Time Series Analysis & Recurrent Neural Networks

lecturer: Daniel Durstewitz
tutors: Georgia Koppe, Manuel Brenner, Daniel Kramer, Leonard Bereska
WS2020/2021

Exercise 3

To be uploaded before the exercise group on December 2nd, 2020

Task 1. Univariate AR models
In file ex3file1.mat (ex3file1.xls) you will find 4 time-series obtained from human fMRI recordings from the dorsolateral prefrontal cortex (DLPFC) and the parietal cortex (Parietal), obtained during a working memory task. For task 1, consider the first time series of DLPFC (termed 'DLPFC1').

1. Compute the log-likelihood of an AR(4) model (e.g, by employing lecture script 1, eq. 7.34). Please write down explicitly.

2. Determine the optimal order p of the AR model by computing the log-likelihood-ratio test statistic (Wilk's D, see 'ASMN_Ch1_Durstewitz.pdf', eq. 1.35, and/or lecture script 1, eq. 7.35). Start with the comparison between an AR(2) vs. AR(1) model (look up D in the χ^2 distribution with appropriate degrees of freedom, Matlab function chi2cdf.m, or scipy.stats.chi2.cdf() in Python), and then keep on computing log-likelihood-ratio test statistics for models of consecutive orders until the difference between models is no longer significant. Do this up to order $p = 5$.

Task 2. Multivariate (vector) AR (=VAR) processes
For this task, use all four time series contained in the data file ('DLPFC1', 'DLPFC2', 'Parietal1', 'Parietal2').

1. Estimate a VAR(1) model by performing multivariate regression on the 4-variate time series. What do the coefficients in matrix A tell you about the coupling between the DLPFC and parietal cortex? Is the resulting VAR(1) model stationary or not?

Task 3. AR Poisson processes

1. Create your own order 2 Poisson time series with T=1000 time steps and the following given parameters: $A_1 = \begin{pmatrix} 0.2 & -.1 \\ 0.1 & 0.1 \end{pmatrix}$, $A_2 = \begin{pmatrix} 0.1 & -.1 \\ 0.1 & 0.1 \end{pmatrix}$, and $\mu_0 = \begin{pmatrix} .5 & .5 \end{pmatrix}^\intercal$. Assume no base rate (i.e. $a_0 = \begin{pmatrix} 0 & 0 \end{pmatrix}^\intercal$).

2. Given the data generated in (a), vary the parameters $A_1(1, 1)$ and $A_2(2, 1)$ between 0 and 0.4 with 0.01 increments. For each parameter value pair, compute the log-likelihood of the data (keeping all other parameters fixed!). Plot the log-likelihood landscape surface as a function of these two parameters. Does the real parameter pair value correspond (or is close) to an extreme point in the approximate log-likelihood landscape? What kind of an extreme point is it?