DESCRIPTION OF THE PRATICAL PROJECT FOR THE COURSE DATA-DRIVEN MODELLING OF DYNAMICAL SYSTEMS AND OPTIMAL CONTROL

a.y. 2021-2022

INPUT DATA:

kobe.mat → seismic activity from Kobe earthquake in Japan, 1995

INFORMATION

The data contain three phases:

1. Data from 1 to 1200 → normal activity
2. Data from 1201 to 1700 → transition to the earthquake
3. Data from 1701 till the end → the real earthquake period

GOAL: USE THE DATA TO REPLICATE ALL THE MAIN DATA-ANALYSIS PHASES SEEN IN THE COURSE!

More precisely, you have to perform the following steps, writing and commenting the Matlab code you develop for it, or describing precisely the use of the GUI for those steps that you can take through it.

1. Preliminary data analysis

- Plot the time series
- From the data, detect the sequences associated with different seismic events (known a priori and described above), and save them in three different time series.
- Plot the data showing in different colours the different parts
- Compute the mean and the variance of the data (in the normal seismic activity only)
- Remove the bias and plot the original and the unbiased data for the normal seismic activity
- Again for the normal seismic activity data:
    - Compute and plot the output spectra
    - Compute and plot the covariance function

2. Learning models from data

- Again for the normal seismic activity data:
    - o using 80% of data for training and 20% for validation, train a set of models and motivate the choice of the optimal model, both the model class and its order.
    - o Check if the optimal model is stationary
    - o Use the model to predict the training data and compare the real and predicted outputs and the prediction horizon increases. Comment the results
    - o Do the same on the validation data and compare the results
    - o Compute and plot the prediction error on training and validation data and its variance as the prediction error horizon increases. Comment the results
    - o Check whether the residuals of the prediction are white and comment.

It can be interesting, at the end of the analysis, to investigate whether it is possible, from the model built on normal activity data, to predict the earthquake measurements.

To see what happens take the data from phases 2 and 3, detrend it with the previously computed mean value for phase 1, and apply to it the 1-step ahead predictor that you previously obtained.

Does it work? What happens to the prediction error?