# Statistical Signal & Data Processing - COM500
# Final Exam

Friday July 1st, 2022, 9h15-12h15.

**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 (3 PTS)
We have recorded $N = 200000$ samples $x[1], \ldots, x[200000]$ of a w.s.s. signal. Using the empirical mean $\widehat{m}$, we observe that the mean of the signal is **clearly not zero**, *i.e.*, $|\widehat{m}| \gg 0$.

We would like to compute the empirical correlation $\widehat{R}_X[k]$ for $k = 0, 1, \ldots, 4$.

Which form of the empirical correlation should we use? (*Only one answer is correct, and you have to justify it precisely*)

- Absolutely only the unbiased correlation!
- Absolutely only the biased correlation!
- Either one of the two, it does not make a relevant difference!

Write the chosen expression(s) of the empirical correlation, given the samples $x[1], \ldots, x[200000]$ (both if you choose "either one of the two").

**Solution 1**.

- Biased correlation $\widehat{R}_b(k) = \dfrac{1}{N} \displaystyle\sum_{n=1}^{N-k} x[n+k]x[n]^*$;
- Unbiased correlation $\widehat{R}_u(k) = \dfrac{1}{N-k} \displaystyle\sum_{n=1}^{N-k} x[n+k]x[n]^*$.

The two equations hold for $k \geq 0$, and we set $\widehat{R}(-k) = \widehat{R}_b^*(k)$.
Give that $N = 100000$ and $k = 0, 1, \ldots, 4$, we have $\dfrac{1}{N} \approx \dfrac{1}{N-k}$, consequently the two estimates of the correlation provide values that do not preset a relevant difference.

**Exercise 2**. MARKOV CHAIN (3 PTS)
Let $X[n]$ be a Markov chain with 3 possible states. We saw that a Markov chain is a parametric signal model where the parameters are probabilities, *i.e.*, in $[0, 1]$.
How many of the model parameters can be freely defined (in $[0, 1]$)?

**Solution 2**.
A 3-state Markov chain is characterised by:

- 3 initial probabilities $\pi_i$, $i = 1, \ldots, 3$.
- 9 transition probabilities $p_{ij}$, $i = 1, \ldots, 3$, $j = 1, \ldots, 3$.

These 12 parameters cannot be all freely defined. Indeed they are constrained:

- 1 constraint for the initial probabilities $\sum_{i=1}^{3} \pi_i$.
- 3 constraints for the transition probabilities $\sum_{j=1}^{3} p_{ij}$, $i = 1, \ldots, 3$.

So finally we can freely define 8 parameters (2 of the 3 transition probabilities and 6 of the 9 transition probabilities).

If in addition we consider the constraint on the stationarity of the Markov chain, namely $\sum_{i=1}^{3} p_{ij}\pi_i = \pi_j$, $j = 1, \ldots, 3$, we can freely define only 5 parameters.

# 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**. Preventive Maintenance of Road Pavement

Preventive maintenance of road infrastructures is of foremost importance in order to guarantee road safety and to optimally cope with infrastructure ageing and fatigue, therefore limiting costs and enabling planification of maintenance interventions and budget.

The midterm exam discussed the preventive maintenance of particular road infrastructures, namely bridges. We saw that an advanced non-destructive inspection method is based on vibration measurements obtained by an accelerometer.

Here we will focus on another particular road infrastructure, namely, road pavement made of asphalt. In order to monitor fatigue damages and to detect abnormalities, the characterization of the pavement response to pressure is needed. This can be done by inserting a strain gauge (Fig. 1) into the asphalt and by measuring the pavement modifications when pressure is applied on it.

A typical strain gauge response to the passage of a truck is depicted in Fig. 2. We shall denote with $x[n]$ the samples of the strain gauge response, *i.e.*, the samples of the pressure measurement signal.

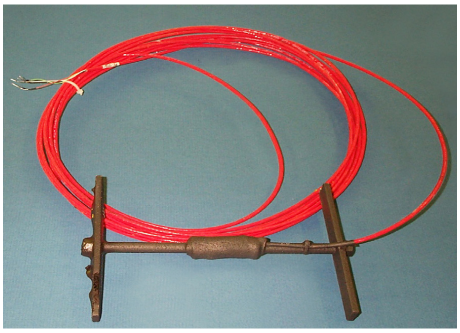


*Fig. 1: Strain gauge.*

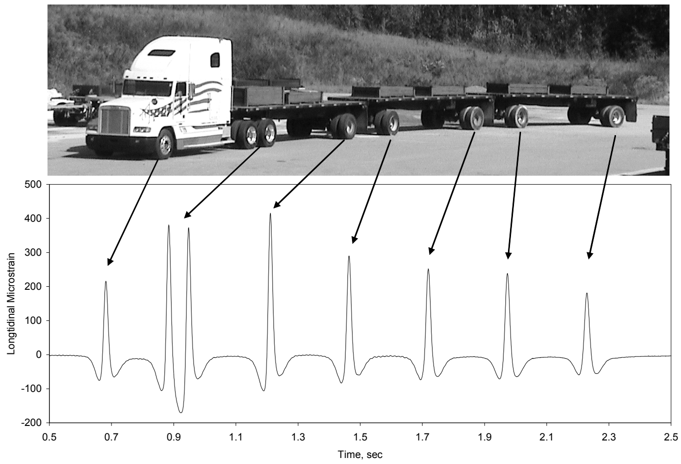



*Fig. 2: Strain gauge response to the passage of a truck.*

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

### Part A: Analysis of Pressure Measurements (10 points)

We start by focusing on the analysis of the pressure measurement signal.

Fig. 1 shows that, over a defined finite time interval $T$, the strain gauge signal is characterized by a countable number of events. Call $\alpha$ that average number of events per unit of time.

A.1) Which stochastic process best describes a countable number of events on a finite time interval $T$?

A.2) Based on such a model, what is the probability of $k$ events in the interval $[0, T]$?

Denote with $x[n]$ the samples of the strain gauge measurement. Suppose we can approximate it to a continuous time signal $x(t)$ corresponding to Dirac pseudo-functions. We would like to estimate the times at which pressure events have occurred and the amplitudes of such events.

A.3) Propose a parametric method to estimate the times at which pressure events have occurred and the amplitudes of the corresponding strain gauge measurements. Precisely describe such method. We consider a recording of the signal $x(t)$ over the time interval $[0, T]$. We assume that over that interval a maximum of $M$ pressure events might occur. You are asked to detail each step as if you have to implement the method on a computer. Precisely indicate the input and output of each step.

**Part B: Characterization of Pressure Events (25 points)**

The strain gauge signal $x[n]$ has been recorded during a year on a road where the traffic is limited to cars. Here each pressure event is represented by a pulse with a certain width and amplitude (like in Fig. 2).

The passage of a car corresponds to two pressure events (front and rear axles/tires). A total of $K = 100'000$ car passages have been identified. The two pressure events corresponding to each passage have been characterized using the following variables:

- Average intensity of the two pressure events $i_a[k]$;
- Intensity difference of the two pressure events $i_d[k]$;
- Time interval between the two pressure events $t_i[k]$;
- Average width of the pressure pulses of the two pressure events $w_a[k]$;
- Width difference of the pressure pulses of the two pressure events $w_d[k]$;

B.1) Describe in detail, step by step, how to compute the principal components given the variables $\boldsymbol{v}[k] = [i_a[k], i_d[k], t_i[k], w_a[k], w_d[k]]$, $k = 1, \ldots, 100'000$. We shall denote the principal components as $p_1[k], \ldots, p_5[k]$, $k = 1, \ldots, 100'000$. 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**.

B.2) What do the principal components represent?

After analyzing the variance of the principal components, it clearly appears that 2 principal components, namely $p_1[k]$ and $p_2[k]$, $k = 1, \ldots, 100'000$ account for most of the total sample variation.

B.3) What does it mean, in relation to the variables $\boldsymbol{v}$, that only 2 principal components account for most of the total sample variation?

By looking at the 2D plot of the 2 principal components we can isolate 3 different clusters.

B.4) What does it mean, in relation to the variables $\boldsymbol{v}$, that the 2D plot of the 2 principal components presents 3 different clusters?

B.5) Provide a Gaussian Mixture model based on the principal components $p_1[k]$ and $p_2[k]$, $k = 1, \ldots, 100'000$, describing the 3 clusters. More precisely, provide its cumulative distribution function and do not forget that the cluster plot has a total of 100′000 points!

B.6) Propose a method (seen in class) to estimate the center of the clusters. Precisely describe such method. You are given the realizations $p_1[k]$ and $p_2[k]$, $k = 1, \ldots, 100'000$ of the principal components. You are asked to detail each step as if you have to implement the method on a computer. Precisely indicate the input and output of each step.

Call $m_1$, $m_2$, and $m_3$, respectively, the centers of the 3 clusters.

B.7) What do these centers represent, in relation to the variables $\boldsymbol{v}$?

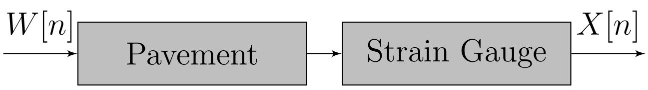
B.8) Now that you have the centers of the 3 clusters, let's go back to the variable $\boldsymbol{v}$ space. How can you do so? Write the corresponding equations.

B.9) You have now obtained values variable $\boldsymbol{v}$ space. Given that we started with 100′000 measurements of the variables $\boldsymbol{v}[k] = [i_a[k], i_d[k], t_i[k], w_a[k], w_d[k]]$, how can you interpret the values obtained in B.8)? What do they represent?

## Part C: Analysis of the Pavement Response (15 points)

The analysis of the pavement response to specific pressure patterns provides insight on the pavement structure and on its ageing behaviour.

The specific pressure pattern considered here is a white noise signal $W[n]$. The pavement response is obtained by measuring the strain gauge signal $X[n]$.




We collect a very large amount $N = 1'000'000$ of samples $x[n]$, $n = 1, \ldots, 1'000'000$ of $X[n]$, and we adopt an advanced periodogram estimation method to obtain a very low variance spectrum, as depicted in Fig. 3.

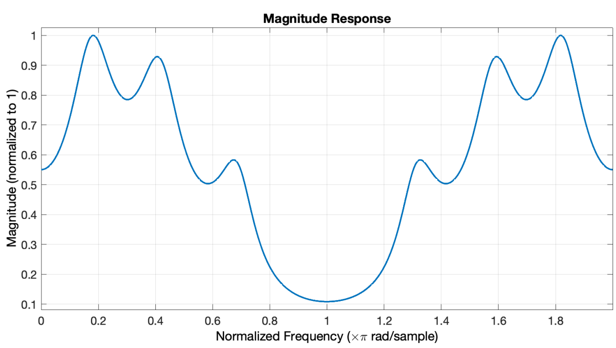
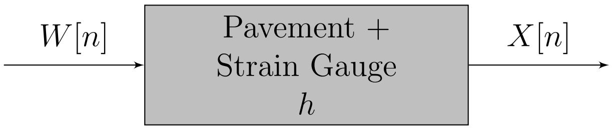



*Fig. 3: Spectrum of the strain gauge signal in response to a white noise pressure pattern.*

We can model the pavement and the strain gauge as a system with a specific impulse response $h$, and therefore a $z$ transform $H(z)$. Assuming the pavement does not have critical resonance modes, the impulse response $h$ can be taken to be absolutely summable, *i.e.*, $h \in l_1$.




C.1) Given the spectrum above, propose a parametric method to estimate the coefficients of the $z$ transform $H(z)$. Precisely describe such method. You are given the realizations $x[n]$, $n = 1, \ldots, 1'000'000$ of $X[n]$. You are asked to detail each step as if you have to implement the method on a computer. Precisely indicate the input and output of each step.

C.2) The method you have just described requires a variable to be defined, namely the order of the model. Can you estimate the order of the model?

C.3) Is the signal $X[n]$ a w.s.s. process? Justify precisely your answer.

C.4) What kind of process is $X[n]$? Justify precisely your answer.

**Solution 3**.

**Part A: Analysis of Pressure Measurements (10 points)**

Fig. 1 shows that, over a defined finite time interval $T$, the strain gauge signal is characterized by a countable number of events. Call $\alpha$ that average number of events per unit of time.

A.1) A (random) countable number of events on a finite interval $(0, T]$ are best described by a Poisson process $N((0, T])$ defined as follows

- Given the average number of events per unit of time $\alpha$, the probability of number of events in $(0, T]$ is
$$\mathrm{P}\left(N((0, T]) = k\right) = \frac{(\alpha T)^k e^{-\alpha T}}{k!} \, ;$$
- Given two disjoint intervals $(t_1, t_2]$ and $(t_3, t_4]$, that is, two not overlapping intervals $(t_1 < t_2 < t_3 < t_4)$, the number of arrivals $N((t_1, t_2])$ is independent of the number of arrivals $N((t_3, t_4])$.

A.2) Given the average number of events per unit of time $\alpha$, the probability of $k$ events in the interval $[0, T]$ (or equivalently $(0, T]$) is given by

$$\mathrm{P}\left(N((0, T]) = k\right) = \frac{(\alpha T)^k e^{-\alpha T}}{k!} \, ;$$

A.3) Given the *a priori* absence of noise, and the fact the the pressure signal has been modelled as a sequence of Dirac pseudo fonctions, the optimal parametric method to estimate the positions of the pressure events is the annihilating filter method (simple and computationally efficient). Over a period of $T$ s, the given maximum number of received pulses is $M$. We shall denote the time instant of the pressure events as $\tau_k$, $k = 1, \ldots, M$.

- We have $x(t)$ for $0 \le t \le T$ s
- Recalling that the annihilating filter works on harmonic signals, we first need to transform the sequences of Deltas into a harmonic signal by taking the Fourier transformation (Fourier series) of $x(t)$ (considered periodic with a period of $T$ s)

$$\widehat{x}[n] = \frac{1}{T}\int_0^T x(t)e^{-j2\pi n\frac{t}{T}} = \frac{1}{T}\sum_{k=1}^{M} \alpha_k e^{-j2\pi n\frac{\tau_k}{T}} .$$

- Targeting the estimation of the position of $M$ spikes, we need an annihilating filter with impulse response of length $M$. The corresponding system reads

$$\begin{bmatrix} \widehat{x}[M-1] & \dots & \widehat{x}[0] \\ \vdots & \ddots & \vdots \\ \widehat{x}[2(M-1)] & \dots & \widehat{x}[M-1] \end{bmatrix} \begin{bmatrix} h[1] \\ \vdots \\ h[M] \end{bmatrix} = - \begin{bmatrix} \widehat{x}[M] \\ \vdots \\ \widehat{x}[2M-1] \end{bmatrix} .$$

  By solving the system we obtain $h[1], \dots, h[M]$ (Toeplitz system, requiring $M^2$ multiplications), and therefore, with $h[0] = 1$, we have the impulse response of the annihilating filter.
- Having the impulse response $h[n]$ we compute the z-transform

$$H(z) = \sum_{n=0}^{M} h[n] = 1 + h[1]z^{-1} + \dots + h[M]z^{-M} .$$

- Compute the zeros of the z-transform $H(z)$, that we shall call $z_1, \dots, z_M$, that is

$$H(z) = \prod_{k=1}^{M}(1 - z_k z^{-k}) .$$

- By taking the argument of the zeros we obtain the positions $\tau_k$, $k = 1, \dots, M$, with the following formula

$$\tau_k = T\frac{\arg(z_k)}{2\pi} ,$$

  where $\arg(z_k)$ is constrained in $[0, 2\pi]$.

**Part B: Characterization of Pressure Events (25 points)**

B.1) We have $100'000$ realizations (data size) of 5 variables, *i.e.*, $\boldsymbol{v}[k] = [i_a[k], i_d[k], t_i[k], w_a[k], w_d[k]]$ (dimensions $1 \times 5$), $k = 1, \dots, 100'000$.

The principal components are denoted as $\boldsymbol{p}[k] = [p_1[k], \dots, p_5[k]]$, $k = 1, \dots, 100'000$.

*Principal Components*

- Given $\boldsymbol{v}[k] = [i_a[k], i_d[k], t_i[k], w_a[k], w_d[k]]$ (dimensions $1 \times 5$), $k = 1, \dots, 100'000$, compute the mean of the variables

$$\boldsymbol{m}_v = \frac{1}{100'000}\sum_{k=1}^{100'000} \boldsymbol{v}[k] , \quad \text{(dimensions } 1 \times 5) .$$

– Given the mean $\boldsymbol{m}_v$, center the variables

$$\bar{\boldsymbol{v}}[k] = \boldsymbol{v}[k] - \boldsymbol{m}_v \,, \quad \text{(dimensions } 1 \times 5) \,, \quad k = 1, \ldots, 100'000 \,.$$

– Given the centred variables $\bar{\boldsymbol{v}}[k]$, $k = 1, \ldots, 100'000$, compute the empirical correlation matrix of the centred variables

$$\widehat{\boldsymbol{R}}_v = \frac{1}{100'000} \sum_{k=1}^{100'000} \bar{\boldsymbol{v}}[k]^T \bar{\boldsymbol{v}}[k] \,, \quad \text{(dimensions } 5 \times 5) \,.$$

– Given empirical correlation matrix $\widehat{\boldsymbol{R}}$ diagonalize it

$$\boldsymbol{V}^t \widehat{\boldsymbol{R}}_v \boldsymbol{V} = \boldsymbol{\Lambda} \,, \quad \text{(dimensions } 5 \times 5) \,,$$

where $\boldsymbol{V}$ is the matrix of eigenvectors (dimensions $5 \times 5$).

– Given the the centred variables $\bar{\boldsymbol{v}}[k]$, $k = 1, \ldots, 100'000$ and the matrix of eigenvectors $\boldsymbol{V}$, compute the principal components

$$\boldsymbol{p}[k] = \bar{\boldsymbol{v}}[k] \boldsymbol{V} \,, \quad \text{(dimensions } 1 \times 5) \,, \quad k = 1, \ldots, 100'000 \,.$$

B.2) For every $k = 1, \ldots, 100'000$, the 5 principal components $\boldsymbol{p}[k]$ represents the one-to-one projection of the 5 centred variable $\bar{\boldsymbol{v}}[k]$ into a 5 dimensional orthogonal space (orthogonalisation of the centred variable)

B.3) It means that for each $k = 1, \ldots, 100'000$ the centred version $\bar{\boldsymbol{v}}[k]$ of the 5 variables $\boldsymbol{v}[k]$ are a linear combination of only 2 of the 5 principal components $\boldsymbol{p}[k]$. That is, for each $k = 1, \ldots, 100'000$, the projection of the 5 centred version $\bar{\boldsymbol{v}}[k]$ of the variables $\boldsymbol{v}[k]$ into the principal component space results in only two components.

B.4) It means that the $100'000$ car passages can be divided in 3 groups, as if there are only 3 types of car.

B.5) Provided that each cluster can be modelled as a 2 dimension Gaussian distribution, the 3 2D clusters can be modelled as a 2 dimension Gaussian mixture model. The variables of the clusters are the two principal components $\boldsymbol{p}[k] = [p_1[k], p_2[k]]$, $k = 1, \ldots, 100'000$. They can be considered at the realization of a 2D stochastic process $\boldsymbol{P}[k] = [P_1[k], P_2[k]]$. For each $k = 1, \ldots, 100'000$, the cumulative distribution reads

$$F_{\boldsymbol{P}[k]}(\boldsymbol{p}[k]) = \sum_{c[k]=1}^{3} \mathrm{P}\left(\boldsymbol{P}[k] \leq \boldsymbol{p}[k], C[k] = c[k]\right) \overset{\text{Bayes}}{=} \sum_{c[k]=1}^{3} \mathrm{P}\left(\boldsymbol{P}[k] \leq \boldsymbol{p}[k] \,|\, C[k] = c[k]\right) \mathrm{P}\left(C[k] = c[k]\right)$$

where $C[k]$ is the random variable indicating the cluster, *i.e.*, indicating which one of the 3 types of car (groups) has generated the 2 pressure events at the passage number $k$.

When we consider all the $100'000$ passages and we indicate $\boldsymbol{P} = [\boldsymbol{P}[1], \ldots, \boldsymbol{P}[100'000]]$, $\boldsymbol{p} = [\boldsymbol{p}[1], \ldots, \boldsymbol{p}[100'000]]$ and $\boldsymbol{C} = [C[1], \ldots, C[100'000]]$, we have

$$F_{\boldsymbol{P}}(\boldsymbol{p}) = \sum_{\boldsymbol{c} \in \mathcal{C}} \mathrm{P}\left(\boldsymbol{P} \leq \boldsymbol{p}, \boldsymbol{C} = \boldsymbol{c}\right) \overset{\text{Bayes}}{=} \sum_{\boldsymbol{c} \in \mathcal{C}} \mathrm{P}\left(\boldsymbol{P} \leq \boldsymbol{p} \,|\, \boldsymbol{C} = \boldsymbol{c}\right) \mathrm{P}\left(\boldsymbol{C} = \boldsymbol{c}\right)$$

where $\mathcal{C}$ represents all the possible combinations of types of car (groups) in the $100'000$ passages.

Notice that:

- We can assume that the types of car passing by, form an independent sequence (no reason to suppose the contrary), *i.e.*

$$\mathrm{P}\left(\boldsymbol{C}=\boldsymbol{c}\right)=\prod_{k=1}^{100'000}\mathrm{P}\left(C[k]=c[k]\right)\ ;$$

- Given the type of car that passed by, the two pressure events can be considered to be equal to a pair of constant values plus noise, *i.e.*

$$\mathrm{P}\left(\boldsymbol{P}\leq\boldsymbol{p}\,|\,\boldsymbol{C}=\boldsymbol{c}\right)=\prod_{k=1}^{100'000}\mathrm{P}\left(\boldsymbol{P}[k]\leq\boldsymbol{p}[k]\,|\,C[k]=c[k]\right)\ ,$$

  where $\mathrm{P}\left(\boldsymbol{P}[k]\leq\boldsymbol{p}[k]\,|\,C[k]=c[k]\right)$ is a Gaussian (bivariate) cumulative distribution. The latter admits a Gaussian (bivariate) density $\mathcal{G}_{\boldsymbol{m}[c[k]],\Sigma^2[c[k]]}(\boldsymbol{p}[k])$ with mean $\boldsymbol{m}[c[k]]=[m_1[c[k]],m_2[c[k]]]$ and covariance matrix $\Sigma^2[c[k]]$;
- Given that the principal components are independent, the covariance matrix is diagonal

$$\Sigma^2[c[k]]=\begin{bmatrix}\sigma_1^2[c[k]] & 0\\ 0 & \sigma_2^2[c[k]]\end{bmatrix}\ .$$

  Therefore, calling $\boldsymbol{\sigma}^2[c[k]]=[\sigma_1^2[c[k]],\sigma_w^2[c[k]]]$

$$\mathcal{G}_{\boldsymbol{m}[c[k]],\Sigma^2[c[k]]}(\boldsymbol{p}[k])=\mathcal{G}_{\boldsymbol{m}[c[k]],\boldsymbol{\sigma}^2[c[k]]}(\boldsymbol{p}[k])\,,$$

  and

$$\begin{aligned}\mathcal{G}_{\boldsymbol{m}[c[k]],\boldsymbol{\sigma}^2[c[k]]}(\boldsymbol{p}[k])=&\frac{1}{\sqrt{2\pi\sigma_1^2[c[k]]}}\exp\left(\frac{(p_1[k]-m_1[c[k]])^2}{\sigma_1^2[c[k]]}\right)\\&\frac{1}{\sqrt{2\pi\sigma_2^2[c[k]]}}\exp\left(\frac{(p_2[k]-m_2[c[k]])^2}{\sigma_2^2[c[k]]}\right)\end{aligned}$$

Finally, the cumulative distribution of the (random variables associated to the) two principal components $p_1[k]$ and $p_2[k]$, $k=1,\dots,100'000$, reads

$$\begin{aligned}F_{\boldsymbol{P}}(\boldsymbol{p})=&\prod_{k=1}^{100'000}\sum_{c[k]=1}^{3}\mathrm{P}\left(\boldsymbol{P}[k]\leq\boldsymbol{p}[k]\,|\,C[k]=c[k]\right)\mathrm{P}\left(C[k]=c[k]\right)\\&\prod_{k=1}^{100'000}\sum_{l=1}^{3}\mathrm{P}\left(\boldsymbol{P}[k]\leq\boldsymbol{p}[k]\,|\,C[k]=l\right)\mathrm{P}\left(C=l]\right)\ ,\end{aligned}$$

where $\mathrm{P}\left(C=l]\right)$, $l,1,\dots,3$, that we shall denote $\pi_l$, $l,1,\dots,3$, are the mixture proportions. Such a cumulative distribution admits the density

$$f_{\boldsymbol{P}}(\boldsymbol{p})=\prod_{k=1}^{100'000}\sum_{l=1}^{3}\mathcal{G}_{\boldsymbol{m}[l],\boldsymbol{\sigma}^2[l]}(\boldsymbol{p}[k])\mathrm{P}\left(C=l\right)$$

B.6) The center of the 3 clusters are given my the 3 means of the bivariate Gaussian distribution, *i.e.*, $\boldsymbol{m}[l]=[m_1[l],m_2[l]]$, $l=1,\dots,3$. Given the realizations of the two principal components $\boldsymbol{p}=[\boldsymbol{p}[1],\dots,\boldsymbol{p}[100'000]]$, the means can be obtained by estimating the parameters of the density $f_{\boldsymbol{P}}(\boldsymbol{p})$ using the maximum likelihood method.

- Set the likelihood function

$$h(\boldsymbol{p}\,;\,\boldsymbol{\theta}) = \prod_{k=1}^{100'000} \sum_{l=1}^{3} \mathcal{G}_{\boldsymbol{m}[l],\boldsymbol{\sigma}^2[l]}(\boldsymbol{p}[k])\pi_l\,,$$

  function of the parameters $\boldsymbol{\theta} = \{m_1[l], m_2[l], \sigma_1^2[l], \sigma_2^2[l], \pi_l, l = 1, \ldots, 3\}$
- Given the realizations of the two principal components $\boldsymbol{p} = [\boldsymbol{p}[1], \ldots, \boldsymbol{p}[100'000]]$, by mean of the EM algorithm, compute the set of the parameters that maximises the likelihood function

$$\hat{\boldsymbol{\theta}} = \arg\max_{\theta} h(\boldsymbol{p}\,;\,\boldsymbol{\theta})\,.$$

B.7) Each center represent the value of the principal components $p_1$ and $p_2$ corresponding the centred version of the variables $i_a, i_d, t_i, w_a, w_d$ describing a characteristic 2 pressure event.

B.8) Set $p_1[l] = m_1[l]$, $p_2[l] = m_2[l]$, $l = 1, \ldots, 3$. Then the 3 set of variables $\boldsymbol{v}[l] = [i_a[l], i_d[l], t_i[l], w_a[l], w_d[l]]$, $l = 1, \ldots, 3$, describing the 3 characteristic 2 pressure events are given by

$$\boldsymbol{v}[l] = \boldsymbol{p}[l]\boldsymbol{V}^T + \boldsymbol{m}_v\,, \quad \text{(dimensions } 1 \times 5)\,, \quad l = 1, \ldots, 3\,.$$

B.9) The variable given in B.8), *i.e.*,

$$\boldsymbol{v}[l] = \boldsymbol{p}[l]\boldsymbol{V}^T + \boldsymbol{m}_v\,, \quad \text{(dimensions } 1 \times 5)\,, \quad l = 1, \ldots, 3\,,$$

represent the three characteristic 2 pressure events. In other words, the 100′000 sets of variables $\boldsymbol{v}[k] = [i_a[k], i_d[k], t_i[k], w_a[k], w_d[k]]$, $k = 1, \ldots, 100'000$ can be seen as only 3 sets of values $\boldsymbol{v}[l] = [i_a[l], i_d[l], t_i[l], w_a[l], w_d[l]] = \boldsymbol{p}[l]\boldsymbol{V}^T + \boldsymbol{m}_v$, $l = 1, \ldots, 3$, plus noise.

## Part C: Analysis of the Pavement Response (15 points)

C.1) The spectrum is a smooth spectrum that can be approximated as a rational spectrum

$$S_X(\omega) = \frac{1}{C(z)} = \left.\frac{\sigma_W^2}{|P(z)|^2}\right|_{z=e^{j\omega}}\,,$$

where $P(z) = 1 + p_1 z^{-1} + \ldots + p_M z^{-M}$ is assumed strictly minimum phase. Then $H(z) = \frac{1}{P(z)}$ is a stable filter. Given that $W[n]$ is *w.s.s.*, by the fundamental filtering formula, the process $X[n] = H(z)W[n]$ is w.s.s., and, given its expression, it corresponds to an autoregressive process.

The parameters of $H(z)$ can be then estimated using the Yule-Walker method for the estimation of the parameters of an autoregressive process.

- Set the order of the model $M$.

- Using the $N = 1'000'000$ samples $x[n]$, $n = 1, \ldots, 1'000'000$, compute

$$\widehat{R}_X[k] = \frac{1}{N} \sum_{l=1}^{N-|k|} x[l+k]x^*[l]\,, \quad k = 0, \ldots, N-1\,.$$

  Notice that *a priori* an AR is zero mean. Nevertheless, since we work by approximation, we should then first center the process $x[n]$ before computing the correlation.

- Write the Yule-Walker equations for the estimation of the $M$ parameters and do not forget that $\sigma^2$ is given since it is the power of the given specific pressure pattern

$$\begin{aligned} \widehat{R}_X[0]p_1 + \ldots + \widehat{R}_X[M-1]p_M &= \widehat{R}_X[1] \\ \vdots \quad &= \quad \vdots \\ \widehat{R}_X[M-1]p_1 + \ldots + \widehat{R}_X[0]p_M &= \widehat{R}_X[M] \end{aligned}$$

- Solve the equations to obtain $\widehat{p}_1, \ldots, \widehat{p}_M$.
- Then $H(z) = \frac{1}{1+\widehat{p}_1 z^{-1} + \ldots + \widehat{p}_M z^{-M}}$

C.2) The order of the model can be estimated in two ways:

- Approximately, by considering the number of peaks of the spectrum;
- Statistically, by using the Levinson algorithm. The Levinson algorithm subsequently estimate the model parameters by starting with a model order of 1 and then increasing it. The analysis of the prediction error enables to determine at which value of the order model further increasing of the model order does not improve the model estimation.

C.3) By the fundamental filtering formula the signal $X[n]$ is w.s.s. process. Indeed, the filter $H(z)$ is stable and the input of the filter is teh w.s.s. process $W[n]$.

C.4) The process is w.s.s. (see point C.3) ) and it satisfy the recursive equation

$$X[n] + p_1 X[n-1] + \ldots + p_M X[n-M] = W[n]\,.$$

Recalling the definition of an AR process, *i.e.*, a w.s.s. process satisfying the above equation, we can affirm that the process $X[n]$ is indeed an autoregressive process.

**Grade Scale**.

The exams accounts for a total of 56 points (exact response to each question).

The grading has been done on a 50 points scale (50 points = 6/6), according to the following formula

$$\text{grade over } 6 = 1 + (5 * \text{points}/50)$$

The result is then constrained to be at maximum 6. The maximum obtained score was 51/50.

The grade of the final exam has been added to the midterm grade and to the mini-project grades and then rounded to .25 steps, that is

$$\text{rounded grade over } 6 = (\text{round-to-0-digit}(4 * \text{grade over } 6))/4\,.$$