

# Assignment 2 (ML for TS) - MVA

Colin Coërchon [colin180302@gmail.com](mailto:colin180302@gmail.com)  
Chloé Habasque [habasquechloe29@gmail.com](mailto:habasquechloe29@gmail.com)

27 Novembre 2024

## 1 Introduction

**Objective.** The goal is to better understand the properties of AR and MA processes and do signal denoising with sparse coding.

**Warning and advice.**

- Use code from the tutorials as well as from other sources. Do not code yourself well-known procedures (e.g., cross-validation or k-means); use an existing implementation.
- The associated notebook contains some hints and several helper functions.
- Be concise. Answers are not expected to be longer than a few sentences (omitting calculations).

**Instructions.**

- Fill in your names and emails at the top of the document.
- Hand in your report (one per pair of students) by Monday 2<sup>nd</sup> December 11:59 PM.
- Rename your report and notebook as follows:  
FirstnameLastname1\_FirstnameLastname1.pdf and  
FirstnameLastname2\_FirstnameLastname2.ipynb.  
For instance, LaurentOudre\_CharlesTruong.pdf.
- Upload your report (PDF file) and notebook (IPYNB file) using this link:  
<https://docs.google.com/forms/d/e/1FAIpQLSfCqMXSDU9jZJbYUMmeLCXbVeckZYNiDpPl4hRUwcJ2>

## 2 General questions

A time series  $\{y_t\}_t$  is a single realisation of a random process  $\{Y_t\}_t$  defined on the probability space  $(\Omega, \mathcal{F}, P)$ , i.e.  $y_t = Y_t(w)$  for a given  $w \in \Omega$ . In classical statistics, several independent realizations are often needed to obtain a “good” estimate (meaning consistent) of the parameters of the process. However, thanks to a stationarity hypothesis and a “short-memory” hypothesis, it is still possible to make “good” estimates. The following question illustrates this fact.

## Question 1

An estimator  $\hat{\theta}_n$  is consistent if it converges in probability when the number  $n$  of samples grows to  $\infty$  to the true value  $\theta \in \mathbb{R}$  of a parameter, i.e.  $\hat{\theta}_n \xrightarrow{\mathcal{D}} \theta$ .

- Recall the rate of convergence of the sample mean for i.i.d. random variables with finite variance.
- Let  $\{Y_t\}_{t \geq 1}$  a wide-sense stationary process such that  $\sum_k |\gamma(k)| < +\infty$ . Show that the sample mean  $\bar{Y}_n = (Y_1 + \dots + Y_n)/n$  is consistent and enjoys the same rate of convergence as the i.i.d. case. (Hint: bound  $\mathbb{E}[(\bar{Y}_n - \mu)^2]$  with the  $\gamma(k)$  and recall that convergence in  $L_2$  implies convergence in probability.)

## Answer 1

1. Soit  $\{X_1, X_2, \dots\}$  une suite de variables aléatoires i.i.d. avec  $\mathbb{E}[X_1] = \mu$  et  $\text{Var}(X_1) = \sigma^2$ . La moyenne empirique  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$  est un estimateur de  $\mu$ .

Nous savons que :

$$\mathbb{E}[\bar{X}_n] = \mu \quad \text{et} \quad \text{Var}(\bar{X}_n) = \frac{\sigma^2}{n}$$

La variance de  $\bar{X}_n$  décroît donc comme  $\frac{1}{n}$ . En appliquant l'inégalité de Bienaymé-Tchebychev, pour tout  $\varepsilon > 0$ , nous avons :

$$\mathbb{P}(|\bar{X}_n - \mu| \geq \varepsilon) \leq \frac{\text{Var}(\bar{X}_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}.$$

Cela montre que  $\bar{X}_n$  converge en probabilité vers  $\mu$  à un taux  $\mathcal{O}(1/\sqrt{n})$  :

$$\boxed{\bar{X}_n \xrightarrow{\mathbb{P}} \mu}$$

En d'autres termes, la moyenne empirique est consistante et converge en probabilité vers  $\mu$  à un taux de convergence en  $\mathcal{O}(1/\sqrt{n})$ .

2. Calcul de  $\mathbb{E}[(\bar{Y}_n - \mu)^2]$

Nous commençons par calculer la variance de  $\bar{Y}_n$ . Comme  $\mathbb{E}[Y_t] = \mu$  pour tout  $t$ , nous avons :

$$\mathbb{E}[\bar{Y}_n] = \mu.$$

La variance de  $\bar{Y}_n$  est donnée par :

$$\text{Var}(\bar{Y}_n) = \mathbb{E}[(\bar{Y}_n - \mu)^2] = \frac{1}{n^2} \sum_{t=1}^n \sum_{s=1}^n \text{Cov}(Y_t, Y_s).$$

En utilisant la stationnarité du processus,  $\text{Cov}(Y_t, Y_s) = \gamma(t - s)$ , nous obtenons :

$$\text{Var}(\bar{Y}_n) = \frac{1}{n^2} \sum_{t=1}^n \sum_{s=1}^n \gamma(t - s).$$

Pour développer cette somme, on effectue le changement de variables  $h = t - s$ . Plus précisément, la sommation double s'effectue dans l'ensemble  $\llbracket 1, n \rrbracket^2$ , que l'on peut partitionner comme suit :

$$\llbracket 1, n \rrbracket^2 = \bigcup_{h=-(n-1)}^{n-1} \{(t, s) \in \mathbb{N}^2 \mid t - s = h\} \stackrel{\text{not.}}{=} \bigcup_{h=-(n-1)}^{n-1} A_h$$

avec

$$|A_h| = n - |h|$$

Ainsi, nous pouvons réécrire la variance comme suit :

$$\text{Var}(\bar{Y}_n) = \frac{1}{n^2} \sum_{h=-(n-1)}^{n-1} (n - |h|) \gamma(h)$$

Nous devons maintenant analyser cette somme asymptotiquement pour  $n \rightarrow \infty$ , en tenant compte de l'hypothèse  $\sum_{k=-\infty}^{\infty} |\gamma(k)| < +\infty$ , qui implique la convergence absolue de la série  $\sum_{k=-\infty}^{\infty} \gamma(k)$ .

- **Étape 1 : Contrôle des termes pour  $|h| > K$  :** L'hypothèse  $\sum_{k=-\infty}^{\infty} |\gamma(k)| < +\infty$  garantit qu'il existe un entier  $K$  tel que la contribution des termes  $\gamma(h)$  pour  $|h| > K$  est négligeable. Plus précisément, nous avons :

$$\sum_{|h|>K} |\gamma(h)| < \epsilon \quad \text{pour tout } \epsilon > 0$$

Cela signifie que pour  $n$  suffisamment grand, la somme sur  $|h| > K$  contribue très peu à la somme totale. Nous pouvons donc séparer la somme en deux parties :

$$\sum_{h=-(n-1)}^{n-1} (n - |h|) \gamma(h) = \sum_{|h| \leq K} (n - |h|) \gamma(h) + \sum_{|h| > K} (n - |h|) \gamma(h)$$

- **Étape 2 : Contribution des termes  $|h| \leq K$  :**

Pour  $|h| \leq K$ , les termes  $\gamma(h)$  sont bornés, et la contribution des termes  $(n - |h|) \gamma(h)$  est dominée par la somme finie des  $\gamma(h)$  :

$$\sum_{|h| \leq K} (n - |h|) \gamma(h) \sim n \sum_{|h| \leq K} \gamma(h) \quad \text{pour } n \text{ grand}$$

- **Étape 3 : Contribution des termes  $|h| > K$  :**

Pour  $|h| > K$ , comme  $\sum_{|h|>K} |\gamma(h)|$  est petit (par l'hypothèse de sommabilité), la contribution de cette somme devient négligeable lorsque  $n$  devient grand. Plus précisément, cette contribution est en  $o(n)$ , ce qui signifie que pour  $n \rightarrow \infty$ , nous avons :

$$\sum_{|h|>K} (n - |h|)\gamma(h) = o(n)$$

- **Étape 4 : Approximation asymptotique :**

En combinant les résultats des deux étapes précédentes, nous obtenons l'approximation asymptotique suivante pour la somme totale :

$$\sum_{h=-(n-1)}^{n-1} (n - |h|)\gamma(h) \sim n \sum_{|h|\leq K} \gamma(h) \quad \text{pour } n \text{ grand.}$$

Ainsi, la variance de  $\bar{Y}_n$  devient asymptotiquement :

$$\text{Var}(\bar{Y}_n) \sim \frac{1}{n} \sum_{|h|\leq K} \gamma(h)$$

Cela montre que, pour  $n$  grand, la variance de  $\bar{Y}_n$  décroît comme  $\frac{1}{n} \sum_{h=-\infty}^{\infty} \gamma(h)$ , car la contribution des termes  $|h| > K$  est négligeable. Donc :

$$\text{Var}(\bar{Y}_n) \sim \frac{1}{n} \sum_{h=-\infty}^{\infty} \gamma(h)$$

### Convergence en probabilité et taux de convergence

Nous avons montré que la variance de  $\bar{Y}_n$  est asymptotiquement proportionnelle à  $1/n$ , ce qui donne :

$$\mathbb{E}[(\bar{Y}_n - \mu)^2] = \mathcal{O}\left(\frac{1}{n}\right)$$

Cela implique que l'écart quadratique moyen entre  $\bar{Y}_n$  et  $\mu$  décroît comme  $1/n$ . Par conséquent, en appliquant l'inégalité de Bienaymé-Tchebychev, pour tout  $\varepsilon > 0$ , nous avons :

$$\mathbb{P}(|\bar{Y}_n - \mu| \geq \varepsilon) \leq \frac{\mathbb{E}[(\bar{Y}_n - \mu)^2]}{\varepsilon^2} = \mathcal{O}\left(\frac{1}{n\varepsilon^2}\right)$$

Cela montre que  $\bar{Y}_n$  converge en probabilité vers  $\mu$  à un taux de  $\mathcal{O}(1/\sqrt{n})$ .

### 3 AR and MA processes

#### Question 2 Infinite order moving average MA( $\infty$ )

Let  $\{Y_t\}_{t \geq 0}$  be a random process defined by

$$Y_t = \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots = \sum_{k=0}^{\infty} \psi_k \varepsilon_{t-k} \quad (1)$$

where  $(\psi_k)_{k \geq 0} \subset \mathbb{R}$  ( $\psi = 1$ ) are square summable, i.e.  $\sum_k \psi_k^2 < \infty$  and  $\{\varepsilon_t\}_t$  is a zero mean white noise of variance  $\sigma_\varepsilon^2$ . (Here, the infinite sum of random variables is the limit in  $L_2$  of the partial sums.)

- Derive  $\mathbb{E}(Y_t)$  and  $\mathbb{E}(Y_t Y_{t-k})$ . Is this process weakly stationary?
- Show that the power spectrum of  $\{Y_t\}_t$  is  $S(f) = \sigma_\varepsilon^2 |\phi(e^{-2\pi i f})|^2$  where  $\phi(z) = \sum_j \psi_j z^j$ . (Assume a sampling frequency of 1 Hz.)

The process  $\{Y_t\}_t$  is a moving average of infinite order. Wold's theorem states that any weakly stationary process can be written as the sum of the deterministic process and a stochastic process which has the form (1).

#### Answer 2

##### 1. Espérance de $Y_t$ :

$$\mathbb{E}[Y_t] = \mathbb{E} \left[ \sum_{k=0}^{\infty} \psi_k \varepsilon_{t-k} \right] = \sum_{k=0}^{\infty} \psi_k \mathbb{E}[\varepsilon_{t-k}] = 0$$

Car par définition de  $\{\varepsilon_t\}_t$ , on a  $\mathbb{E}[\varepsilon_{t-k}] = 0$  pour tout  $t \geq 0$ .

##### Covariance de $Y_t$ et $Y_{t-k}$ :

$$\mathbb{E}[Y_t Y_{t-k}] = \mathbb{E} \left[ \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i} \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-k-j} \right] = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \psi_i \psi_j \mathbb{E}[\varepsilon_{t-i} \varepsilon_{t-k-j}]$$

En utilisant les propriétés de la variance du bruit blanc, on obtient :

$$\mathbb{E}[\varepsilon_{t-i} \varepsilon_{t-k-j}] = \begin{cases} \sigma_\varepsilon^2 & \text{si } i = j + k \\ 0 & \text{sinon} \end{cases}$$

Cela donne :

$$\mathbb{E}[Y_t Y_{t-k}] = \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+k}$$

**Conclusion :** Le processus  $\{Y_t\}$  est faiblement stationnaire, avec  $\mathbb{E}[Y_t] = 0$  et la fonction d'autocovariance qui dépend uniquement de  $k$ , et non de  $t$ .

## 2. Spectre de puissance de $\{Y_t\}$ :

Le spectre de puissance  $S(f)$  est défini comme la transformée de Fourier de la fonction d'autocovariance  $\gamma(\tau)$  :

$$S(f) = \sum_{\tau=-\infty}^{\infty} \gamma(\tau) e^{-2i\pi f\tau}$$

où

$$\gamma(\tau) = \text{Cov}(Y_t, Y_{t-\tau}) = \mathbb{E}[Y_t Y_{t-k}]$$

En utilisant l'expression pour  $\gamma(\tau) = \sigma_\epsilon^2 \sum_{k=0}^{\infty} \psi_k \psi_{k+\tau}$ , nous avons :

$$\begin{aligned} S(f) &= \sum_{\tau=-\infty}^{\infty} \gamma(\tau) e^{-2i\pi f\tau} \\ &= \sigma_\epsilon^2 \sum_{\tau=-\infty}^{\infty} \left( \sum_{k=0}^{\infty} \psi_k \psi_{k+\tau} \right) e^{-2i\pi f\tau} \\ &= \sigma_\epsilon^2 \left| \sum_{k=0}^{\infty} \psi_k e^{-2i\pi f k} \right|^2 \\ &= \sigma_\epsilon^2 \left| \phi(e^{-2i\pi f}) \right|^2 \end{aligned}$$

où  $\phi(z) = \sum_{k=0}^{\infty} \psi_k z^k$  est la fonction génératrice des coefficients  $\psi_k$ .

car,

$$\begin{aligned} \left| \sum_{j=0}^N \psi_j e^{-2i\pi f j} \right|^2 &= \left( \sum_{k=0}^N \psi_k e^{-2i\pi f k} \right) \left( \sum_{l=0}^N \psi_l e^{2i\pi f l} \right) \\ &= \sum_{k=0}^N \sum_{l=0}^N \psi_k \psi_l e^{-2i\pi f(k-l)} \\ &= \sum_{\tau=-N+1}^{N-1} \sum_{n=0}^{N-\tau-1} \psi_n \psi_{n+\tau} e^{-2i\pi f \tau} \end{aligned}$$

Donc,

$$\left| \sum_{j=0}^{\infty} \psi_j e^{-2i\pi f j} \right|^2 \xrightarrow{n \rightarrow \infty} \sum_{\tau=-\infty}^{\infty} \sum_{n=0}^{\infty} \psi_n \psi_{n+\tau} e^{-2i\pi f \tau}$$

**Conclusion :** Le spectre de puissance du processus  $\{Y_t\}_t$  est donné par  $S(f) = \sigma_\epsilon^2 \left| \phi(e^{-2i\pi f}) \right|^2$ , où  $\phi(z) = \sum_{k=0}^{\infty} \psi_k z^k$  est la fonction génératrice des coefficients  $\psi_k$ .

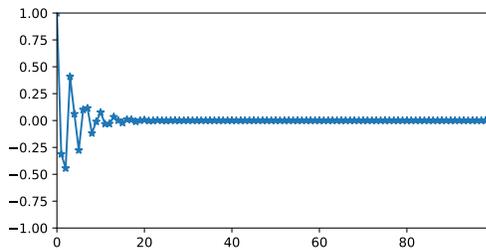
### Question 3 AR(2) process

Let  $\{Y_t\}_{t \geq 1}$  be an AR(2) process, i.e.

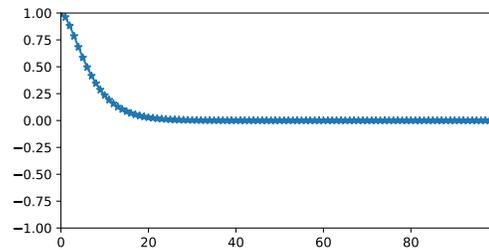
$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t \quad (2)$$

with  $\phi_1, \phi_2 \in \mathbb{R}$ . The associated characteristic polynomial is  $\phi(z) := 1 - \phi_1 z - \phi_2 z^2$ . Assume that  $\phi$  has two distinct roots (possibly complex)  $r_1$  and  $r_2$  such that  $|r_i| > 1$ . Properties on the roots of this polynomial drive the behavior of this process.

- Express the autocovariance coefficients  $\gamma(\tau)$  using the roots  $r_1$  and  $r_2$ .
- Figure 1 shows the correlograms of two different AR(2) processes. Can you tell which one has complex roots and which one has real roots?
- Express the power spectrum  $S(f)$  (assume the sampling frequency is 1 Hz) using  $\phi(\cdot)$ .
- Choose  $\phi_1$  and  $\phi_2$  such that the characteristic polynomial has two complex conjugate roots of norm  $r = 1.05$  and phase  $\theta = 2\pi/6$ . Simulate the process  $\{Y_t\}_t$  (with  $n = 2000$ ) and display the signal and the periodogram (use a smooth estimator) on Figure 2. What do you observe?



Correlogram of the first AR(2)



Correlogram of the second AR(2)

Figure 1: Two AR(2) processes

### Answer 3

1. Déterminons la fonction d'autocovariance  $\gamma(\tau)$  :

Nous partons de la définition de l'autocovariance :

$$\begin{aligned} \gamma(\tau) &= \mathbb{E}[Y_t Y_{t-\tau}] && \text{(par définition)} \\ &= \mathbb{E}[(\phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t) Y_{t-\tau}] \\ &= \mathbb{E}[\phi_1 Y_{t-1} Y_{t-\tau} + \phi_2 Y_{t-2} Y_{t-\tau} + \varepsilon_t Y_{t-\tau}] \\ &= \phi_1 \mathbb{E}[Y_{t-1} Y_{t-\tau}] + \phi_2 \mathbb{E}[Y_{t-2} Y_{t-\tau}] + \mathbb{E}[\varepsilon_t Y_{t-\tau}] \\ &= \phi_1 \gamma(\tau - 1) + \phi_2 \gamma(\tau - 2) + \mathbb{E}[\varepsilon_t Y_{t-\tau}] \end{aligned}$$

Or, étant donné que  $\varepsilon_t$  est un bruit blanc et indépendant de  $Y_{t-\tau}$ , nous avons :

$$\mathbb{E}[\varepsilon_t Y_{t-\tau}] = 0$$

Nous obtenons ainsi la relation de récurrence suivante pour la fonction d'autocovariance  $\gamma(\tau)$  :

$$\gamma(\tau) = \phi_1 \gamma(\tau - 1) + \phi_2 \gamma(\tau - 2)$$

Cette relation est une **suite récurrente linéaire d'ordre 2**.

L'équation caractéristique associée à cette suite récurrente est donnée par :

$$r^2 - \phi_1 r - \phi_2 = 0 \quad (3)$$

Cette équation caractéristique correspond à  $\phi\left(\frac{1}{r}\right) = 0$ , ce qui nous permet de relier directement les racines  $r_1$  et  $r_2$  à l'équation caractéristique (qui sont supposées distinctes).

En effet, si on note  $x_1$  et  $x_2$  les deux solutions de l'équation caractéristique (3), on remarque aisément qu'on a alors :

$$r_1 = \frac{1}{x_1} \quad \text{et} \quad r_2 = \frac{1}{x_2}$$

Ainsi, les solutions de la suite récurrente sont de la forme :

$$\gamma(\tau) = C_1 x_1^\tau + C_2 x_2^\tau = C_1 r_1^{-\tau} + C_2 r_2^{-\tau}$$

où  $C_1$  et  $C_2$  sont des constantes déterminées par les conditions initiales  $\gamma(0)$  et  $\gamma(1)$ . En particulier, nous avons :

$$\gamma(0) = C_1 + C_2 \quad \text{et} \quad \gamma(1) = C_1 r_1 + C_2 r_2$$

**Conclusion :** Les coefficients d'autocovariance du processus AR(2) peuvent être exprimés sous la forme :

$$\gamma(\tau) = C_1 r_1^{-\tau} + C_2 r_2^{-\tau}$$

où  $r_1$  et  $r_2$  sont les racines du polynôme caractéristique, et  $C_1, C_2$  sont déterminés par les conditions initiales  $\gamma(0)$  et  $\gamma(1)$ .

2. Nous savons maintenant que la fonction d'autocovariance s'écrit  $\gamma(\tau) = C_1 r_1^{-\tau} + C_2 r_2^{-\tau}$ .

Comme  $|r_1| > 1$  et  $|r_2| > 1$ , on remarque facilement qu'on a bien :

$$\gamma(\tau) \xrightarrow{\tau \rightarrow +\infty} 0$$

Cette écriture de  $\gamma(\tau)$  permet de mieux comprendre les résultats des corrélogrammes donnés dans la figure (1) :

- **Premier AR(2) - Racines complexes :** le corrélogramme montre un comportement oscillatoire atténué. Cela est dû à la partie imaginaire des racines, qui introduit des oscillations périodiques.
- **Second AR(2) - Racines réelles :** le corrélogramme diminue de manière monotone, ce qui reflète une décroissance purement liée aux valeurs absolues des racines.

3. Calculons le spectre de puissance du processus AR(2)  $\{Y_t\}_{t \geq 1}$ .

Pour cela, on pose  $s(z) = \sum_{k=0}^{+\infty} \gamma(k)z^k$ . On peut alors remarquer que :

$$s(z) = \frac{\gamma(0) - z\phi_1\gamma(0) + \gamma(1)z}{\phi(z)}$$

En effet, on a :

$$\begin{aligned} s(z) &= \sum_{k=0}^{+\infty} \gamma(k)z^k \\ &= \gamma(0) + \gamma(1)z + \sum_{k=2}^{+\infty} \gamma(k)z^k \\ &= \gamma(0) + \gamma(1)z + \sum_{k=2}^{+\infty} \phi_1\gamma(k-1)z^k + \sum_{k=2}^{+\infty} \phi_2\gamma(k-2)z^k \\ &= \gamma(0) + \gamma(1)z + \phi_1z \sum_{k=2}^{+\infty} \gamma(k-1)z^{k-1} + \phi_2z^2 \sum_{k=2}^{+\infty} \gamma(k-2)z^{k-2} \\ &= \gamma(0) + \gamma(1)z + \phi_1z(s(z) - \gamma(0)) + \phi_2z^2s(z) \\ &= s(z) [\phi_1z + \phi_2z^2] + \gamma(1)z + \phi(0)(1 - \phi_1z) \end{aligned}$$

On en déduit alors le résultat.

Et alors :

$$S(f) = \sum_{\tau=-\infty}^{+\infty} \gamma(\tau) e^{-2i\pi f\tau} = s(e^{2i\pi f}) + s(e^{-2i\pi f}) - \gamma(0)$$

Après simplification, et en posant  $\sigma^2 = \gamma(0)$ , on a ainsi :

$$S(f) = \frac{\sigma^2}{|\phi(e^{-2i\pi f})|^2}$$

4. Pour commencer, exprimons  $\phi_1$  et  $\phi_2$  en fonction de  $r_1$  et  $r_2$  en sachant que  $\phi(r_1) = \phi(r_2) = 0$

$$\begin{aligned} \begin{cases} \phi(r_1) = 0 \\ \phi(r_2) = 0 \end{cases} &\implies \begin{cases} 1 - \phi_1r_1 - \phi_2r_1^2 = 0 \\ 1 - \phi_1r_2 - \phi_2r_2^2 = 0 \end{cases} \\ &\implies \begin{cases} \phi_1(r_2 - r_1) + \phi_2(r_2^2 - r_1^2) = 0 \\ 1 - \phi_1r_2 - \phi_2r_2^2 = 0 \end{cases} \\ &\implies \begin{cases} \phi_1 = -\phi_2(r_2 + r_1) \\ 1 + \phi_2(r_2 + r_1)r_1 - \phi_2r_1^2 = 0 \end{cases} \implies \begin{cases} \phi_1 = \frac{r_1+r_2}{r_1r_2} \\ \phi_2 = \frac{-1}{r_1r_2} \end{cases} \end{aligned}$$

Avec  $r_{1/2} = re^{\pm i\theta}$ , on en déduit que :

$$\phi_1 = \frac{2 \cos(\theta)}{r} \quad \text{et} \quad \phi_2 = \frac{-1}{r^2}$$

Et voici les résultats que l'on obtient :

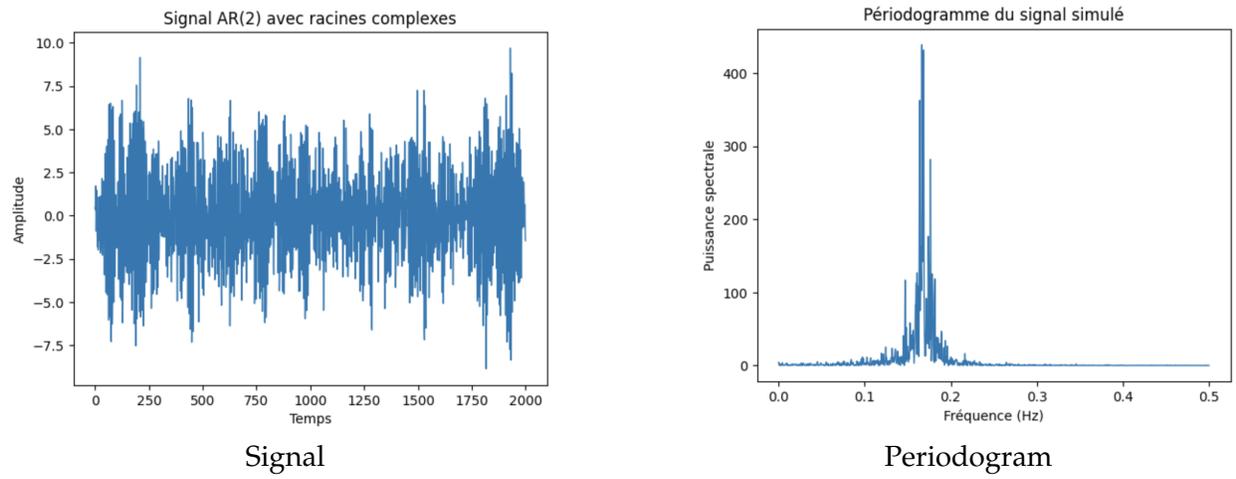


Figure 2: AR(2) process

## 4 Sparse coding

The modulated discrete cosine transform (MDCT) is a signal transformation often used in sound processing applications (for instance, to encode an MP3 file). A MDCT atom  $\phi_{L,k}$  is defined for a length  $2L$  and a frequency localisation  $k$  ( $k = 0, \dots, L - 1$ ) by

$$\forall u = 0, \dots, 2L - 1, \quad \phi_{L,k}[u] = w_L[u] \sqrt{\frac{2}{L}} \cos\left[\frac{\pi}{L} \left(u + \frac{L+1}{2}\right) \left(k + \frac{1}{2}\right)\right] \quad (4)$$

where  $w_L$  is a modulating window given by

$$w_L[u] = \sin\left[\frac{\pi}{2L} \left(u + \frac{1}{2}\right)\right]. \quad (5)$$

### Question 4 Sparse coding with OMP

For the signal provided in the notebook, learn a sparse representation with MDCT atoms. The dictionary is defined as the concatenation of all shifted MDCT atoms for scales  $L$  in  $[32, 64, 128, 256, 512, 1024]$ .

- For the sparse coding, implement the Orthogonal Matching Pursuit (OMP). (Use convolutions to compute the correlation coefficients.)
- Display the norm of the successive residuals and the reconstructed signal with 10 atoms.

### Answer 4

L'algorithme **Orthogonal Matching Pursuit (OMP)** consiste à reconstruire un signal  $x$  en utilisant une combinaison parcimonieuse d'atomes du dictionnaire  $D$ .

- **Initialisation:**

- $r \leftarrow x$  ▷ Initialisation du résidu
- $A \leftarrow \emptyset$  ▷ Matrice des atomes sélectionnés
- $\mathbf{z}_K \leftarrow \mathbf{0}$  ▷ Vecteur des coefficients initiaux
- $\text{atoms\_idx} \leftarrow \emptyset$  ▷ Liste des indices des atomes sélectionnés

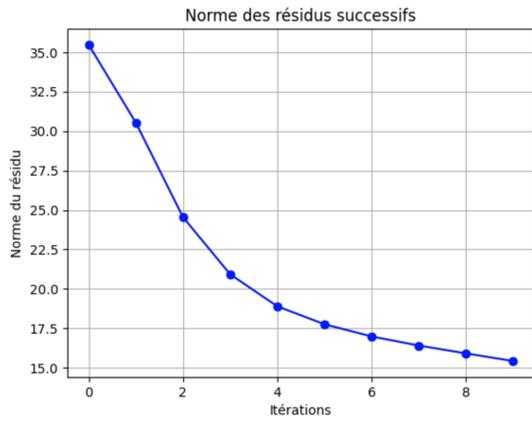
- **Répéter pour  $k = 1$  à  $10$ :**

- $j_{\max} \leftarrow \arg \max_{j \notin \text{atoms\_idx}} |\langle r, D_j \rangle|$  ▷ Sélectionner l'atome  $j_{\max}$  avec la plus grande corrélation avec le résidu
- $\text{atoms\_idx} \leftarrow \text{atoms\_idx} \cup \{j_{\max}\}$  ▷ Marquer l'atome comme utilisé
- $A \leftarrow [A, D_{j_{\max}}]$  ▷ Ajouter l'atome sélectionné à la matrice  $A$
- Résoudre le problème des moindres carrés pour trouver les coefficients:

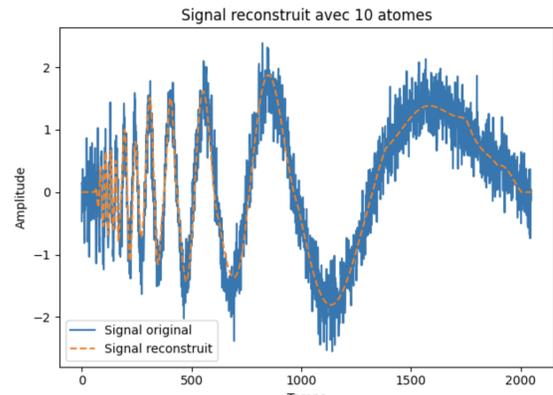
$$\mathbf{z}_k = \underset{\mathbf{z}}{\operatorname{argmin}} \|x - A\mathbf{z}\|_2^2$$

- $r \leftarrow x - A\mathbf{z}_k$  ▷ Mettre à jour le résidu,  $A\mathbf{z}_k$  représentant le signal reconstruit

- **Sortie:**
  - $A$  : Matrice des atomes sélectionnés
  - $\mathbf{z}_K$  : Coefficients associés aux atomes
  - $r$  : Résidu final



Norms of the successive residuals



Reconstruction with 10 atoms

Figure 3: Question 4