# Statistical Signal & Data Processing - COM500
# Final Exam

June 24 2024, Duration 3h

## Read Me First!

**You are allowed to use:**

- The given cheatsheet summarizing the most important formulas;
- 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.**

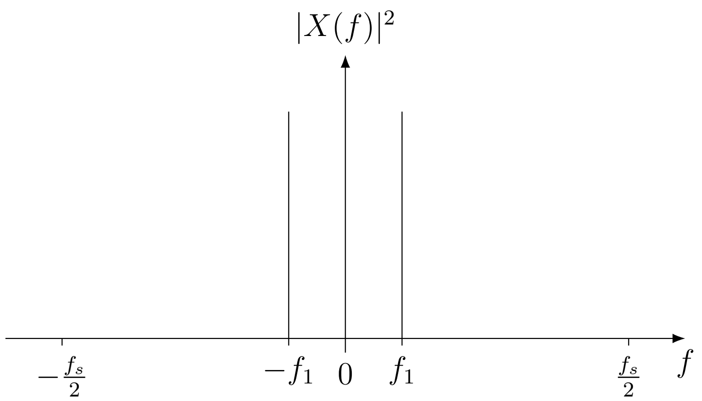
**Return your sheets ordered according to problem (solution) numbering.**

## 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**. JUST A REAL SINUSOID (2 POINTS)
Consider a sinusoid at frequency $f_1 = 440$ Hz that is sampled with a sampling frequency $f_s = 8000$ Hz (telephone quality). The theoretical power spectrum is depicted in the picture below.




We mesure $N$ samples of the sinusoid and we use the periodogram to compute its power spectral density.

1) What is the minimum number of samples that are necessary to correctly distinguish the two peaks (at frequencies $\pm f_1$) in the periodogram?

2) Sketch the spectrum computed using the periodogram with the minimum number of samples.

**Exercise 2**. JUST AN AUTO-REGRESSIVE PROCESS (2 POINTS)
Consider an auto-regressive process of order 1

$$X[n] = aX[n-1] + W[n]\,,$$

where the variance of the white noise is $\sigma_W^2$.
Prove the expression of its correlation function.

## Main Problem

*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 you are rather asked to describe how to apply such tools to the given problem. If needed, you can add assumptions to the problem setup.*

**Exercise 3**. Acoustic of Organ Pipes

Plots and some contents of this problem are taken from the following two scientific papers:

[1] J. Angster, P. Rucz, A. Miklós. Acoustics of Organ Pipes and Future Trends in the Research. Acoustics Today, March 2017, Volume 13(1), pp. 12-20.

[2] J. Prezelj, M. Čudina. Quantification of Aerodynamically-induced Noise and Vibration-induced Noise in a Suction Unit. Journal of Mechanical engineering science- Proceedings of the institution of mechanical engineers, 2011, Volume: 225(3), pp. 617-624.

The pipe organ produces a majestic sound that differs from all other musical instruments. The richness and variety of sound color (timbre) produced by a pipe organ is very unique because of the almost uncountable possibilities for mixing the sounds from different pipes.

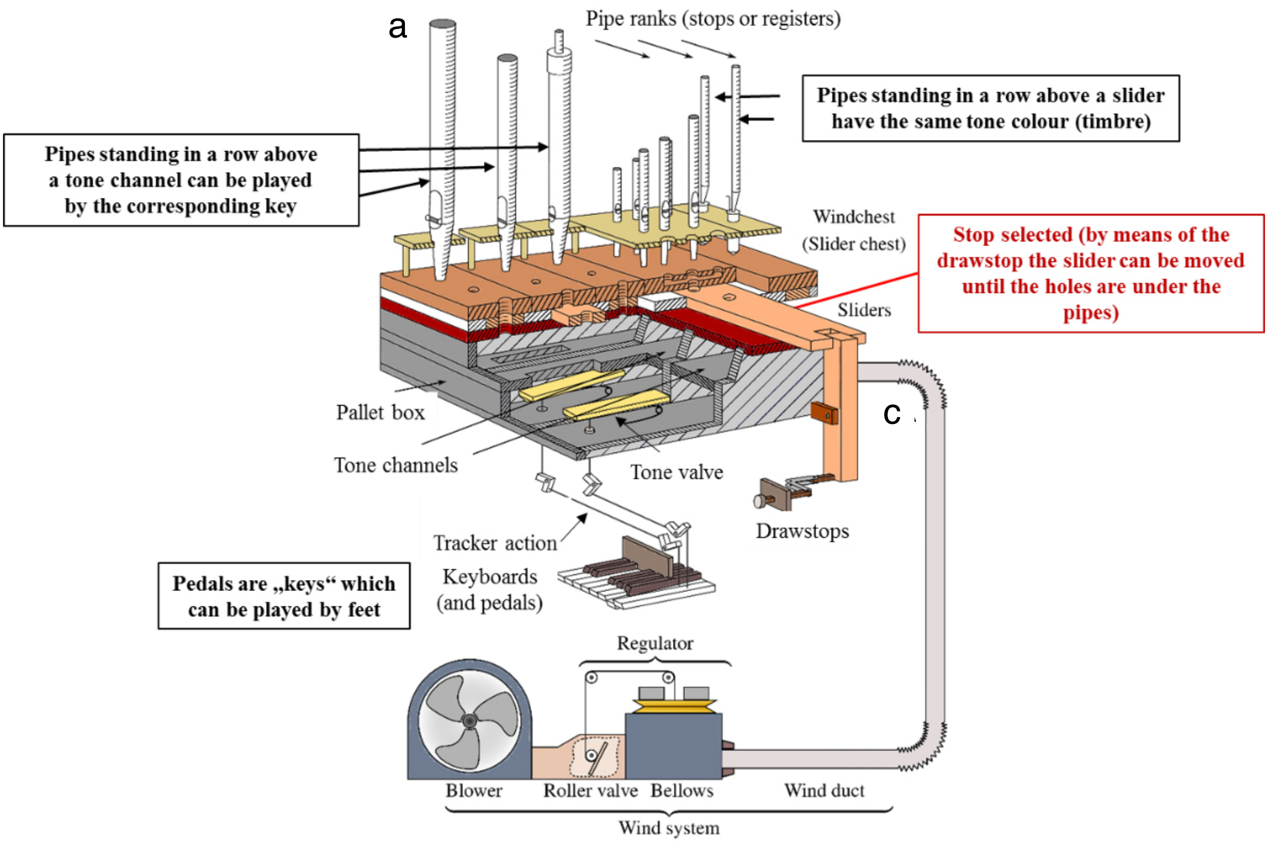



*Fig. 1: A sketch of a pipe organ and its most important parts [1].*

The so called **stationary spectrum** is the spectrum of the sound of a continuously sounding pipe. Such a sound can be measured by microphones at mouth or at the open end of a pipe.

Figure 2 depicts the symmetric stationary spectrum of a continuously sounding pipe playing a C note at 261.63 Hz. Notice that: The amplitude scale is in dB; Only half of the symmetric spectrum is depicted; Only the informative part of the spectrum is depicted.

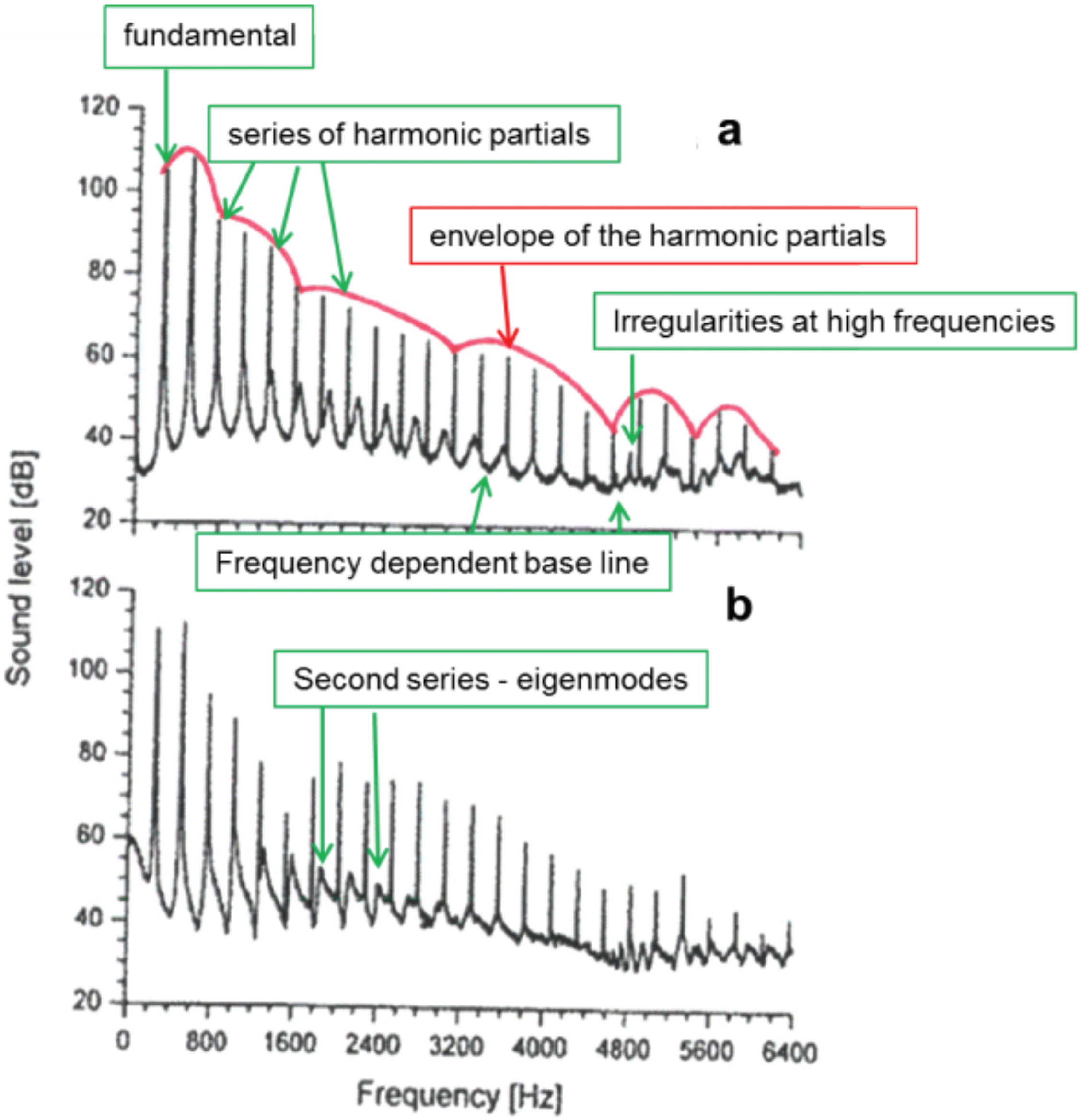



*Fig. 2: Typical stationary spectrum of a flue organ pipe at the open end (a) and at the mouth (b) [1].*

The spectrum is composed of:

- The spectrum of a harmonic signal $X_S[n]$ corresponding to the played note and composed of a fundamental frequency (here at 261.63 Hz) and its harmonics;
- The spectrum of a harmonic signal $X_E[n]$ corresponding to the acoustic eigen-modes of the pipe, that is, to the presence of standing waves. A so-called standing wave occurs in a pipe when the sound waves reflected back and forth in the pipe are combined such that each location along the pipe axis has constant but different amplitude. The locations with minimum and maximum amplitude are called nodes and antinodes, respectively.

The frequency of the standing wave is the resonance frequency or eigen-frequency of the tube. Standing waves occur in a tube on several frequencies. Notice that the standing waves frequencies are **not harmonically related** because of the signal shape at the mouth of the pipe (see Fig. 3).

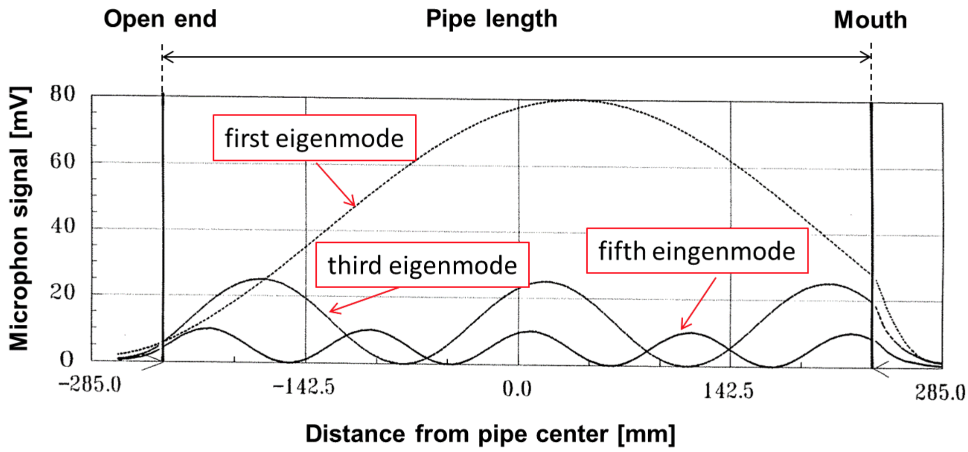



*Fig. 3: Standing waves in an organ pipe. Sound pressure distributions of the first, third, and fifth eigen-modes in a wide pipe are shown.*

- A frequency-dependent baseline, corresponding to the "colored" noise $W_C[n]$ that is produced by the airflow at the mouth of the pipe (called aerodynamically-induced noise). Such a noise is typical of air blowers. It is "colored" in the sense that it does not correspond to a flat spectrum (baseline). Figure below depicts the typical spectrum of such a noise.

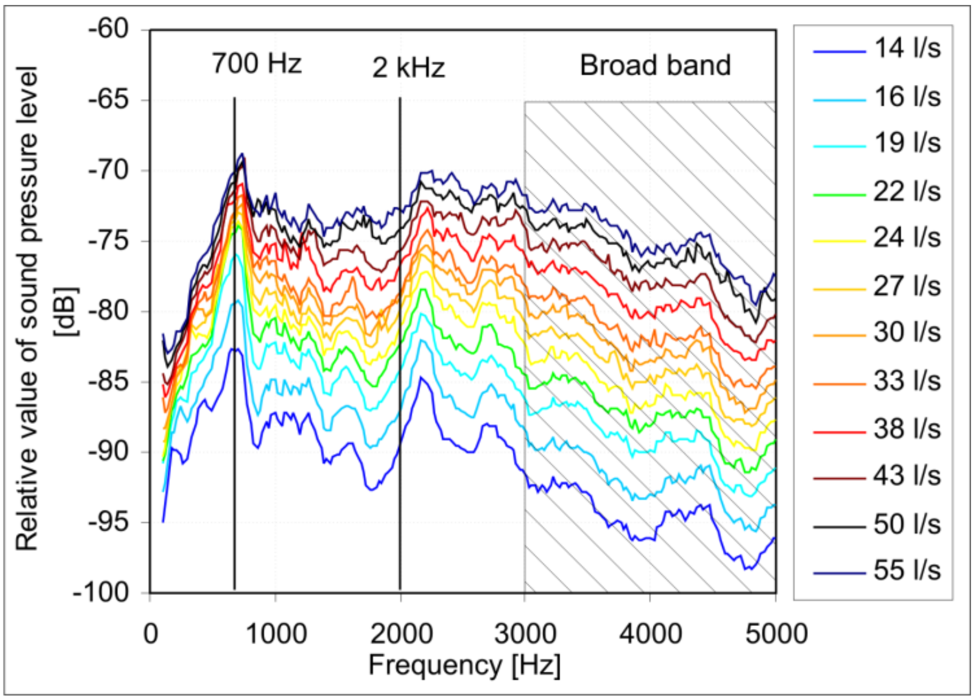



*Fig. 4: Spectrum of air flow noise at different flow rates [2].*

The following three parts are independent.

**Part A: Modelling and Analysis (18 pts)**

We proceed to the modelling of the signal generated by a continuously sounding pipe.

A.1 Based on the spectrum of Fig. 1 (please notice that the amplitude scale is in dB), provide a w.s.s. model for each of the three signals:

- $X_S[n]$ (based on Fig. 1, try to set, approximately, the order of the model);
- $X_E[n]$ (based on Fig. 1, try to set, approximately, the order of the model);
- $W_C[n]$ (to model a colored noise consider the filtering of a noise with an all poles filter, and based on Fig. 1, try to set, approximately, the order of the model).

A.2 Prove that the models for $x_E[n]$ (limited to 2 harmonics) and $W_C[n]$ describe indeed a w.s.s. process.

Then we proceed to the analysis.
We suppose to be able to mesure the three signals $x_S[n]$, $x_E[n]$ and $W_C[n]$ separately.

A.3 Propose a method to estimate the parameters of the model of the signal $x_E[n]$. 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 Propose a method to estimate the parameters of the model of the signal $W_C[n]$. 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.

**Part B: Characterization (20 pts)**

Although the main features of the sound of organ pipes have been investigated extensively, the connection between sound character and pipe shape and the dimensions are still not well understood.
We would like to characterize different pipes, corresponding to the same note, but made with different material, dimensions, shape, etc. etc..
We therefore proceed in testing $M = 1000$ pipes corresponding to the note $C$ at 261.63 Hz. Each pipe is continuously played and the generated sound $x[n]$ recorded.
For each recorded sound the spectrum and the following characteristics are computed (please refer to Fig. 1)

- The average amplitude of the first three harmonics $y_{A3}[n]$;
- The energy of the eigen-frequencies (eigen-modes) $y_{EF}[n]$;
- The energy of the irregularities at high frequencies $y_I[n]$;
- The energy of the envelope of the harmonics $y_{EH}[n]$;
- The energy of the frequency dependent (noise) baseline $y_{NB}[n]$;

Given that we do not if all the characteristics are individually relevant, *i.e.*, independent, we apply the principal component analysis.

B.1) Describe in detail, step by step, how to compute the principal components given the variables $\boldsymbol{y}[n] = [y_{A3}[n], y_{EF}[n], y_I[n], y_{EH}[n], y_{NB}[n]]$, $n = 1, \ldots, 1000$. We shall denote the principal components as $z_1[n], \ldots, z_5[n]$, $n = 1, \ldots, 1000$. 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 $z_1[n]$ and $z_2[n]$, $n = 1, \ldots, 1000$, account for most of the total sample variation.

B.3) What does it mean, in relation to the variables $\boldsymbol{y}$, 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 4 different clusters.

B.4) Provide a Gaussian Mixture model based on the principal components $z_1[n]$ and $z_2[n]$, $n = 1, \ldots, 1000$, describing the 4 clusters. More precisely, develop its cumulative distribution function and do not forget that the cluster plot has a total of 1000 points!

Call $m_1$, $m_2$, $m_3$ and $m_4$ the centers of the 4 clusters, respectively.

B.5) What do these centers represent, in relation to the variables $\boldsymbol{y}$?

B.6) Now that you have the centers of the 4 clusters, let's go back to the variable $\boldsymbol{y}$ space. How can you do so? Write the corresponding equations.

B.7) You obtain 4 sets of variables $\boldsymbol{y}[n] = [y_{A3}[n], y_{EF}[n], y_I[n], y_{EH}[n], y_{NB}[n]]$, $n = 1, \ldots, 4$. What do these 4 sets represent? Why do we have 4 sets now and not 1000 as in the beginning of the exercise?

## Part C: Denoising (8 pts)

We would like to get rid of the aerodynamically-induced noise that comes from the wind system. To do so, we mesure the noise $w[n]$ where the air flow enter the organ (point **c** in Fig. 1) and the sound generated by the pipe $x[n]$ at the open end of the pipe (point a in Fig. 1). As clearly appears from Fig. 2, the frequency content of the noise is different between the open end (**a**) and the mouth of the pipe (**b**), suggesting that the pipe has a particular impulse response $h$. We suppose that such impulse response might vary over time being dependent to the intensity at which each note is played.

C.1) Provide the scheme (draw a block diagram) of the adaptive filter $f_n$ capable of reducing the aerodynamically-induced noise at the open end of the pipe. Show the signal that is used for adaptation and give the quantity $J(f_n)$ 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) Consider that, as mentioned, the impulse response of the pipe depends on the intensity at which a note is played. Give an example of a situation where the adaptive filtering will not work (*i.e.*, it will need time to correctly denoise the sound)

# COM500 CheatSheet



## Stochastic Processes

w.s.s.

$$\mathrm{E}[X[n]]=\text{const.}\,,\mathrm{Var}(X[n])=\text{const.}<\infty$$

$$\mathrm{E}[X[k]X^*[l]]=R(k-l)\,,\forall k,l\in\mathbb{Z}\,,$$

PSD: w.s.s. + $R_X(k)\in\ell_1$ (summable)

$$S_X(\omega)=\sum_{k=-\infty}^{\infty} R_X(k)e^{-i\omega k}\,.$$

Fundamental Filtering Formula

$X[n]$ w.s.s., $R_X(k)\in\ell_1$, and $h_k\in\ell_1$, then

$$Y[n]=\sum_{k=-\infty}^{\infty} h_{n-k}X[k]\,,\text{ is w.s.s. with}$$

$$\mathrm{E}[Y]=\mathrm{E}[X]\sum_{k=-\infty}^{\infty} h_k\,,\quad R_Y(k)\in\ell_1$$

$$S_Y(\omega)=|H(\omega)|^2 S_X(\omega)\,,H(\omega)=\text{DTFT of } h_k$$

Markov Chain

$\{X[n]\}_{n\in\mathbb{Z}}$ (considered stationary) with discrete values in $\mathcal{D}$, |

$$\mathrm{P}\big(X[n]=i_n \mid X[n-1]=i_{n-1},X[n-2]=i_{n-2},\ldots\big)$$
$$=\mathrm{P}\big(X[n]=i_n \mid X[n-1]=i_{n-1}\big)\,,\forall i_n,i_{n-1},\ldots\in\mathcal{D}$$

Hidden Markov Chain

$\{X[n]\}_{n\in\mathbb{Z}}$ Markov chain, $\{W[n]\}_{n\in\mathbb{Z}}$ Gaussian white noise

$$Y[n]=X[n]+W[n]\,.$$

Bayes' Rule

$A$ and $B$ with discrete values in $\mathcal{D}$,

$$\mathrm{P}(A=k \mid B=l)=\frac{\mathrm{P}(A=k\,,\ B=l)}{\mathrm{P}(B=l)}\,,\quad k,l\in\mathcal{D}\,.$$

AR Process

A w.s.s. process $X[n]$, with values in $\mathbb{R}$, |

$$\sum_{k=0}^{M} p_k X[n-k]=W[n]\,,\quad n\in\mathbb{Z}\,,$$

$W[n]$, is a zero mean Gaussian white noise

$p_k\,,k=0,\ldots,M$ bounded coefficients (real or complex). We assume $p_0=1$.

Filtering Interpretation ($z^{-1}$ delay operator)

$$P(z)X[n]=W[n]$$

Canonical form: $P(z)$ strict. min. phase, $p_0=1$.

Correlation:

$$R_X[m]+\sum_{k=1}^{M-1} p_k R_X[m-k]=\delta_m\sigma_W^2\,,\quad m\geq 0\,.$$

PSD (fundamental filtering formula):

$$S_X(\omega)\big|P\big(e^{j\omega}\big)\big|^2=\sigma_W^2\,.$$

Harmonic Processes

$$X[n]=\sum_{k=1}^{K}\alpha_k e^{j(\omega_k n+\Theta_k)}\,,n\in\mathbb{N}\,,$$

$\Theta_k$ i.i.d. uniformly distributed over $[0,2\pi]$.

$$R_X[l]=\sum_{k=1}^{K}|\alpha_k|^2 e^{j\omega_k l}\,,\ S_X(\omega)=\sum_{k=1}^{K}|\alpha_k|^2\delta(\omega-\omega_k)\,.$$

Poisson Random Process $N((0,t])$

$N((0,t])$ obeys the Poisson distribution $\mathrm{P}(N=k)=\frac{(a\lambda)^k e^{-a\lambda}}{k!}$, ($\lambda$ is the rate), and given two disjoint intervals $(t_1,t_2]$ and $(t_3,t_4]$, $N((t_1,t_2])$ is independent of $N((t_3,t_4])$.

Inter-arrival time $S_n=T_n-T_{n-1}$. i.i.d. with density $f_S(t)=\lambda e^{-t\lambda}$.

## Hilbert Spaces

Projection Theorem

$E,S$ Hilbert spaces with $S\subset E$, then

$$\forall v\in E\,,\exists!\,b\in S\mid$$

$$b=\arg\min_{c\in S}\|v-c\|\,,\ \langle v-b,c\rangle=0\,,\forall\ c\in S\,,$$

Projection Theorem w.s.s.

$E\,S$ Hilbert spaces of w.s.s. processes with $S\subset E\subseteq \mathrm{L}^2(P)$, then

$$\forall X[n]\in E\,,\exists!\,Y[n]\in S\mid$$

$$Y[n]=\arg\min_{U[n]\in S}\mathrm{E}[|X[n]-U[n]|^2]\,,$$

$$\mathrm{E}[(X[n]-Y[n])U^*[n]]=0\,,\forall\ U[n]\in S\,,$$

## Empirical Statistics

Bias & Variance

$\widehat{S}(x[1],\ldots,x[N])$ empirical statistics of a probabilistic moment $S$.

Bias $\mathrm{E}\big[\widehat{S}(X[1],\ldots,X[N])\big]-S\,,$

Variance $\mathrm{Var}\big(\widehat{S}(X[1],\ldots,X[N])-S\big)\,.$

Unbiased & Biased Correlation

$$\widehat{R}_X^{\mathrm{NB}}(k)=\frac{1}{N-|k|}\sum_{n=1}^{N-|k|} x[n+k]x^*[n]\,,$$

$$\widehat{R}_X^{\mathrm{B}}(k)=\frac{1}{N}\sum_{n=1}^{N-|k|} x[n+k]x^*[n]\,.$$

## Methods

Linear Estimation of w.s.s.: Wiener Filter

Estimation of $X[n]$ given $Y[n]$

Normal equations $R_{XY}[u]=\sum_{m\in\mathbb{Z}} h[m]R_Y[u-m]$

Wiener Filter $H(e^{j\omega})=\frac{S_{XY}(\omega)}{S_Y(\omega)}$

Linear Prediction of w.s.s.: Yule-Walker

Prediction of $X[n]$ as linear combination of $X[n-1],\ldots,X[n-N]$.

Coefficients $a_k$ solutions of

$$\begin{bmatrix} R_X[0] & \ldots & R_X[N-1]\\ \vdots & \ddots & \vdots\\ R_X[N-1] & \ldots & R_X[0]\end{bmatrix}\begin{bmatrix}a_1\\ \vdots\\ a_N\end{bmatrix}=\begin{bmatrix}R_X[1]\\ \vdots\\ R_X[N]\end{bmatrix}.$$

Linear Estimation of AR: Yule-Walker

$\sum_{k=0}^{N} p_k X[n-k]=W[n]$. Coeff. $p_k$ solution of

$$-\begin{bmatrix} R_X[0] & \ldots & R_X[N-1]\\ \vdots & \ddots & \vdots\\ R_X[N-1] & \ldots & R_X[0]\end{bmatrix}\begin{bmatrix}p_1\\ \vdots\\ p_N\end{bmatrix}=\begin{bmatrix}R_X[1]\\ \vdots\\ R_X[N]\end{bmatrix}.$$

$$\sigma_W^2=R_X[0]+R_X[1]p_1+\ldots+R_X[N]p_N$$

Linear Pretiction of AR: Projection Theorem

$H(X,n)=H(W,n)\,,\quad\forall\ n\in\mathbb{Z}\,.$

Intuitive property:

$$Y\in H(X,n+k)\,,\ Y=A+B\,,\ A\perp H(X,n)\,,\ B\in H(X,n)$$

orthogonal projection of $Y$ onto $H(X,n)$ is $B$.

Estimation Param. Prob.: MLE

$\boldsymbol{\theta}$ parameters of the prob. function $f_Y$, $y[1],\ldots,y[N]$ realization of the process $Y[n]$, then

$$\widehat{\boldsymbol{\theta}}=\arg\max_{\boldsymbol{\theta}} f_Y(y[1],\ldots,y[N],\boldsymbol{\theta})$$

## Spectral Estimation

Periodogram: General w.s.s. process

$$P_X^N(\omega)=\frac{1}{N}\Big|\sum_{n=1}^{N} x[n]e^{-j\omega n}\Big|^2=\frac{1}{N}|\widehat{x}_N(\omega)|^2\,,$$

Bias $\sum_{k=-N+1}^{N-1}\frac{N-|k|}{N}R_X[k]e^{-j\omega k}-S_X(\omega)$

Variance constant

Resolution $\Delta f>\frac{1}{N}$

Annihilating Filter: Line Spectra

Estimation of line spectrum frequencies and amplitudes of a Harmonic w.s.s. process in absence of noise

1) Given $2K$ observations, solve the system

$$\begin{bmatrix} x[K-1] & \ldots & x[0]\\ \vdots & & \vdots\\ x[2K-2] & \ldots & x[K-1]\end{bmatrix}\begin{bmatrix}h[1]\\ \vdots\\ h[K]\end{bmatrix}=-\begin{bmatrix}x[K]\\ \vdots\\ x[2K-1]\end{bmatrix}$$

2) Compute $H(z)$ and the zeros of $H(z)$

3) Compute the argument of the zeros of $H(z)$

4) Compute $\omega_k$ from the zeros' arguments

5) Compute the amplitudes $|\alpha_k|^2$ by solving

$$\begin{bmatrix} 1 & \ldots & 1\\ e^{j\omega_1} & \ldots & e^{j\omega_K}\\ \vdots & & \vdots\\ e^{j\omega_1(K-1)} & \ldots & e^{j\omega_K(K-1)}\end{bmatrix}\begin{bmatrix}\alpha_1 e^{j\Theta_1}\\ \alpha_2 e^{j\Theta_2}\\ \vdots\\ \alpha_K e^{j\Theta_K}\end{bmatrix}=\begin{bmatrix}x[0]\\ x[1]\\ \vdots\\ x[K-1]\end{bmatrix}$$

MUSIC: Line Spectra

Estimation of line spectrum frequencies and amplitudes of a Harmonic w.s.s. process in the presence of noise

1) Given $M$ observations with $M>>N>K$ center the process and compute the empirical correlation matrix

$$\widehat{\boldsymbol{R}}_{\boldsymbol{Y}}^{NN}=\frac{1}{M-N+1}\sum_{n=1}^{M-N+1}\boldsymbol{y}^{N1}[n]\boldsymbol{y}^{N1}[n]^H\,;$$

2) Compute the eigendecomposition $\widehat{\boldsymbol{G}}^{N(N-K)}$ of $\widehat{\boldsymbol{R}}_{\boldsymbol{Y}}^{NN}$ corresponding to $\lambda_{K+1}^R$ to $\lambda_N^R$.

3.a) Determine the peaks of

$$\frac{1}{\boldsymbol{e}^{N1}(\omega)^H\widehat{\boldsymbol{G}}^{N(N-K)}\widehat{\boldsymbol{G}}^{N(N-K)\,H}\boldsymbol{e}^{N1}(\omega)}\,;$$

where

$$\boldsymbol{e}^{N1}(\omega)=\begin{bmatrix}1 & e^{-j\omega} & \ldots & e^{-j(N-1)\omega}\end{bmatrix}^T,$$

3.b) Determine the minimum values of

$$\boldsymbol{e}^{N1}(\omega)^H\widehat{\boldsymbol{G}}^{N(N-K)}\widehat{\boldsymbol{G}}^{N(N-K)\,H}\boldsymbol{e}^{N1}(\omega)\,.$$

4) Compute the modulus of the amplitudes using

$$\boldsymbol{R}_{\boldsymbol{Y}}^{NN}=\boldsymbol{E}^{NK}\boldsymbol{A}^{KK}\boldsymbol{E}^{NK\,H}+\sigma_W^2\boldsymbol{I}^{NN}\,.$$

Yule-Walker: Smooth Spectra

1) Given $N$ observations with $N>>M$ center the process and

compute the empirical correlation

$$\widehat{R}_X[k]=\frac{1}{N}\sum_{n=1}^{N-k} x_0[n+k]x_0[n]^*\,, \quad k=0,\ldots,M$$

2) Solve the Yule Walker equations to obtain $\hat{p}_1,\ldots,\hat{p}_M$
3) Compute the estimate of the spectrum as

$$\widehat{S}_X(\omega)=\left.\frac{\widehat{\sigma}^2_W}{|\widehat{P}(z)|^2}\right|_{z=e^{j\omega}}.$$

## Mixture Models

Sequence of samples $\boldsymbol{y}=[y[1],\ldots,y[N]]$, sequence of corresponding classes $\boldsymbol{c}=[c[1],\ldots,c[N]]$

$$F_{\boldsymbol{Y}}(\boldsymbol{y})=\sum_{\boldsymbol{c}\in\mathcal{C}} \mathrm{P}(\boldsymbol{Y}\leq\boldsymbol{y},\boldsymbol{C}=\boldsymbol{c})=\sum_{\boldsymbol{c}\in\mathcal{C}} \mathrm{P}(\boldsymbol{Y}\leq\boldsymbol{y}\mid\boldsymbol{C}=\boldsymbol{c})\mathrm{P}(\boldsymbol{C}=\boldsymbol{c})$$
$$=\sum_{\boldsymbol{c}\in\mathcal{C}}\prod_{n=1}^{N} \mathrm{P}(Y[n]\leq y[n]\mid C[n]=c[n])\mathrm{P}(\boldsymbol{C}=\boldsymbol{c})$$

i.i.d. Mixtures

$$\mathrm{P}(\boldsymbol{C}=\boldsymbol{c})=\mathrm{P}(C[1]=c[1])\ldots\mathrm{P}(C[N]=c[N])=\pi_{c[1]}\cdots\pi_{c[N]}\,,$$

Markovian Mixtures

$$\mathrm{P}(\boldsymbol{C}=\boldsymbol{c})=\pi_{c[1]}p_{c[1]c[2]}\cdots p_{c[N-1]c[N]}\,,$$

Discrete value Process + Noise

$\boldsymbol{Y}=\boldsymbol{X}+\boldsymbol{W}$ where $\boldsymbol{X}$ is a discrete value process and $\boldsymbol{W}$ a white Gaussian noise.

$$f_{\boldsymbol{Y}}(\boldsymbol{y})=\sum_{\boldsymbol{x}\in\mathcal{C}}\prod_{n=1}^{N}\mathcal{G}_{x[n],\sigma^2}(y[n])\mathrm{P}(\boldsymbol{X}=\boldsymbol{x})\,.$$

Denoising a Discrete Value Process

Estimate the parameters of the mixture model using the maximum likelihood approach; Estimate the original signal using the maximum *a posteriori* approach, *i.e.*, find $\boldsymbol{x}$ maximizing the *a posteriori* distribution

$$\mathrm{P}(\boldsymbol{X}=\boldsymbol{x}\mid\boldsymbol{y})=\frac{f_{\boldsymbol{Y}}(\boldsymbol{y}\mid\boldsymbol{X}=\boldsymbol{x})\mathrm{P}(\boldsymbol{X}=\boldsymbol{x})}{f_{\boldsymbol{Y}}(\boldsymbol{y})}$$

## PCA

Principal Components Computaiton

*M data vectors*, each characterized of $N$ variables (realization of a zero mean w.s.s. process) $\boldsymbol{c}_m=[c_m[1],\ldots,c_m[N]]^T\,, \quad m=1,\ldots,M$.
Empirical correlation matrix

$\widehat{\boldsymbol{R}}_{\boldsymbol{c}}=\frac{1}{M}\sum_{m=1}^{M}\boldsymbol{c}_m*\boldsymbol{c}_m^H=\frac{1}{M}\boldsymbol{C}*\boldsymbol{C}^H\,,(N\times N)$, where $\boldsymbol{C}=\begin{bmatrix}\boldsymbol{c}_1 & \ldots & \boldsymbol{c}_M\end{bmatrix}$,
$V$ solution of the equation $\widehat{\boldsymbol{R}}_{\boldsymbol{c}}\boldsymbol{V}=\boldsymbol{V}\boldsymbol{\Lambda}$, where $\boldsymbol{\Lambda}=\mathrm{diag}(\lambda_1,\ldots,\lambda_N)$ and $\boldsymbol{V}^H\widehat{\boldsymbol{R}}_{\boldsymbol{c}}\boldsymbol{V}=\boldsymbol{\Lambda}$.
Principal components $\boldsymbol{Z}=\boldsymbol{V}^T\boldsymbol{C}\,,(N\times M)$, uncorrelated.
Invertible transformation $\boldsymbol{C}=\boldsymbol{V}\boldsymbol{Z}\,,$

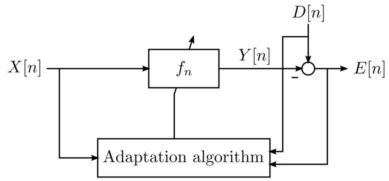
Analysis

$K<<N$ eigenvalues with highest values (lossy/lossless reduction of variables)

## Adaptive Filtering / Echo cancellation

Wiener-Hopf equations $\sum_{k\in\mathbb{Z}} h[n-k]R_Y(k-l)=R_{XY}(n-l)\,, \quad \forall\, l\,.$

Echo cancellation setup




$$E[n]=D[n]-f_n*X[n]=S[n]+h*X[n]-f_n*X[n]=S[n]+(h-f_n)*X[n]\,.$$

Cost function & normal equations

Cost function for a $k$-tap filter

$$J(f_n)=\mathrm{E}\left[|E[n]|^2\right]=\mathrm{E}\left[|D[n]-f_n*X[n]|^2\right], \quad \text{w.r.t.} \quad f_n[l]\,,l=0,1,\ldots,K$$

Minimum of the cost function = normal equations

$$\sum_{l=0}^{K} f_n[l]R_X(n-l,n-i)=R_{DX}(n,n-i)\,, \quad \boldsymbol{R}_{X,n}\boldsymbol{f}_n=\boldsymbol{r}_{DX,n}\,.$$

Iterative solution

$$\boldsymbol{f}^{(i+1)}=\boldsymbol{f}^{(i)}+\mu\boldsymbol{p}\,, \quad i=0,1,\ldots\,, \quad \mu \;\&\; \boldsymbol{p} \text{ such that } J(\boldsymbol{f}^{(i+1)})<J(\boldsymbol{f}^{(i)})$$

Convergence conditions

$$0<\mu<2/\lambda_{\max}\,, \quad \boldsymbol{p}=\tfrac{1}{2}(\boldsymbol{r}_{DX}-\boldsymbol{R}_X\boldsymbol{f}^{(i)}) \text{ or } \boldsymbol{p}=4\boldsymbol{R}_X^{-1}(\boldsymbol{r}_{DX}-\boldsymbol{R}_X\boldsymbol{f}^{(i)})(\text{Newton})$$

Convergence rate

- $0\leq 1-\mu\lambda_j<1$, monotonic decay to zero;
- $-1<1-\mu\lambda_j<0$ oscillatory decay to zero.
- $K+1$ modes $\{1-\mu\lambda_j, j=0,\ldots,K\}$. The modes with maximum magnitude (slowest rate of convergence), determine the convergence rate of the algorithm. One can select $\mu$ optimally by minimizing the value of the slowest mode $\min_\mu \max_{j=0,\ldots,K}|1-\mu\lambda_j|$, with the constraint that each of the modes is stable, i.e., $|1-\mu\lambda|<1$.

Computational burden reduction

Merging interation and adaptation

$$\boldsymbol{f}_{n+1}=\boldsymbol{f}_n+\mu(\boldsymbol{r}_{DX,n+1}-\boldsymbol{R}_{X,n+1}\boldsymbol{f}_n)\,,$$
$$\boldsymbol{f}_{n+1}=\boldsymbol{f}_n+\mu\boldsymbol{R}_{X,n+1}^{-1}(\boldsymbol{r}_{DX,n+1}-\boldsymbol{R}_{X,n+1}\boldsymbol{f}_n) \quad (\text{Newton})\,,$$

Replacing statistics with individual values

$$\boldsymbol{f}_{n+1}=\boldsymbol{f}_n+\mu\boldsymbol{X}_n(D[n]-\boldsymbol{X}_n^T\boldsymbol{f}_n)=\boldsymbol{f}_n+\mu\boldsymbol{X}_nE[n]\,,$$