DS-GA 3001.001 Special Topics in Data Science: Probabilistic Time Series Analysis
Homework 1

Due date: September 27, by 6pm

Problem 1. Consider the *sample mean* of a stationary time series x_t, defined as:

$$\hat{\mu} = \frac{1}{T} \sum_t x_t. \tag{1}$$

Compute the variance of this estimate $Var[\hat{\mu}]$, as a function of T, and the autocovariance function $\gamma(h)$.
Hint: The empirical mean is also a linear combination of random variables, so you can use the formula for the covariance of linear combinations of random variables from the lecture.

Problem 2. Confidence bounds for the autocorrelation function: show that the variance of the empirical ACF for white noise with variance σ^2 estimated given T data points is $\frac{1}{T}$.
Hint: Use theorem A.7 from tsa4.pdf; alternatively, you can just show it numerically by plotting empirical estimates of the ACF as a function of T.

Problem 3. For an MA(1), $x_t = w_t + \theta w_{t-1}$ show that the autocorrelation function $|\rho_x(1)| \leq 0.5$, for all θ. For which values θ is it maximum/minimum?

Problem 4. Identify the following models as ARMA(p, q):

- $x_t = 0.8x_{t-1} - 0.15x_{t-2} + w_t - 0.3w_{t-1}$
- $x_t = x_{t-1} - 0.5x_{t-2} + w_t - w_{t-1}$

Note: watch out for parameter redundancy!

Problem 5. Having observed a sequence $\{x_1, x_2 ... x_t\}$ we are trying to predict a future observation x_{t+h}, with $h \geq 1$. How well / far can one predict into the future if the data comes from a a) MA(3) and b) AR(1) model.
Hint: think of the functional form of the optimal estimator and/or the corresponding graphical model.

Problem 6. Given the AR(2) process with $P(B) = (1 - 0.2B)(1 - 0.5B)$, what is $\rho(h)$? Check your analytical solution against an empirical estimate obtained using the code from the lab.
Hint: Difference equations + initial conditions.