCH(1,1) and modified GARCH(1,1) models on the in-sample data. Determine the log likelihood, AIC and BIC (in table 3).

Table 3: Model evaluation

	GARCH(1,1)	mod GARCH(1,1)
Log likelihood	-929.882751	-897.003825
AIC	1867.765503	1806.007651
BIC	1895.704779	1847.916565
In-sample log likelihood	-1051.947105	-1021.277591
In-sample AIC	2111.894210	2054.555183
In-sample BIC	2137.061399	2092.305965
Out-of-sample log likelihood	136.021351	152.474125

Exercise: Determine if the difference in out-of-sample log likelihood values between the models are statistically significant. What are the conclusions from the out-of-sample test?

Statistiskt signifikant med knofidensnivå ca 93.99%. Slutsatsen är att mod GARCH(1,1) troligtvis beskriver

avkastningarna bäst.

Preparation: Is $-\ln x$ a concave function? Is $-\frac{1}{x} - \frac{1}{a+x}$ a concave function for all values a?

Derivering två ggr ger att första funktionen är konvex ty andraderivatan alltid positiv.

PSS för andra funktionen, med a = -1.5x, ger positiv andraderivata för vissa x. Svar: nej och nej.

Exercise: Solve the ML estimation for the modified GARCH(1,1) model from multiple starting solutions. Are all the solutions plausible? Explain the results!

Nej, alla är inte tänkbara lösningar. Vissa får markant lägre loglikelihood och konstiga parametrar.

Log-likelihood funktionen (2) är inte konkav så vi kan hitta lokalt optimum.