Universitetet i Oslo
Matematisk institutt

Obligatory assignment: **STK4060/9060 – Time series**

With: **Emil Aas Stoltenberg**,

Email: **emilas@math.uio.no**,

Handed out: **October 20, 2025**

Due: **November 6, 2025**. Complete guidelines about delivery of mandatory assignments, along with a 'log on to Canvas', can be found here:
uio.no/english/studies/examinations/compulsory-activities/mn-math-mandatory.html

Permitted aids: All aids, including collaboration, are allowed, but the submission must be written by you and reflect your understanding of the subject. If we doubt that you have understood the content you have handed in, we may request that you give an oral account.

Format for your answer: **A pdf.**

Exercise 1 The autocovariance function. The autocovariance function and the autocorrelation function are fundamental tools for stationary time series. In this exercise we look at certain dangers lurking when the empirical autocovariance is applied to a time series that is *not* stationary.

(a) Please go to R and simulate a mean zero multinormal vector $(X_1, \ldots, X_n)$ with $\mathrm{Cov}(X_t, X_s) = \rho^{|t-s|}$ for all t, s, where $\rho = 0.4$ and $n = 100$. An convenient way of simulating this involves the Cholesky decomposition of the desired covariance matrix using `chol(Sigma)` in R, etc. You do not need to make a plot of these data, but do report the values you find for the empirical autocovariance $\widehat{\gamma}_X(h)$ for $h = 0, 1, 2, 3$. You might use the command `acf(xx,"covariance",lag=3)$acf` in R.

(b) Then change the setup above slightly and introduce a modest linear trend; in particular, consider $Y_t = \beta t + X_t$ for $t = 1, \ldots, n$, where the X_t are as above. Find the autocovariance function $\gamma_Y(h)$ of the Y process. Next, simulate $Y_1, \ldots, Y_n$ for $\rho = 0.4$, $\beta = 0.05$, and $n = 100$. Plot the Y_t series as well as its empirical autocovariance $\widehat{\gamma}_Y(h)$.

(c) The empirical autocovariance $\widehat{\gamma}_Y(h)$ does not look at all like the real underlying $\gamma_Y(h)$. For the more general case where $Y_t = \mu_t + X_t$, with X_t as above and μ_t a deterministic sequence, find an expression for the expected value of $\widehat{\gamma}_Y(h)$ for $0 \le h \le n$. For the specific case of $\mu_t = \beta t$, as above, see what these expectation formulae amount to, and comment on your findings.

(d) This little demonstration illustrates that the empirical autocovariance might be incorrectly intepreted if applied to a series that is not stationary. One standard fix is to 'detrend first', and then work with the detrended series $Y_t^* = Y_t - \widehat{\mu}_t$. Do this for your simulated data (assuming that you known the trend is linear $\mu_t = \beta t$), and give $\widehat{\gamma}^*(0), \widehat{\gamma}^*(1), \widehat{\gamma}^*(2), \widehat{\gamma}^*(3)$ for your detrended series. Comment on your findings.

Exercise 2 LIKELIHOOD AND WHITTLE LIKELIHOOD. Let $X = \{X_t : t \in \mathbb{Z}\}$ be the causal AR(1) process

$$X_t = \phi X_{t-1} + W_t, \tag{1}$$

where $W = \{W_t : t \in \mathbb{Z}\}$ consists of i.i.d. $\mathrm{N}(0,1)$ random variables, i.e., we take this variance as known. Suppose that we observe a sample $X_1, \dots, X_n$ from X.

(a) Go to R and simulate a sample of size $n = 100$ from the model above, with $\phi = 0.543$ (remember that we need som 'burn-in' to reach stationarity). Make a plot of your simulated sample path, that is, reproduce a version of the plot in Figure 1.

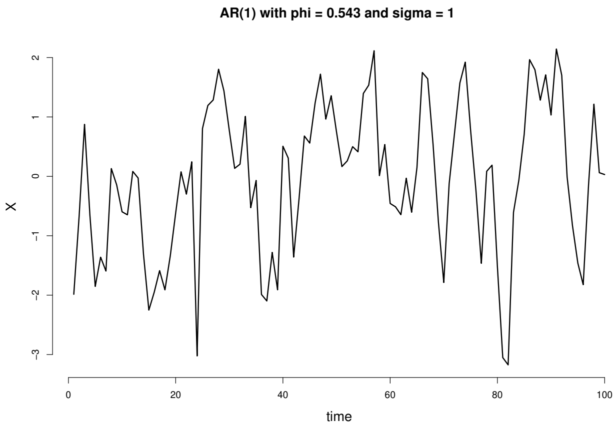

FIGURE 1. A sample path of an AR(1) process, as described in Ex. 2(a).

(b) Show that the log-likelihood function based on $X_1, \dots, X_n$ can be written

$$\ell_n(\phi) = \tfrac{1}{2}\log(1-\phi^2) - \frac{1-\phi^2}{2}X_1^2 - \frac{1}{2}\sum_{t=2}^{n}(X_t - \phi X_{t-1})^2.$$

Find an expression for the score function $\partial \ell_n(\phi)/\partial\phi$, and also for

$$J_n = -\frac{1}{n}\frac{\partial^2}{\partial\phi^2}\ell_n(\phi).$$

(c) The following is useful: Establish that $\{X_t^2 : t \in \mathbb{Z}\}$ is a stationary process with an absolutely summable covariance function, and from this deduce that

$$\frac{1}{n}\sum_{t=1}^{n} X_t^2 \to_p \frac{1}{1-\phi^2}.$$

(d) Show that $J_n \to_p J$ for a positive J, that you are supposed to find an expression for. It can be shown that the maximum likelihood estimator $\widehat{\phi}_n$ is such that $\sqrt{n}(\widehat{\phi}_n - \phi) \to_d \mathrm{N}(0, 1/J)$ (and you are free to show it, but need not). Comment on how this compares to the limit distribution of the least squares estimator for ϕ.

(e) We now turn to the Whittle estimator for ϕ. Provide a not-too-technical explanation for why it makes sense to use as an estimator for ϕ the minimiser $\widetilde{\phi}_n$ of the function

$$\ell_n^w(\phi) = -\sum_{j=1}^{\lfloor (n-1)/2 \rfloor}\Big\{\log f(j/n,\phi) + \frac{|d_X(j/n)|^2}{f(j/n,\phi)}\Big\},$$

where $d_X(j/n) = n^{-1/2}\sum_{t=1}^n X_t \exp(-i2\pi tj/n)$, $j = 1,\dots,\lfloor (n-1)/2\rfloor$ is the discrete Fourier transform of $X_1,\dots,X_n$; and $f(\omega,\phi)$ is the spectral density of X.

(f) With the simulated data from **(a)** plot a realisation of the Whittle log-likelihood. Comment on what you find.

(g) It can be shown that $\sqrt{n}(\widetilde{\phi}_n - \phi)$ converges in distribution to a mean zero normal. Perform a simulation study to compare the Whittle estimator and the maximum likelihood estimator for ϕ. Try some different sample sizes $n = 100$, $n = 500$ and $n = 1000$, for example, and comment on what you observe.

(h) Under the assumption that $|d_X(j/n)|^2$ is exponential with mean $f(j/n,\phi)$ for all $j = 1,\dots,\lfloor (n-1)/2\rfloor$, show that

$$-\frac{1}{n}\mathrm{E}\,\frac{\partial^2}{\partial\phi^2}\ell_n^w(\phi) = \frac{4}{n}\sum_{j=1}^{\lfloor (n-1)/2\rfloor}\{\phi - \cos(2\pi j/n)\}^2 f(j/n,\phi)^2,$$

Find its limit as n tends to infinity (by doing the mathematics, or by other perhaps more unscrupulous means, or by starting with the latter then doing the former).

Exercise 3 Linear regression with correlated errors. Linear regression with i.i.d. noise is canonical. In this exercise we consider the case where the noise stem from a causal AR(1) process. Let $X = \{X_t\colon t\in\mathbb{Z}\}$ be a sequence of i.i.d. random variables with mean μ_X and variance σ_X^2; and let $\{W_t\colon t\in\mathbb{Z}\}$ be i.i.d. with mean zero and variance σ_W^2. Suppose further that X and W are independent. Introduce $\varepsilon_t = \phi\varepsilon_{t-1} + W_t$ for $t\in\mathbb{Z}$ where $|\phi| < 1$, and assume that

$$Y_t = \alpha + \beta X_t + \varepsilon_t, \quad t\in\mathbb{Z}.$$

Based on data $(X_1,Y_1),\dots,(X_n,Y_n)$ we are to make inference on β. Let $\widehat{\beta}_n$ be the least squares estimator of β, so that (which you may verify)

$$\sqrt{n}(\widehat{\beta}_n - \beta) = \frac{n^{-1/2}\sum_{t=1}^n (X_t - \bar{X}_n)\varepsilon_t}{n^{-1}\sum_{t=1}^n (X_t-\bar{X}_n)^2}, \tag{2}$$

where $\bar{X}_n = n^{-1}\sum_{t=1}^n X_t$.

(a) To show that the denominator in this expression converges in probability we cannot rely on Chebyshev's inequality (why?). As you may know from before, the law of large numbers holds for i.i.d. random variables with finite mean, with no assumptions about higher moments: Let $Z_1,\dots,Z_n$ be i.i.d. replicates of Z, and suppose that $\mathrm{E}\,Z = 0$ (w.l.o.g., as one says). Show that for any $\varepsilon > 0$ and any $K > 0$,

$$\Pr(|n^{-1}\sum_{i=1}^n Z_i| \geq \varepsilon) \leq \frac{2}{\varepsilon}\mathrm{E}\,|Z|I\{|Z|\geq K\} + \frac{4}{\varepsilon^2}\frac{K}{n}\mathrm{E}\,|Z|,$$

or a version thereof. Explain why this inequality entails that $n^{-1}\sum_{i=1}^n Z_i \to_p 0$; and deduce that for the denominator on the right in (2), we have $n^{-1}\sum_{t=1}^n (X_t - \bar{X}_n)^2 \to \sigma_X^2$.

(b) The convergence in probability result from **(a)** combined with the Cramér–Slutsky rules, allow us to concentrate on the numerator on the right in (2). Explain why we are allowed to write

$$n^{-1/2}\sum_{t=1}^n (X_t - \bar{X}_n)\varepsilon_t = n^{-1/2}\sum_{t=1}^n (X_t - \mu_X)\varepsilon_t + r_n,$$

where r_n tends in probability to zero as n tends to infinity.

(c) For some $m \geq 1$, let $\varepsilon_t^m = \sum_{j=0}^m \phi^j W_{t-j}$, and argue that $(X_t - \mu)\varepsilon_t^m$ for $t \in \mathbb{Z}$ is a strictly stationary m-dependent sequence of mean zero random variables. Finally, deduce from the above efforts that

$$\sqrt{n}(\widehat{\beta}_n - \beta) \to_d \mathrm{N}\big(0, \frac{\sigma_W^2}{\sigma_X^2(1-\phi^2)}\big).$$

Discuss how this nicely generalises the standard central limit result for linear regression with i.i.d. noise. Provide also a brief interpretation.

(d) For the convergence in distribution result from (c) to be practically useful, we need an estimator for the limiting variance $(\sigma_W^2/\sigma_X^2)(1-\phi^2)^{-1}$. Please propose a consistent estimator of this quantity.

Exercise 4 Price of Euro in NOK. On the website of Norges Bank one can find the price of one Euro in Norwegian kroner over time. Here is a link. Download the data from October 21, 2024 up until the last working day. In R one can do:

```
# download data from Norges Bank
eur_link <- "https://data.norges-bank.no/api/data/EXR/B.EUR.NOK.SP?
    format=csv&bom=include&apisrc=
    nbi&startPeriod=2024-10-21&endPeriod=2025-10-17&locale=no"
eur <- read.csv(eur_link,sep=";",header=TRUE)
xx <- as.numeric(gsub(",",".",eur$OBS_VALUE))
n <- length(xx)
```

Let $X_1, \ldots, X_n$ denote the random variables whose realisations we have just downloaded. Figure 2 contains a plot of these data from October 21, 2024 up to October 17, 2025.

(a) Display a plot of the downloaded currency data. Fit an AR(1) model $X_t = \phi X_{t-1} + W_t$ to these data, and report your estimate of ϕ. Comment on what you find. If feasible, provide a 95 percent confidence interval for ϕ.

Figure 2. The price of an Euro in Norwegian kroner, from October 21, 2024 to October 17, 2025. Data from Norges Bank.

(b) In view of the above, it might be relevant to test whether we are dealing with an AR(1) process or a random walk. Let the null hypothesis be random walk ,$\phi = 1$; and the alternative hypothesis be a causal AR(1), so $|\phi| < 1$. With $\widehat{\phi}_n$ the least squares estimator of ϕ, it can be shown that under the null hypothesis

$$n(\widehat{\phi}_n - 1) \to_d \frac{\frac{1}{2}(\chi_1^2 - 1)}{\int_0^1 B(t)^2 \, \mathrm{d}t}, \tag{3}$$

where $\chi_1^2 \sim \text{Gamma}(1/2, 1/2)$ is independent of B, and B is a Brownian motion on $[0,1]$, i.e., the (process) limit in distribution of $m^{-1/2} \sum_{j=1}^{\lfloor mt \rfloor} \xi_j$, where $\xi_1, \xi_2, \ldots$ are i.i.d. mean zero with unit variance. Via simulations or by other means, test the unit root null hypothesis $\phi = 1$. Comment on what you find.

(c) Perhaps the currency time series is a random walk? Suppose that $X_t = X_{t-1} + W_t$ for i.i.d. $W_t \sim \mathrm{N}(0, \sigma^2)$. Simulate i.i.d. data from $\mathrm{N}(0, \widehat{\sigma}_n^2)$, where $\widehat{\sigma}_n^2$ is an estimator of σ^2, and plot the simulated W_t alongside the observed differences $X_t - X_{t-1}$ for $t = 2, \ldots, n$. Comment on similarities and differences between these plots.

(d) Another model for the currency data is $X_t = X_{t-1} + \sigma_t W_t$, where the W_t are i.i.d. standard normals, independent of the σ_t^2, which are themselves i.i.d. random variables. Specifically, we take $\lambda_t = 1/\sigma_t^2$ to be i.i.d. $\text{Gamma}(\lambda_0/c, 1/c)$. Provide a brief interpretation of the role played by the parameter c. Then, show that the likelihood function based on the differences $X_t - X_{t-1}$ for $t = 2, \ldots, n$ takes the form

$$L_n(\lambda_0, c) = (2\pi)^{-n/2} c^{-n\lambda_0/c} \frac{\Gamma(\lambda_0/c + \frac{1}{2})^n}{\Gamma(\lambda_0/c)^n} \prod_{t=2}^{n} \big(1/c + \tfrac{1}{2}(X_t - X_{t-1})^2\big)^{-(\lambda_0/c + \frac{1}{2})}.$$

(e) Estimate the parameters λ_0 and c and report their values. With the estimated values of these parameters, simulate data from the model introduced in **(d)**, and plot these alongside the observed $X_t - X_{t-1}$. Comment on what you find.

(f) Build you own model for the currency data. On the real data, try to compare the one-step-ahead predictions of the random-variance random walk model from **(d)** with those given by the model devised by you. Which one is the winner?

Slutt

Please send me an email as soon as possible if (when?) you spot mistakes. Thank you!