EE512 – Applied Biomedical Signal Processing

Practical session – Linear Models II

Instructions

- Please submit your report as a single PDF file.
- We recommend working in a group of 3–4 students; you can prepare one single report for the group (`name1_name2_name3_lab_LinearModelsII.pdf`), but every member needs to upload the same file individually.
- Include in the report the code you used when addressing the questions, together with your answers.
- The session is prepared to be done in Matlab – other languages like Python would be possible, but would require much more coding.
- We recommend creating a Matlab script – see "Important recommendation" below.

Useful commands

Type `help function_name` for more information on inputs and outputs.

`resample`	Downsample/upsample/resample a signal
`AR_order`	Estimate optimal model order using MDL (and `arcov` for estimation)
`AR_psd`	Parametric estimation of power spectral density (PSD)
`pwelch`	Non-parametric PSD estimation via Welch's method
`pisarenko`	Sinusoidal component estimation using the Pisarenko method
`aryule`	AR model estimation, autocorrelation method; uses Levinson-Durbin

`figure, plot`	Plot signals for visualization. The following example code illustrates how it could be used for a given signal `x` in time:
	`n = size(x,1);`
	`fs = 100; % Sampling frequency in Hz`
	`t = (0:n-1)'/fs; % Sampling times in sec`
	`figure('Units','centimeters','Position',[0,0,17,17],'Color','w');`
	`subplot(3,1,1); plot(t,x); title('Signal'); xlabel('Time (s)');`
	`subplot(3,1,2); etc…`

Important recommendation: Write the commands of your analysis in a Matlab script, starting with a `clear all` command so that your functions start on an empty Matlab workspace. This helps keeping track of what you are doing and can avoid getting confusing results when variables are changed by accident with repeated commands (which can often happen in the command window). At the end, you can simply copy the relevant code from the script to your report.

Experiment 1: Parametric spectral estimation of cardiorespiratory signals

The file `heart_1.dat` contains three physiological signals: RR-interval, arterial pressure and respiration, in this order (4 Hz sampling frequency). The recordings were made in a patient at rest. In the spectral domain, this RR-interval signal is expected to show a peak around 0.08 Hz, corresponding to the baroreflex (regulation between arterial blood pressure and heart rate), and a peak at 0.225 Hz, corresponding to the breathing rate (heart rhythm modulation by respiration, the so-called "sinus arrhythmia") – see `readme_heart.txt` for more details. In contrast, file `heart_2.dat` contains recordings from the same individual after alcohol ingestion.

a) Extract the three signals of `heart_1.dat`, remove their mean values, and subsample by a factor of 4 (i.e. to 1 Hz sampling frequency). Perform a parametric spectral estimation of all signals (`AR_psd`), trying different AR orders (5, 15, 25), and plot the results. How does the order affect the spectra? Given the above description of the expected peaks for the RR-interval, what is the best order choice?

b) Now perform a non-parametric spectral estimation on the same signals, using Welch's method (`pwelch`; for example: `pwelch(x,100,[],0:2.5e-4:0.5,1)`). How do the spectra compare with the parametric approach (order 15)? Given the previous RR-interval description, which approach performs most reliably?

c) Now extract the signals from `heart_2.dat` and perform the same processing steps as in a), to then obtain a parametric spectral estimation (order 15). Which of the physiological mechanisms described for the RR-interval was most disrupted by alcohol? And what other changes do you observe in the physiological signals?

Experiment 2: Pisarenko harmonic estimation of brain signals

The file `EEG_stim.dat` contains a matrix with three columns, corresponding to three recordings of scalp electroencephalography (EEG; left frontal electrode, 4-s duration, 512 Hz sampling frequency) from a Parkinsonian patient. The patient was also implanted with a deep brain stimulation (DBS) electrode. The 1st column corresponds to a baseline (rest) state, the 2nd was obtained during 1-Hz stimulation, and the 3rd during 100-Hz stimulation.

a) Plot the three EEG segments in the time domain, with a correct time axis.

b) Apply the Pisarenko harmonic estimation approach (`pisarenko`), considering a model with 5 sinusoids, to the three recordings to get their harmonic content. Report the estimated parameters (frequencies, amplitudes, excitation variance). Which stimulation frequency appears to effectively influence brain activity?

Experiment 3: Estimating the effects of fatigue on muscle signals

The files `emg1.dat` and `emg2.dat` contain three surface electromyography (sEMG) signals each (1024 Hz sampling frequency). These signals were recorded from the thigh muscle during a 10-min exercise on an ergocycle (average to high intensity), respectively from an athlete (`emg1`) and from a moderately trained subject (`emg2`). In each file, the columns correspond to recordings from the start, middle and end of the exercise, respectively.

a) For each subject, plot the three sEMG segments in the time domain, with a correct time axis. You can observe four bursts in each recording, which correspond to four pedal

strokes. The goal of the experiment was to assess the effects of fatigue on the signal properties of the sEMG.

b) Consider, for example, the start sEMG recording of the athlete. In principle, it would be more correct to segment each burst and analyze it separately, but it turns out that using the entire recording will yield a relatively similar spectral profile. Check that it is the case: obtain a parametric power spectral density (PSD) estimate for the entire signal and for the first burst alone (`AR_psd`, order of 30, max. frequency for display at 500 Hz), and plot them together. How do their profiles compare? Where does the difference in peak amplitude come from?

c) Remove the mean from each signal, compute their PSD, and plot all PSDs (the three of each subject can be combined in a single plot). Additionally, from each PSD, compute the mean frequency using `f' * Px / sum(Px)`. How does this formula work? Based on the PSD plots and the mean frequencies, what marks the difference between the athlete and the moderately trained subject?

Experiment 4: Assessment of PPG signal structure

Photoplethysmography (PPG) is a simple, low-cost, non-invasive optical monitoring technique that can be used to detect volumetric changes in blood in the peripheral circulation, by measuring on the skin surface. It can be integrated in wearable devices (e.g. a watch) to monitor a person's heart rate throughout daily life, and detect potential issues like atrial fibrillation. The file `PPG.dat` contains a matrix where the 1st column is a PPG recording from a normal cardiac rhythm, and the 2nd column is from a subject in atrial fibrillation (both sampled at 21.3 Hz).

a) Plot the two PPG recordings, with a correct time axis. Which one looks more regular?

After removing the mean value from each signal, quantify the difference in structure between the two signals based on:

b) The optimal AR order (`AR_order`);

c) The ratio between the excitation variance (`aryule`) and signal variance;

d) The power spectral densities (`AR_psd`, order 20);

e) The entropy of the PSDs, using: `Pn = Px / sum(Px); S = -Pn' * log(Pn);`
(the entropy can be seen as a measure of organization)

f) What do these measures (b–e) tell us about the two signals? Interpret what you obtained in terms of signal structure/organization.

Have a good session, and don't hesitate to ask questions!