02471 Machine Learning for Signal Processing

Problem set 2

This problem set is based on the teaching material and exercises from week 2–7.

You are free to use any code that has been handed out during the course, Try to work on these problems using your own code though as the code you write in this homework will make your life easier down the road.

Hand in **one pdf** with your answers. This problem set will be approved/not approved problem set, and you can consider this as a mid-way test. Late submission will be considered as failed. All problems are weighted equally and you must obtain a score of at least 50% to get it approved.

It is fine to work in small groups (max 3 people), in that case, clearly state when you hand the group composition. All answers must be submitted individually and everyone are accountable for the entire solution.

Notation

- $r(k)$ refers to a correlation between two signal (called cross-correlation), or the correlation between the signal itself (called auto-correlation). The function $\hat{r}(k)$ denotes the estimate of $r(k)$ from data.
- $r_{xy}(k) = \mathbb{E}[x(n)y(n-k)] = \mathbb{E}[\mathrm{x}_n\mathrm{y}_{n-k}]$ refers to the (unnormalized) cross-correlation between the two stochastic processes $x(n)$ and $y(n)$. Be aware that we use the unnormalized version in this course.
- $r_x(k) = \mathbb{E}[x(n)x(n-k)] = \mathbb{E}[\mathrm{x}_n\mathrm{x}_{n-k}]$ refers to the (unnormalized) autocorrelation of $x(n)$. Be aware that we use the unnormalized version in this course.
- x refers to a scalar, $\boldsymbol{x}$ refers to a vector, and X refers to a matrix.
- x refers to a random variable, $\mathbf{x}$ refers to a random vector, and X refers to a random matrix. $\hat{\mathrm{x}}$ refers to the estimate of the variable x.

Loading data

The data is stored in matlab files, and can be loaded in python using scipy loadmat, see here: https://docs.scipy.org/doc/scipy/reference/tutorial/io.html

Problem 2.1 Signal Processing is Linear Algebra

As have described in class, signal processing can be thought of as linear algebra, and in particular the filtering operation. Consider the case where you have a discrete time signal, x_n of length N, where n is the time index, and a linear FIR filter of length L, with the coefficients, θ_l.

Ordinarily, the output would be calculated as a convolution between the input signal and filter weights, but it can also be described as a matrix product, $\boldsymbol{y} = X\boldsymbol{\theta}$. Describe exactly how you would construct X and $\boldsymbol{\theta}$ in order to compute $\boldsymbol{y}$.

Problem 2.2 Biased versus unbiased parameter estimation

Consider the estimation task of a parameter to a model, $\hat{\theta} = f(\mathbf{y}, X)$, where $\mathbf{y}$ is the output vector, X is our input matrix, θ_o is the true parameter and $\hat{\theta}$ denotes the random variable of the associated estimator.

First show that the mean-square error, MSE $= \mathbb{E}[(\hat{\theta} - \theta_o)^2]$ can be decomposed into the bias and the variance.

Next, argue whether a biased or unbiased estimator is best suited for minimizing the MSE. Include formulas as you see necessary to support your analysis.

Problem 2.3 Convexity of norms

An important property that has been used is the convexity of the mean squared loss if a linear model is used. That means we can assume that minimizer of the function is also a global minimizer. A related and important result is that a function that is a norm is also a convex function.

A function, $\|\cdot\|_p : \mathbb{R}^l \longmapsto [0, \infty)$ is a valid norm if the following properties hold:

1. $\|\boldsymbol{\theta}\|_p \geq 0, \quad \|\boldsymbol{\theta}\|_p = 0 \Leftrightarrow \boldsymbol{\theta} = \mathbf{0}$.
2. $\|\alpha\boldsymbol{\theta}\|_p = |\alpha|\|\boldsymbol{\theta}\|_p, \forall \alpha \in \mathbb{R}$.
3. $\|\boldsymbol{\theta}_1 + \boldsymbol{\theta}_2\|_p \leq \|\boldsymbol{\theta}_1\|_p + \|\boldsymbol{\theta}_2\|_p$.

By using the definitions in section 8.2 of a convex function in the book, prove that any function that is a norm is also a convex function.

Problem 2.4 Correlation functions

Consider a stochastic signal $\mathrm{x}_n = a\mathrm{x}_{n-1} + \mathrm{v}_n$, that is generated as a 1st order auto-regressive process (AR(1)), with the auto-regressive coefficient $a \in]-1, 1[$, and a white noise signal with variance σ_v^2. A new signal $\mathrm{y}_n = \mathrm{x}_n + b$ is created, where $b \in \mathbb{R}$ and $\mathbb{E}[\mathrm{x}_n] = 0$. Determine the auto-correlation function $r_y(k)$, and compute $r_y(-2)$ for $a = 0.8$, $b = 0.5$ and $\sigma_v^2 = 1$.

Note: there is some inconsistency on how the correlation function is defined in various textbook. We use the unnormalized version as defined in eq. (2.101), section 2.4.1 in the book.

Problem 2.5 Wiener filtering

In this problem, we will consider a FIR Wiener filter setup and use the filter to remove noise. Explain what the desired signal and the error signal of a filter setup is.

Load the signals `problem2_5_signal.mat` and `problem2_5_noise.mat`, and denote these signal s_n and w_n respectively. The signals were sampled using a sampling frequency of 8 kHz. Create a new signal $x_n = s_n + w_n$, and consider x_n as input to the filter.

Estimate a Wiener filter with an appropriate length, and argue for your choice of length. Plot the frequency responses of the signals and the filter, and evaluate the results.

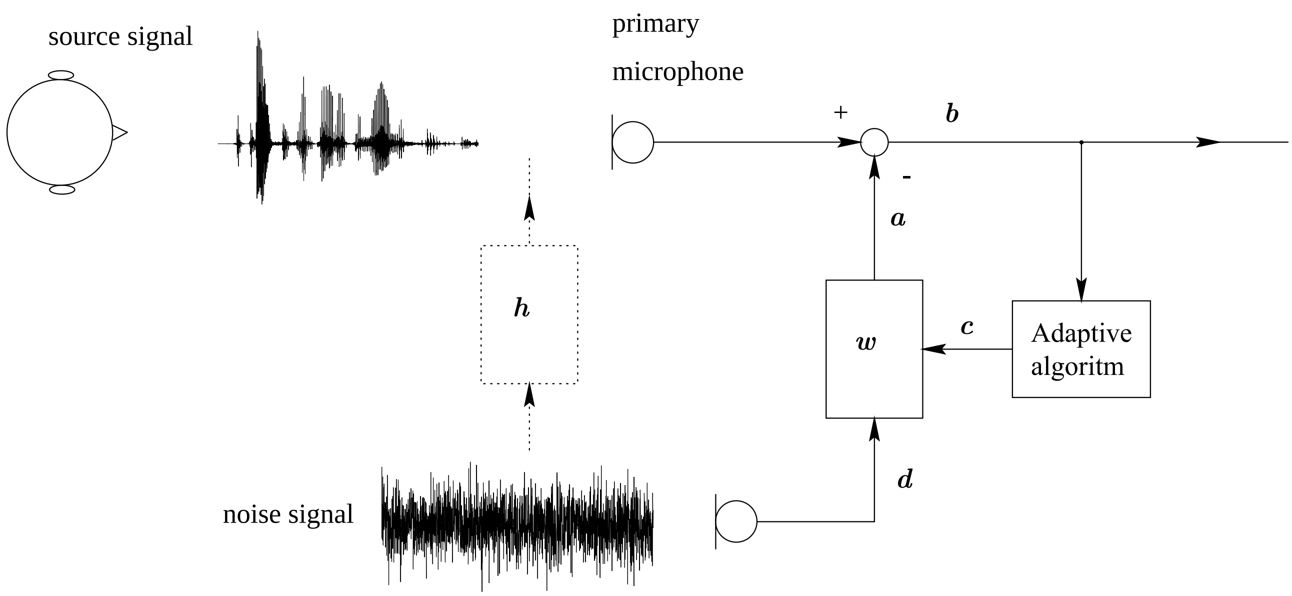
Problem 2.6 Time-frequency analysis of audio

The problem set contains the sound-file `problem2_6.wav`. This is the sound of a synthesizer from the 1980s that used to be very hip. The drumbeat in this sound is composed out of different "instruments". Each one is distinct from the others in its spectral composition and you should be able to tell them apart by listening to it.

Conduct a time-frequency analysis of the signal by plotting the spectrogram of this recording. Use a window size, window overlap, and window function of your choice, and briefly argue for the choices, e.g by relating the time between the beats from the spectrogram and identify the spectral signature of the different "instruments".

Problem 2.7 Adaptive filtering

Consider the following filtering setup



On the figure there are six signals denoted a, b, c, d, h, and w. Identify the signals and describe them in terms relating them to a filter setup.

Describe the functional principle of the noise canceling setup. Why is it important to assume that the source signal is uncorrelated with the noise?

Load the speech signal `problem2_7_speech.mat`, and generate a noise signal with a suitable variance (perform a listening test, if possible). Use 8 kHz as the sampling frequency. Filter the noise signal through one of the filters found in `problem2_7_lpir.mat`, `problem2_7_hpir.mat`, and `problem2_7_bpir.mat`, and add the filtered noise signal to the speech signal. Use the LMS, NLMS and RLS algorithm to build an adaptive noise canceling system. Evaluate the performance of the system and comment on your choices and findings.