STAT 520 – Homework 4 – Fall 2025

For Homework 4A, do Problem 1 and Problem 3 below. Also for Homework 4A, you may choose to do Problems 5 through 7 below which are **extra credit** for everyone (it requires calculus to do these).

For Homework 4B, do Problem 2 and Problem 4 below.

Recall: The `arima` function in the `TSA` package in R fits a mean-centered model by default, if the specified model is an AR, MA or ARMA model (i.e., a stationary model).

1) A data set of 57 consecutive measurements from a machine tool are in the `deere3` object in the `TSA` package. Type `library(TSA); data(deere3); print(deere3)` in R to see the data set.
(a) Estimate the parameters of a (mean-centered) AR(1) model for this series. Use the least squares method and maximum likelihood, and report the estimated parameters from each of these methods. Comment on any similarities and differences.
(b) Estimate the parameters of a (mean-centered) AR(2) model for this series. Use the least squares method and maximum likelihood, and report the estimated parameters from each of these methods. Comment on any similarities and differences.
(c) Compare the results of the ML fits from parts (a) and (b) using AIC. Which model do you believe is preferable? Briefly explain your answer.

2) Consider the (mean-centered) AR(1) model for the `deere3` data in the `TSA` package, estimated using maximum likelihood.
(a) Give a basic plot of the standardized residuals over time and a Q-Q plot of the residuals. Comment on what these tell you about the adequacy of the model.
(b) Give a plot of the sample autocorrelation function of the residuals. Also perform a runs test and a Ljung-Box test (with $K = 8$). Comment on what these tell you about whether the errors are independent in this model.
(c) Diagnose the fit of the AR(1) model by using the overfitting strategy.

3) A data set of 324 measurements of an industrial robot's positions are in the `robot` object in the `TSA` package. Type `library(TSA); data(robot); print(robot)` in R to see the data set.
(a) Estimate the parameters of a (mean-centered) AR(1) model for these data, using maximum likelihood. Give the equation of the estimated model.
(b) Give an approximate 95% confidence interval for ϕ, the coefficient in the AR(1) model.
(c) Estimate the parameters of an IMA(1,1) model (a.k.a. an ARIMA(0,1,1) model) for these data, using maximum likelihood. Give the equation of the estimated model.
(d) Note that the output labeled "aic" given by the `arima()` function output uses a slightly different definition of AIC than the `AIC()` function in R does. Verify this by reporting both the "aic" output and the `AIC()` result for the models in parts (a) and (c). However, as long as you use the same definition when comparing models, you can use either definition (since the formulas only differ by a constant). But why is it not recommended to compare the models in parts (a) and (c) using AIC?

4) Consider the (mean-centered) AR(1) model for the `robot` data in the `TSA` package, estimated using maximum likelihood.
(a) Give a basic plot of the standardized residuals over time and a Q-Q plot of the residuals. Comment on what these tell you about the adequacy of the model.
(b) Give a plot of the sample autocorrelation function of the residuals. Also perform a runs test and a Ljung-Box test (with $K = 30$). Comment on what these tell you about whether the errors are independent in this model.
(c) Diagnose the fit of the AR(1) model by using the overfitting strategy.
(d) Repeat part (b), but with the residuals from a (mean-centered) AR(2) model for the robot data. Comment on whether your conclusions are any different.

Problem 5 (**extra credit**): Look at the conditional sum-of-squares expression $S_c(\phi, \mu)$ in the Chapter 7 notes. Take the partial derivative of $S_c(\phi, \mu)$ with respect to μ, set this equal to zero, and solve for μ to obtain the least squares estimator of μ that is given in the formula on the next slide. Carefully show all your steps.

Problem 6 (**extra credit**): Look at the conditional sum-of-squares expression $S_c(\phi, \mu)$ in the Chapter 7 notes. Take the partial derivative of $S_c(\phi, \mu)$ with respect to ϕ, set this equal to zero, and solve for ϕ to obtain the least squares estimator of ϕ that is given in the formula on the next slide. (You can plug in $\bar{Y}$ in place of μ in the expression for $S_c(\phi, \mu)$ before you begin.). Carefully show all your steps.

Problem 7 (**extra credit**): Look at the likelihood function under the AR(1) that is given in the Chapter 7 notes. We want to find the maximum likelihood estimator of the noise variance. First, take the natural logarithm of the likelihood function. To ease the notation, use v instead of σ_e^2 to denote the noise variance. Then take the derivative of the log-likelihood function with respect to v and set this equal to zero. Solve for v to obtain the maximum likelihood estimator of v (that is, σ_e^2) given on the next slide (note that you can simply replace ϕ and μ with notation for their estimates as is done in the formula). Carefully show all your steps.