S414 Advanced Time Series
Mandatory Assignment 3

Selami Dogan Akansu[a] and Arinjai Das[b]

[a] Eberhard Karls University of Tubingen, Economics and Finance Program, 6315870, sdakansu@gmail.com

[b]Eberhard Karls University of Tubingen, Economics and Finance Program, 6322070, arinjaid@gmail.com

14 December 2022

Abstract

The paper presents a MA(1) process. The paper is divided into two sections. The first section simulates the MA(1) process and examines the conditional log-likelihood functions and conditional log-likelihood contributions. The paper then estimates the parameters of the processes utilising the Conditional Maximum Likelihood (CML) toolbox. The second part of the paper examines the test statistic for the processes for further statistical inference and computes the parameters under varying conditions.

1 Introduction

Forecasting time series is commonly used in different areas like analyzing stock prices in financial markets, interest rates, and studying other long term trends in data. This paper includes time series analysis and prediction of its parameters with generated moving average (MA) processes. A first-order moving average process (MA(1)) is created with a specific parameter vector $\boldsymbol{\theta} \equiv (\mu, \theta, \sigma^2)$. Gaussian white noise $\varepsilon_t \sim$ N(0,σ^2) is used in the standard MA(1) process. The maximum likelihood estimation is a process of finding the value of the vector of population parameters (θ) for which the sample is most likely to be observed.

The Conditional Maximum Likelihood (CML) tool is used to find parameters of the MA(1) process. Conditional maximum likelihood estimates are preferred as they are easier to compute than exact likelihood estimates. T-scores and two sided p-values are examined and the results are interpreted. Same processes are repeated for different specifications of theta (θ), number of observations (T) of the MA(1) process, and minimization algorithms and covariance matrices of CML. The estimation of parameters μ, θ, ans σ^2 are tested with t-test with significance level of 5%. Further, a possible misleading case like finding local minima is detected in the final experiment and interpreted.

2 Experiments

2.1 Generating a MA(1) Process

The First-Order Moving Average process is created from the weighted sum of the two most recent values of ε. The general formula is represented in Equation 2.1 (Hamilton (1994), p. 127).

$$y_t = \mu + \theta\varepsilon_{t-1} + \varepsilon_t; \;\; \varepsilon_t \sim N(0, 1)$$

With the given parameter set, $\boldsymbol{\theta} \equiv (10, 0.2, 1)$, an MA(1) process is obtained for 100 observations in equation 1. The series fluctuates around the mean which is 10 in this experiment. In every observation, fluctuations around the mean are observed which is a direct effect of Gaussian White Noise and θ.

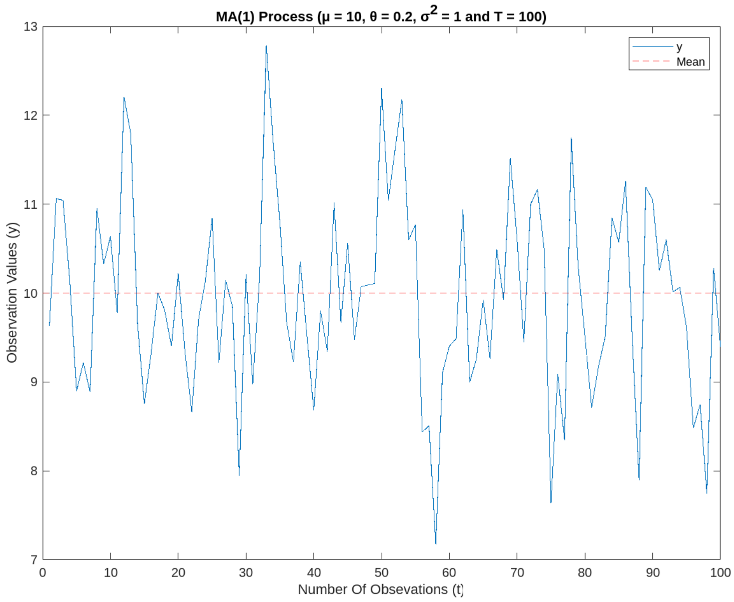

Figure 1: Observations of MA(1) Process of parameters $\boldsymbol{\theta} \equiv (10, 0.2, 1)$ with Gaussian white noise

2.2 Parameter Estimation of an MA(1) Process

This section deals with the creation of the MA(1) process and adapting the process to the set(s) of input variables. The section then covers the formulation and application of maximum log-likelihood and conditional log-likelihood to the MA(1) process for the set(s) of input variables.

2.2.1 Maximum Likelihood Estimation

The maximum likelihood estimate of $\boldsymbol{\theta}$; where $\boldsymbol{\theta} \equiv (\mu, \theta, \sigma^2)$ is the value of $\boldsymbol{\theta}$ for which the sample is most likely to have been observed. In other words, the value of $\boldsymbol{\theta}$ that maximises the probability density in Equation 2.2.1 (Hamilton (1994), p. 117).

$$f_{Y_t, Y_{t-1}, \ldots, Y_1}(y_t, y_{t-1}, \ldots, y_1; \boldsymbol{\theta}),\ where\ \boldsymbol{\theta} \equiv (\mu, \theta_t, \theta_{t-1}, \ldots, \theta_1, \sigma^2)$$

Variance-Covariance Matrix

The variance-covariance matrix is a *(T× T)* matrix containing the variances and covariances associated with the T observations of Y. The diagonal elements of the matrix contain the variances of the observations and the off-diagonal elements contain the covariances between all possible pairs of observations.

For a MA(1) process with a *(T×1)* vector $\mathbf{y} \equiv$ (y_1,y_2,...,y_T), mean $\boldsymbol{\mu} \equiv$ (μ,μ,...,μ), and (T× T) variance-covariance matrix $\boldsymbol{\Omega} = E(Y - \boldsymbol{\mu})(Y - \boldsymbol{\mu})'$, the likelihood function is given in Equation 2.2.1 (Hamilton (1994), p. 128).

$$f_Y(\mathbf{y}; \boldsymbol{\theta}) = (2\pi)^{-\frac{T}{2}} |\boldsymbol{\Omega}|^{-\frac{1}{2}} exp\left[-\tfrac{1}{2}(\mathbf{y} - \boldsymbol{\mu})'\boldsymbol{\Omega}^{-1}(\mathbf{y} - \boldsymbol{\mu})\right]$$

Exact Maximum Likelihood Estimate

The MLE $\hat{\theta}$ is the value which maximizes the log-likelihood of for a sample of size T of a MA(1) process. The exact log-likelihood of a Gaussian MA(1) process of sample size T is given in Equation 2.2.1 (Hamilton (1994), p. 129).

$$\mathcal{L}(\boldsymbol{\theta}) = log f_Y(\mathbf{y}; \boldsymbol{\theta}) - \tfrac{T}{2} log(2\pi) - \tfrac{1}{2}\textstyle\sum_{t=1}^{T} log(d_{tt}) - \tfrac{1}{2}\sum_{t=1}^{T} \frac{y_t^{-2}}{d_{tt}}$$

where d_{tt} are the terms along the principal diagonal of D; which is the triangular factorized matrix of the *(T× T)* variance-covariance matrix of Y.

2.2.2 Conditional Likelihood Function

Conditional likelihood function is an alternative to numerical maximization of exact likelihood function. Calculation of the likelihood function for a moving average process is simpler if the initial values of ε are conditioned. For example, in a general MA(1) process, if the value of ε_{t-1} is known, then $Y_t|\varepsilon_{t-1}$ is normally distributed along $\sim(\mu+\theta\varepsilon_{t-1},\sigma^2)$ In case of conditional likelihood, since the value of ε_0 is known, the calculation of exact likelihood estimate (MLE) becomes easier to compute. If the sample size is very large and the $|\theta|$<1, then the exact likelihood and the conditional likelihood are very similar.

The conditional log likelihood function of the MA(1) function is given in Equation 2.2.2.

$$ln\mathcal{L}\sum_{t=1}^{T} ln\left\{\frac{1}{\sqrt{2\pi\sigma^2}}exp\left[\frac{(y_t-\mu-\theta\varepsilon_{t-1})^2}{-2\sigma^2}\right]\right\} = \sum_{t=1}^{T}\left[ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)-\frac{\varepsilon_t^2}{2\sigma^2}\right]$$

The general conditional log-likelihood function of the MA(1) function is given in Equation 2.2.2 (Hamilton (1994), p. 128).

$$\mathcal{L}(\boldsymbol{\theta}) = logf_{Y_t,Y_{t-1},...,Y_1|\varepsilon_0=0}(y_t,y_{t-1},...,y_1|\varepsilon_0=0;\Theta) = -\frac{T}{2}log(2\pi)-\frac{T}{2}log(\sigma^2)-\sum_{t=1}^{T}\frac{\varepsilon_t^2}{2\sigma^2}$$

2.2.3 Constructing a Conditional Log-likelihood of an MA(1) process

The output of the MA(1) process is in the form of $(T\times 1)$ column vector of parameters $\mathbf{y}\equiv(y_1,y_2,...,y_T)$. The conditional log-likelihood is calculated with equation 2.2.1 using $\mathbf{y}$ and parameter vector , . To compute the process, conditional log-likelihood contributions is constructed that takes inputs $\boldsymbol{\theta}$ and $\mathbf{y}$, and returns an output vector of contribution values for every observation.

For the optimization process, a negative of the conditional log-likelihood is used to maximize the conditional log-likelihood using a minimization algorithm. The resultant conditional log-likelihood contributions are then multiplied with -1 to get the final result. Thus, when the process is coded, a negative conditional log-likelihood function is constructed.

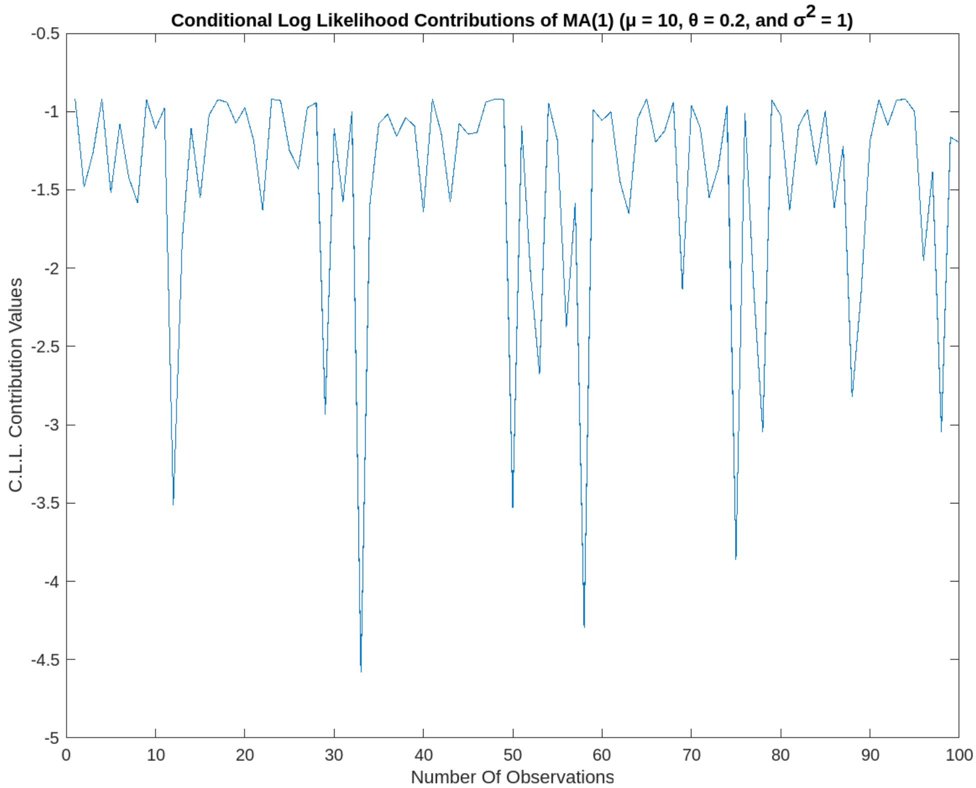

Figure 2: Conditional Log Likelihood Contributions of parameters (10, 0.2, 1)

Conditional Log Likelihood contributions of the MA(1) process with true parameter set is exhibited in Figure 2. The contributions vary below -0.9. Since conditional likelihood is a way to measure the goodness of fit for a model, a higher the value of the log-likelihood indicates a better model fit for the data set. In this result there are a few strong outliers, however, the contributions are stationary around -0.9.

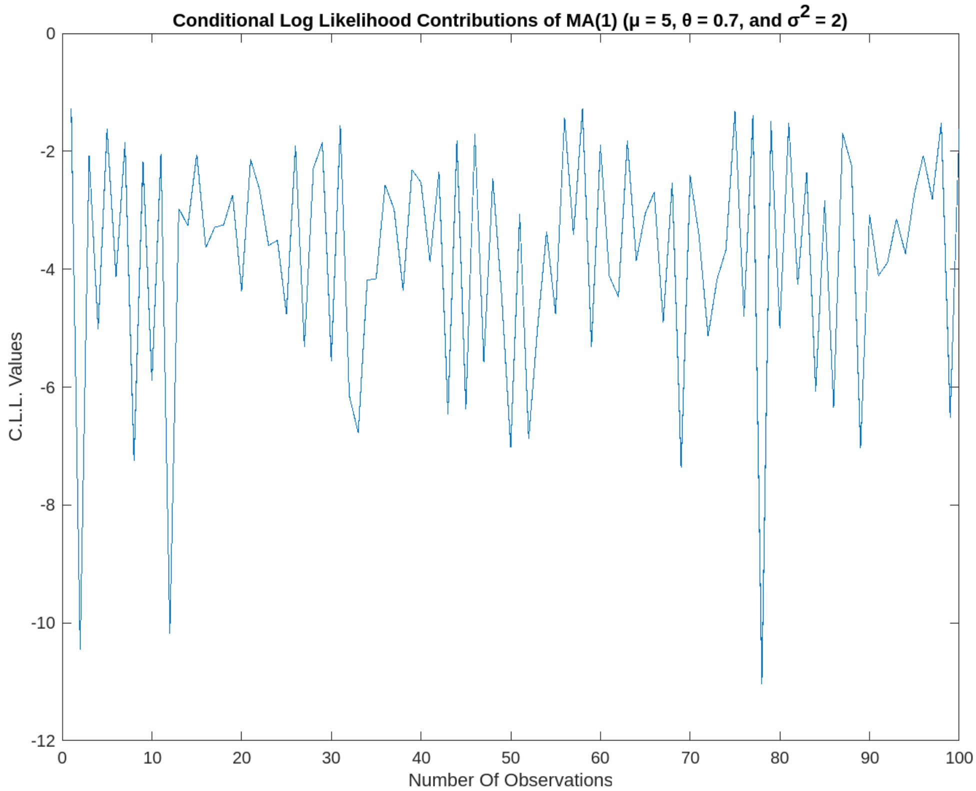

Figure 3: Conditional Log Likelihood Contributions of parameters (5, 0.7, 2)

To observe change of the contributions, a different parameter set, $\boldsymbol{\theta}' = (5, 0.7, 2)$, is used in conditional log likelihood contributions calculation. Even after changing input parameters, there is no discernable change in the fundamental behaviour of the result. The variance is higher here, but that can be explained by the doubling of the standard deviation and increase in θ from 0.2 to 0.7. Apart from that, the graph is stationary.

2.2.4 Estimating the Parameters of MA(1) process

The parameters of the MA(1) process, namely $\boldsymbol{\theta} \equiv (\mu, \theta, \sigma^2)$ are calculated using the CML toolbox. The toolbox utilises the conditional maximum likelihood process.

Parameters of the generated MA(1) are estimated using a Hessian-based covariance matrix in the $\boldsymbol{fminsearch}$ algorithm. The optimization process is based on conditional maximum likelihood process. The estimated parameters are $\boldsymbol{\theta} \equiv (9.9535, 0.2774, 1.0473)$.

Standard errors of parameters, t-test, and p-value are calculated to measure correctness of the estimation. The standard error of estimation for μ is 0.1310, for θ is 0.0850, and for σ^2 is 0.1481. For Null Hypothesis: (H_0): $\theta = 0.4$, the t-score is -1.1980. The t-test determines whether the estimated θ lies inside the confidence interval for a significance level of α=0.05. The confidence interval for θ is [0.1109, 0.4439]. Since the null hypothesis lies inside the confidence interval, the null hypothesis is not rejected.

Table 1: Standard Errors (T=100)

	μ	θ	σ
Hessian	**0.1310**	0.0850	0.1581
OPG	0.1315	0.1066	**0.1460**
QML	0.1318	**0.0681**	0.1511

Table 2: Standard Errors (T=50,000)

	μ	θ	σ
Hessian	0.0054	0.0044	0.0063
OPG	0.0054	0.0044	0.0063
QML	0.0054	0.0044	0.0063

The same processes are repeated for T=50,000. The results are compiled in Table 1 (T=100) and Table 2 (T=50,000). Comparing the two tables, there is a clear decrease in standard errors of two experiments including all standard error calculations from Hessian, OPG and QML covariance matrices. This indicates that the parameter estimations approaches the true values with an increase in number of observations for an MA(1) process.

2.3 Parameter Estimation of MA(1) Processes Ensembles

100 random instances of the MA(1) process are created with the same parameter set and arranged in an ensemble. Confidence intervals are then computed for every series in this ensemble,. For $\mu = 10$, there are 95 series that lie inside the confidence interval. For $\theta = 0.2$, 99 series that lie inside the confidence interval. Therefore, 95% of the estimated μ(s) and 99% of estimated θ(s) are correctly predicted with 5% significance level.

2.3.1 Conditional Log Likelihood For Varying θ

Conditional Log Likelihood values are calculated for θ varying between [-0.6, 0.8] with increments of 0.05.

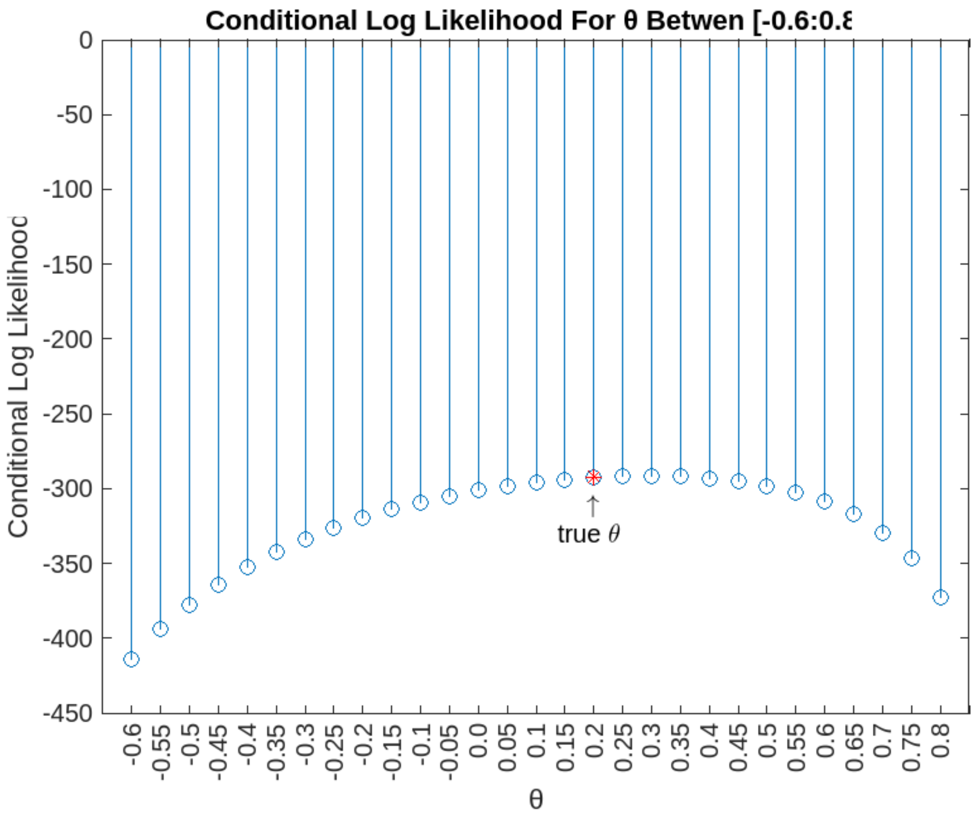

Figure 4: Conditional Log Likelihood Contributions of parameters (5, 0.7, 2)

The value corresponding to the resultant $\hat{\theta}$ is plotted in Figure 4. The estimation of parameter θ is 0.2774 which is obtained in Section 2.2.4. The true value of θ, 0.2, is not equal to the value of θ which maximizes the conditional log likelihood.

2.3.2 Local Maxima in Parameter Estimation

The maximization of log likelihood may end up in the local maxima, in which case the estimated parameters could be false and not equal to true parameters. To demonstrate this, a MA(1) process, created with 200 observations with the parameter set $\boldsymbol{\theta}' = (10, 5, 1)$, . ***fminunc*** function is selected for CML tool to do the estimation. Using the CML tool, the optimization algorithm outputs an estimated parameter set that is different from the true parameter set. The result of the estimation for μ is 9.2285, for θ is 0.1128 and for σ^2 is 22.6860. Compared to the previous findings, the CML algorithm could not find a close estimation to the true parameter set. This implies that the CML estimation estimated a local maxima instead of the global maxima, which is the true parameter set.

3 Conclusion

In the light of the experiments and findings. There are several deductions that are shaped. The MA(1) model is mean and covariance stationary, since it has a mean of μ and a variance of $\sigma^2(1+\theta^2)$. This is also visually confirmed. Further, the correctness of the MA(1) process is indicated by the ensemble, since ¿95% of the parameter values lie within the confidence interval. Results also indicate that increasing the number of observations of an MA(1) process leads to have a better estimation, evident from the inverse correlation of standard errors and number of observations is increased. Confidence intervals are also found to be more stable when higher number of observations are taken into account.

References

Hamilton, J. D. (1994). *Time series analysis.* Princeton University Press.

Appendix

Important functions that are used commonly in the experiments are provided in this section. All codes are enclosed in a compressed file. The experiments can be repeated in one run of Assignment3.m file, alternatively section by section.

```
function [ y ] = moving_average_process(mean, theta_1, sigma_square, T)

e = normrnd(0, sigma_square, [T+1,1]);
y = mean + theta_1.*e(2:end) + e(1:end-1);

end


function [L] = conditional_log_likelihood_contribution(parameters_vector,
    output_vector)

     e_0 = 0;
     mean = parameters_vector(1);
     theta = parameters_vector(2);
     sigma_square = parameters_vector(3);

     [rows, columns] = size(output_vector);

     L = zeros(rows, 1);
     e = e_0;

     for i = 1:rows

          %we have only e_0, so calculations for "e" are needed for next
          %iterations
          if i == 1
               e = e_0;
          else
               e = output_vector(i) - mean - theta*e;
          end

               L(i) = log(1/sqrt(2*pi*sigma_square)) - (e^2 / (2 *
                   sigma_square));
     end
end
```

```
function [sum_log_likelihood] =
    negative_conditional_log_likelihood_function(parameters_vector,
    output_vector)
%This function computes, first, the log likelihood contri- butions vector,
% and then sums it up to get the conditional log likelihood function.

contributions_vector = conditional_log_likelihood_contribution(
    parameters_vector, output_vector);

sum_log_likelihood = -1 * sum(contributions_vector);

end
```