Statistical Signal & Data Processing - COM500
Final Exam

Wednesday June 30, 2021, 8h15-11h15.

You will hand in this sheet together with your solutions.

Write your personal data (please make it readable!).

Seat Number:

Family Name:

Name:

Read Me First!

You are allowed to use:

- A handwritten cheatsheet (two A4 sheets, double sided) summarizing the most important formulas (no exercise text or exercise solutions);
- A pocket calculator.

You are definitively not allowed to use:

- Any kind of support not mentioned above;
- Your neighbor ; Any kind of communication systems (smartphones etc.) or laptops;
- Printed material; Text and Solutions of exercises/problems; Lecture notes or slides.

Write solutions on separate sheets, *i.e.* no more than one solution per paper sheet.

Return your sheets ordered according to problem (solution) numbering.

Return the text of the exam.

All the best for your exam!!

Warmup Exercise

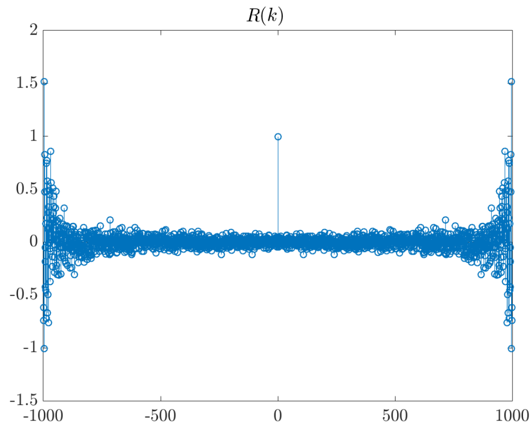
This is a warm up problem .. do not spend too much time on it. Please provide justified, rigorous, and simple answers. If needed, you can add assumptions to the problem setup.

Exercise 1. Correlation (2 pts)

Let $W[n]$ be a centered white noise with $\sigma^2 = 1$, taking real values. Given that the process is i.i.d. and centered, we know that, theoretically,

$$R(k) = \mathrm{E}\left[W[k+n]W[n]\right] = \begin{cases} \sigma^2 = 1 & k = 0\,; \\ 0 & k \neq 0\,; \end{cases}$$

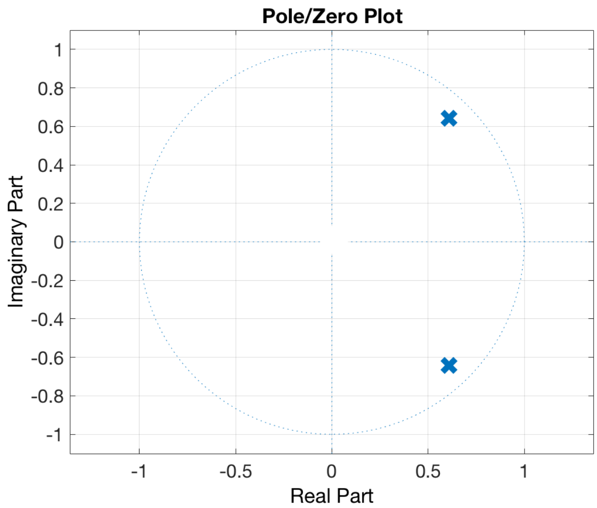
We have measured $N = 1000$ samples of the noise $w[1], \ldots, w[1000]$ and then we have computed the correlation $R(k)$, $k = -999, \ldots, 0, \ldots, 999$. Here's the plot of the correlation.



Can you tell if the plotted correlation has been computed using the empirical un-biased correlation or the empirical biased correlation? Precisely justify your answer. Provide the expression of both correlations.

Exercise 2. A simple system (4 pts)

The figure below depicts the poles of a causal LTI system $P(z)$. the two poles have magnitude $a = 0.9$ and phase $\varphi = \pm\pi/4$. $P(z)$ has also two zeros at the origin (not plotted).



1) Is the system $P(z)$ stable?

2) Sketch the magnitude of the transfer function $|P(e^{jw})|$.

3) Is the inverse system $H(z) = 1/P(z)$ stable?

4) Give the impulse response $h(n)$ of $H(z)$.

5) Can $H(z)$ be the synthesis filter of an autoregressive process?

Main Problems

Here comes the core part of the exam .. take time to read the introduction and each problem statement. Please provide justified, rigorous, and simple answers. Remember that you are not simply asked to describe statistical signal processing tools, but rather to describe how to apply such tools to the given problem. If needed, you can add assumptions to the problem setup. You are asked to comply with the notation given in the problem.

Exercise 3. MOUNTAIN BIKE VIBRATIONS

Despite being equipped with a double suspension system (front and rear) even high tech competition mountain bikes cannot completely attenuate vibrations on irregular terrains.

Vibrations causes the athlete to experience uncomfort and loss of energy and efficiency.

Attenuating vibrations is a complex control problem that requires an adaptive adjustment of the suspension and a trade-off between shock absorption and effective transfer of power from the athlete to the bike wheels.

In order to cope with such a complex control problem we need to:

- Analyze the vibration data;
- Classify the terrains, via the classification of the vibrations;
- Implement the adaptive control system.

The vibration data is collected using sensors (accelerometers) placed on the bike, while the adaptive control is integrated into the suspensions.

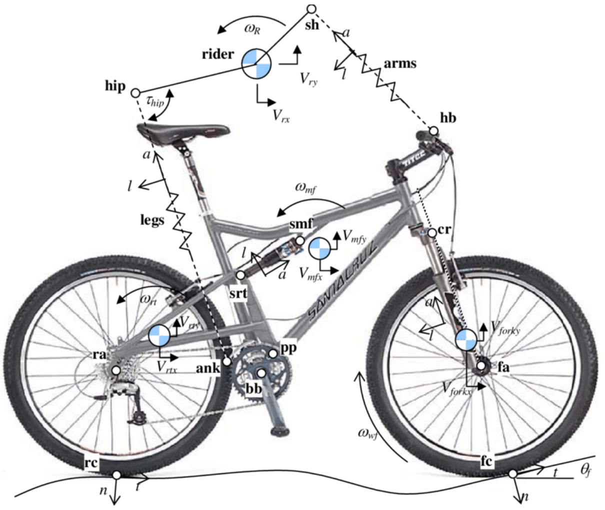


Fig. 1: MTB Vibration Model.

Fig. 2: Competition MTB with Sensor Positions.

As depicted in the figure above (right), vibrations are measure by 4 accelerometers placed in 4 positions on the bike: Handlebar (#1); Fork (#2); Bottom bracket (# 3); Rear axle (# 4). We shall call the corresponding measured signals $x_k[n]$, $k = 1, 2, 3, 4$.

The following parts A, B, and C are independent.

Part A: Analysis of Vibrations (18 points)

The figure below depicts a bit more than 3000 samples of the root mean square RMS of four signals $x_k[n]$, $k = 1, 2, 3, 4$. We shall note the RMS of the four signals as $s_k[n]$, $k = 1, 2, 3, 4$, respectively.

Such samples are obtained with a sampling frequency of 1 kHz.

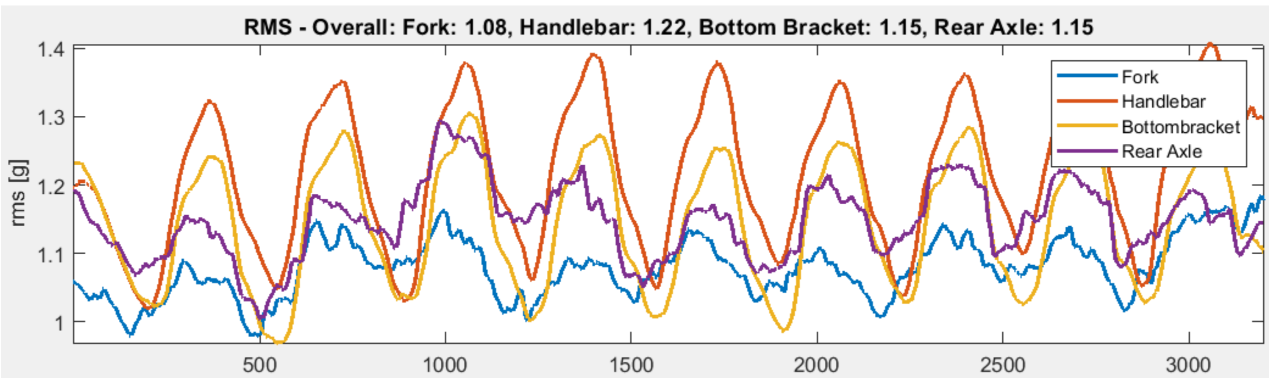


Fig. 3: RMS (Root Mean Square) of the acceleration measured by the 4 sensors.

We can see that the signal are periodic.

- The RMS $s_1[n]$ of the accelerometer 1 (handlebar) and the RMS $s_3[n]$ of the accelerometer 3 (bottom bracket) can each be easily approximated as a sinusoid.
- The RMS $s_2[n]$of the accelerometer 2 (fork) and the RMS $s_4[n]$ of the accelerometer 4 (rear axle) can each be approximated as the sum of two sinusoid: One at the fundamental frequency f_1, and the other at a higher frequency f_2.

Given the short window of time, $s_k[n]$, $k = 1, 2, 3, 4$ can be assumed to be w.s.s..

A.1) Provide w.s.s. stochastic models $S_k[n]$ $k = 1, 2, 3, 4$ for each one of the 4 RMS signals $s_k[n]$, $k = 1, 2, 3, 4$, and precisely prove that each stochastic model is indeed a w.s.s. process. Please provide justified, rigorous, and simple proof.

We now focus on the RMS signal of the sensor 2 (fork) and the RMS signal of the sensor 4 (rear axle), for which we shall consider the first 3000 samples of the above picture.

Remember that such signals can each be approximated as the sum of two sinusoid, where one sinusoid has fondamental frequency, and the other has a higher frequency.

Let's compute the periodogram of the RMS signals of sensor 2 and 4.

A.2) By computing the periodogram of $s_2[n]$ and $s_4[n]$, $n = 1, \ldots, 3000$, which are the conditions to be able to distinguish the fundamental frequency f_1 and the higher frequency f_2? Notice that these frequencies are not given and that you are not asked to compute them. Please provide a justified, rigorous, and simple answer.

We shall improve our model for the two RMS signals by considering an additive Gaussian white noise, *i.e.*,

$$Y_k[n] = S_k[n] + W[n] , \quad k = 2, 4, \quad n = 1, \ldots, 3000 .$$

A.3) Accounting for the presence of noise, propose a parametric method to estimate the two frequencies of $Y_2[n]$ and $Y_4[n]$. Precisely describe such method. You are given $N = 3000$ samples (be careful that they are not a lot). You are asked to detail each step as if you have to implement the method in a computer. Precisely indicate the input and output of each step.

A.4) Given that we can assume $Y_2[n]$ and $Y_4[n]$ to have the same frequencies, how can you exploit such assumption to improve the estimation of the two frequencies? Please justify your answer.

Part B: Classification of the Terrains (16 points)

From vibration measurements we would like to classify the type of terrain. Such a classification is of foremost importance: Based on the type of terrain we can adapt the fork configuration to improve the riding comfort. First of all, we need to understand how to characterises a terrain based on the vibration measurements.

The four acceleration signals $x_k[n]$, $k = 1, 2, 3, 4$, sampled at 1 kHz, are recorded over intervals of 1 s. The athlete ride the bike for about 30 minutes and we obtain $K = 2000$ of such recordings (2000 intervals). For each recording $k = 1, \ldots, 2000$, the four signals are combined to extract 5 characteristic variables:

- Mean $mn[k]$;
- Maximum value $mx[k]$;
- Median $md[k]$;

- Minimum value $mi[k]$;
- Energy $en[k]$.

B.1) Describe in detail, step by step, how to compute the 5 principal components given the variables $\boldsymbol{m}[k] = [mn[k], mx[k], md[k], mi[k], en[k]]$, $k = 1, \ldots, 2000$. We shall denote the 5 principal components as $c_1[k], \ldots, c_5[k]$, $k = 1, \ldots, 2000$. Each step should be able to be interpreted and executed by a computer (in particular the input, the executed operation with corresponding equations, and the output of each step has to be clear). Also, **clearly indicate the dimensions of the matrices and vectors**.

After analyzing the variance of the principal components, it clearly appears that only 2 principal components, namely $c_1[k]$ and $c_2[k]$, $k = 1, \ldots, 2000$ account for most of the total sample variation.

By looking at the 2D plot of the 2 principal components we can isolate 5 different clusters.

B.2) Provide a Gaussian Mixture model describing the 5 clusters based on the principal components, namely $c_1[k]$ and $c_2[k]$, $k = 1, \ldots, 2000$. More precisely, provide its cumulative distribution function and do not forget that the cluster plot has a total of 2000 points!

B.3) Write the corresponding likelihood function clearly indicating which are its parameters.

Finally we can conclude that the recordings show 5 different types of terrains. We can imagine that such types are flat, ondulating, lightly bumpy, bumpy, very bumpy.

Part C: Adaptive Control System (14 points)

Consider the acceleration signal $x_1[n]$ of the sensor in the handlebar, and the acceleration signal $x_2[n]$ of the sensor in the fork. We want to develop and adaptive control system that by adjusting the fork reduces the vibrations in the handlebar, especially those at high frequencies.

To start with we can consider the signal measured in the handlebar (without adaptive control system) to be modelled as

$$x_1[n] = v[n] + (h * x_2)[n]\,,$$

where

- $v[n]$ is low frequency vibration signal accounting for the very smooth variations of the terrain, not affecting the adaptive control system;
- $x_2[n]$ is the vibration signal accounting for the abrupt variations of the terrain, measured by the sensor in the fork
- $h[n]$ is the impulse response accounting for the inertia of the fork (weight and frictions of moving parts).

C.1) By modelling the fork adaptive control system as an adaptive filter g_n, provide the scheme (draw a block diagram) of the adaptive filter capable of reducing the vibrations at the handlebar. Show the signal that is used for adaptation and give the quantity $J(g_n)$ that the adaptive filter minimizes.

C.2) When implementing the adaptive control system (as adaptive filter), which are the parameters to be specified in order to ensure the convergence of the algorithm?

C.3) Can we use the Stochastic Gradient Descent algorithm (LS algorithm with reduced computational burden) for such an adaptive control system? Precisely justify your answer.

C.4) (Bonus) What do you think is the limitating factor in applying such an adaptive approach (Hint: Think of the inertia of the fork).