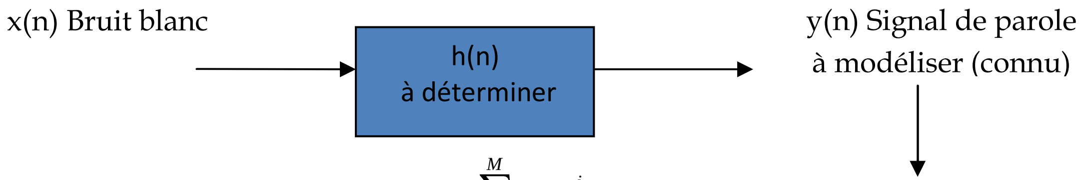
Chapitre 4 : Processus générateurs AR, MA et ARMA

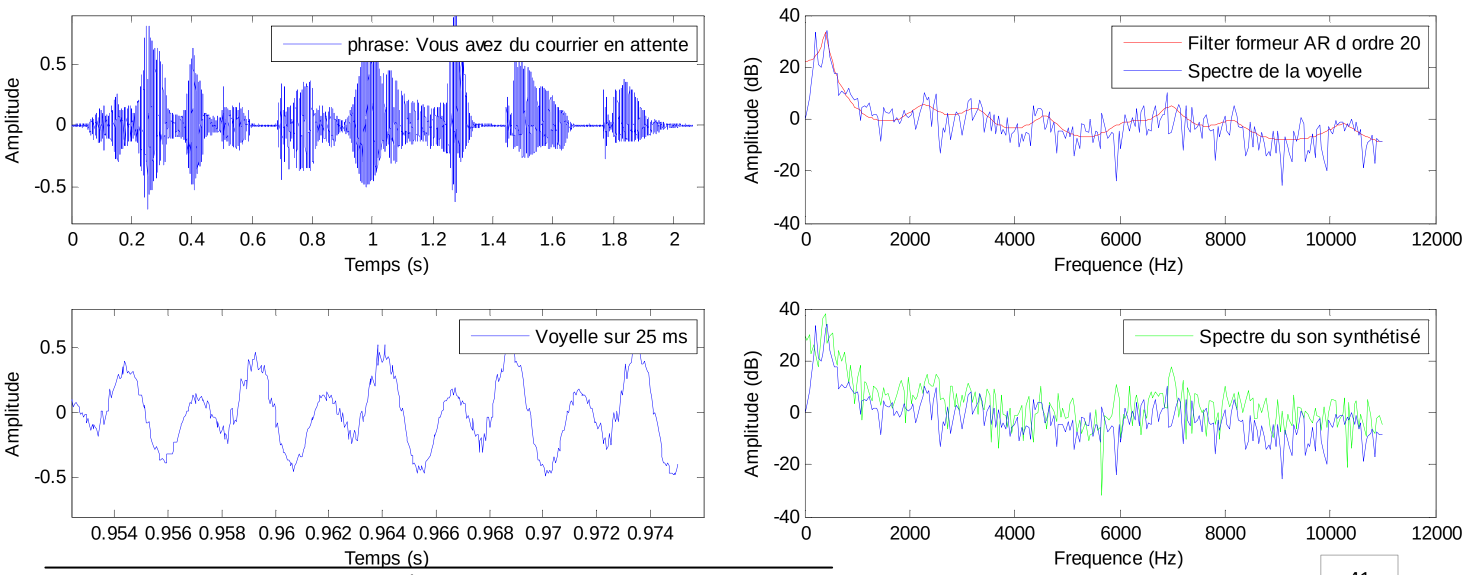
Nous avons vu au chapitre précédent qu'un signal aléatoire peut être modélisé (synthétisé) comme la réponse d'un filtre linéaire à une excitation sous forme de bruit blanc tel que $|H(f)|^2 = S_x(f)/\sigma_b^{\ 2}$. Ce filtre formeur H(f) est aussi dit processus générateur. Ces paramètres associés à la variance du bruit σ_b^2 constituent le modèle mathématique correspondant au signal aléatoire. Le concept de processus générateur de signal a été particulièrement développé et appliqué avec des filtres numériques. Ce sont les modèles de signaux les plus utilisés en traitement statistique du signal (estimation, prédiction...).Selon la nature du filtre, on peut obtenir différents modèles de signaux (AR,MA, ARMA, etc.)

Exemple : Lors d'un appel par GSM (téléphone portable), le portable qui fait office d'un micro- ordinateur regroupant différentes fonctionnalités dont l'analyse, la synthèse, le codage, etc. va nous permettre de modéliser la parole (aléatoire) en opérant un codage LPC par tranches de dizaines de ms. Elle consiste à retrouver les paramètres du filtre formeur *h(n)* pour chaque tranche *y(n)* enregistrée et analysée. Ce sont ces paramètres (a_i et b_j) qui seront transmis pour produire un signal de synthèse approchant le signal original y(n).



$$H(z) = \frac{\sum_{j=0}^{M} b_j z^{-j}}{\sum_{i=0}^{M} a_i z^{-i}} \qquad S_y(f) = \sigma_b^{\ 2} |H(f)|^2$$

Ci-dessous la phrase "vous avez du courrier en attente" échantillonnée à une fréquence fe =22050 (45531 échantillons) et un zoom sur une voyelle de durée 25 ms (500 échantillons).



En comparant l'allure du filtre formeur H(f) à celle de Sy(f), on note que l'on retrouve l'allure générale du spectre de la voyelle, notamment les fréquences dont la puissance est maximale. Sachant que le spectre de la voyelle comporte 500 valeurs et que le filtre H(f) équivalent est obtenu à partir de 20 coefficients, il vaut mieux transmettre 21 coefficients (20 a_i + variance du bruit) que 500 valeurs.

C'est ainsi que les modèles autorégressifs sont d'un emploi de plus en plus répandu en traitement du signal : codage et transmission par prédiction linéaire, synthèse de parole, reconnaissance, etc.

Remarque : Le signal de parole est un processus aléatoire non-stationnaire à long terme, mais il est considéré comme stationnaire dans des fenêtres temporelles d'analyse de l'ordre de 20 à 30ms. Cette propriété de stationnarité à court terme permet donc une analyse et modélisation progressive du signal de parole Pour éviter toutes pertes d'information, on veillera à prendre des fenêtres chevauchantes.

1. Modèle auto-régressif (AR)

Les signaux autorégressifs sont obtenus par passage d'un bruit blanc dans un filtre purement récursif. Ce filtre est donc de réponse impulsionnelle infinie.

$$H(z) = 1/\left(1+\sum_{i=1}^{N} a_i z^{-i}\right)$$

A partir de H(z), on peut déterminer l'équation aux différences : $y(n) = x(n) - \sum_{i=1}^{N} a_i y(n-i)$

Cela signifie que le signal y(n) est supposé être prédictible en fonction d'un certain nombre de ses valeurs antérieures.

Sachant que x(n) est un bruit blanc alors d'où $R_x(k) = \sigma^2 \delta(k)$

- $\mu_x = E\{x(n)\} = 0$
- $R_{xx}(0) = E\{x(n)^2\} = \sigma^2$
- $R_{xx}(k) = E\{x(n)x(n-k)\} = 0 \quad pour \quad k \neq 0$

Calculons alors $R_{yy}(k)$

$$R_{yy}(k) = E\{y(n)y(n-k)\} = E\{x(n)y(n-k)\} - \sum_{i=1}^{N} a_i E\{y(n-i)y(n-k)\}$$

$$= E\left\{x(n)\sum_{k'} x(n-k-k')\,h(k')\right\} - \sum_{i=1}^{N} a_i R_{yy}(k-i) = \sum_{k'} E\{x(n)x(n-k-k')\}h(k') - \sum_{i=1}^{N} a_i R_{yy}(k-i)$$

$$= \sum_{k'} R_{xx}(k+k')\,h(k') - \sum_{i=1}^{N} a_i R_{yy}(k-i) = R_{xx}(k)h(0) + R_{xx}(k+1)h(1) + \ldots\ldots - \sum_{i=1}^{N} a_i R_{yy}(k-i)$$

-Si k=0, $R_{yy}(0) = R_{xx}(0).h(0) - \sum_{i=1}^{N} a_i R_{yy}(-i) = \sigma^2.1 - \sum_{i=1}^{N} a_i R_{yy}(i)$

- Pour k=1 à N, $R_{yy}(k) = 0 - \sum_{i=1}^{N} a_i R_{yy}(k-i) = -\sum_{i=1}^{N} a_i R_{yy}(k-i)$

On peut utiliser une forme matricielle : R. a = S (Rappelons que pour un signal réel $R_{yy}(k)=R_{yy}(-k)$

$$\begin{bmatrix} R_{yy}(0) & R_{yy}(1) & \ldots\ldots & R_{yy}(N) \\ R_{yy}(1) & R_{yy}(0) & \ldots\ldots & R_{yy}(N-1) \\ & & \ldots\ldots & \\ R_{yy}(N) & R_{yy}(N-1) & \ldots\ldots & R_{yy}(0) \end{bmatrix} . \begin{bmatrix} 1 \\ a_1 \\ . \\ a_N \end{bmatrix} = \begin{bmatrix} \sigma^2 \\ 0 \\ . \\ 0 \end{bmatrix}$$

R est la matrice d'autocorrélation dont le terme général r_{ij} ne dépend que de la différence *i-j* (Matrice de Toeplitz). La résolution de ces équations dite de Yule-Walker permet de connaître les paramètres du filtre et la variance du bruit blanc.

Remarques :

- Nous ne disposons pas d'un processus aléatoire mais d'une seule réalisation soit *y(n),* il n'est pas donc pas possible de calculer l'auto-corrélation statistique $R_{yy}(k)$. Cette dernière sera remplacée par l'autocorrélation temporelle en faisant l'hypothèse que le processus est ergodique (voir exo 1 du TP n°4).

- Il existe divers algorithmes (Burg, Levinson) qui permettent d'estimer assez rapidement les a_i et σ sans passer par l'inversion matricielle. Tout comme il est possible de déterminer l'ordre N adéquat (critère AIC).

Exemple d'application

1. On considère le modèle auto-régressif (AR) d'ordre 1 tel que : $x(n)=-a_1.x(n-1)+b(n)$

- Déterminer les équations de Yule-Walker pour ce modèle
- En supposant que x(n) est connu, déterminer les paramètres du modèle.
- Déterminer les $R_x(k)$ (les $a_i$sont supposés connus)

Réponses :

- ($a_1=-Rx(1)/Rx(0)$ $\sigma^2=Rx(0)(1-a_1^2)$)
- ($Rx(0)= \sigma^2/(1-a_1^2)$ $Rx(1)=-a_1\sigma^2/(1-a_1^2)$$Rx(k)=(-a_1)^k\sigma^2/(1-a_1^2)$)

2. Refaire le même travail pour un modèled'ordre 2 tel que : $x(n)=-a_1.x(n-1)-a_2.x(n-2)+b(n)$

Réponses :

- $a_1=Rx(1)[Rx(2)-Rx(1)]/[Rx(0)^2-Rx(1)^2]$ $a_2=[Rx(1)^2-Rx(0)Rx(2)]/[Rx(0)^2-Rx(1)^2]$
- $\sigma^2=Rx(0)+Rx(1)^2[Rx(2)-Rx(0)]/[Rx(0)^2-Rx(1)^2]+Rx(2)[Rx(1)^2-Rx(0)Rx(2)]/[Rx(0)^2-Rx(1)^2]$
- $Rx(1)=-a_1Rx(0)/(a_2+1)$ $Rx(2)=(-a_2+a_1^2/(1+a_2))Rx(0)$$Rx(0)=(1+a_2)\sigma^2/(1+a_2- a_1^2–a_1^2-a_2^3+a_2a_1^2)$

2.Modèle à moyenne ajustée (MA)

Les signaux à moyenne mobile sont obtenus par passage d'un bruit blanc dans un filtre purement transverse.

Ce filtre est aussi appelé filtre à réponse impulsionnelle finie : $H(z) = \sum_{i=0}^{M} b_i z^{-i}$

Le signal *y(n)* est supposé pouvoir s'écrire comme une combinaison linéaire d'échantillons décorrélés entre eux, ce qui peut se formaliser comme une combinaison linéaire d'échantillons d'un bruit blanc *x(n)*.

On a donc : $y(n) = \sum_{i=0}^{M} b_i x(n-i)$

et $\mu_y = E\{y(n)\} = \sum_{i=0}^{M} b_i \mu_x = \mu_x \sum_{i=0}^{M} b_i = \mu_x . H_f(0)$

On cherche les paramètres du filtre qui génèrent y(t) à partir de x(t), bruit blanc centré :

$$R_{yy}(k) = E\{y(n)y(n-k)\} = E\left\{\sum_{i=0}^{M} b_i x(n-i) . \sum_{j=0}^{M} b_j x(n-j-k)\right\}.$$

$$R_{yy}(k) = \sum_{i=0}^{M} b_i . \sum_{j=0}^{M} b_j E\{x(n-i).x(n-j-k)\} = \sum_{i=0}^{M} b_i . \sum_{j=0}^{M} b_j R_{xx}(j+k-i)$$

- Si j+k≠i $\Rightarrow R_{yy}(k) = 0$
- Sinon $\Rightarrow R_{yy}(k) = \sigma^2 \sum_{j=0}^{M-k} b_{j+k} . b_j$

Le problème est non linéaire en fonction des coefficients, il faut un algorithme de programmation non linéaire pour obtenir b_i à partir des $R_{yy}(k)$. Cependant, l'algorithme de Durbin permet d'approcher la solution optimale avec de bons résultats. Le principe de cet algorithme consiste à identifier le modèle MA d'ordre M avec un modèle AR d'ordre N>>M (voir exo 2 du TP n°4). En effet, tout modèle MA peut être identifié à un modèle AR d'ordre infini: $\sum_{i=0}^{M} b_i z^{-i} = 1 / \sum_{i=0}^{\infty} a_i z^{-i}$

<u>Exemple</u>

On considère le modèle à moyenne ajustée (MA) :

A. d'ordre 1 tel que $x(n)=e(n)+b_1 . e(n-1)$
- Calculer μ_x.
- En supposant que x(n) est connu, déterminer les paramètres du modèle.
- Connaissant les paramètres du modèle, déterminer les $R_x(k)$

B. d'ordre 2 tel que : $x(n)=e(n)+b_1. e(n-1)+b_2. e(n-2)$

- Calculer μ_x.
- Connaissant les paramètres du modèle, déterminer les $R_x(k)$

Réponses

$\mu_x = 0$ $Rx(0)=(1+ b_1^2).\sigma^2$ $Rx(1)=b_1.\sigma^2$ $Rx(k)=0$ pour $k\geq 2$

$b_1=(Rx(0)\pm sqrt(Rx(0)^2-4Rx(1)^2)/2Rx(1)$ $\sigma^2=2/(Rx(0)\pm sqrt(Rx(0)^2-4Rx(1)^2)$

$Rx(0)=(1+ b_1^2+ b_2^2).\sigma^2$ $Rx(1)=(b_1+ b_1 b_2).\sigma^2$ $Rx(2)=b_2.\sigma^2 Rx(k)=0$ pour $k\geq 3$

<u>*Remarque*</u> : Il est très important de remarquer que nous ne disposons que d'une seule réalisation du signal aléatoire à modéliser *y(n)*, de ce fait l'auto-corrélation statistique $R_{yy}(k)$ est obtenu de l'auto-corrélation temporelle en considérant le processus ergodique.

3. Modèle ARMA

Les signaux ARMA sont obtenus par passage d'un bruit blanc dans un filtre récursif appelé aussi filtre à réponse impulsionnelle infinie (R.I.I). Ces signaux sont une combinaison des signaux AR etMA. La fonction de transfert du filtre présente un numérateur et un dénominateur:

Soit $y(n)=\sum_{i=0}^{M} b_i x(n-i)-\sum_{i=1}^{N} a_i y(n-i)$ $\qquad H(z)=\sum_{i=0}^{M} b_i z^{-i} \Big/ \left(1+\sum_{i=1}^{N} a_i z^{-i}\right)$

La corrélation statistique de y(n) s'écrit alors : $R_{yy}(k)=-\sum_{i=1}^{N} a_i R_{yy}(k-i)+\sigma^2\sum_{j=0}^{M-k} b_{j+k}.b_j$

C'est une équation non linéaire en a_i et b_j.

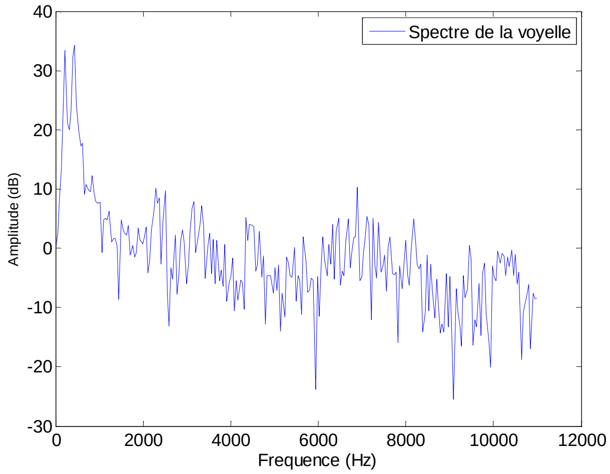
La modélisation ARMA peut se décomposer en une modélisation AR suivie d'une modélisation MA. Le modèle AR présente une simplicité de calcul par rapport aux modèles MA et ARMA du fait où les coefficients AR sont solutions d'un système linéaire d'équations. Alors que la détermination des coefficients MA et ARMA requiert la résolution d'équations non linéaires. Cependant, le modèle ARMA permet de modéliser aussi bien les minima que les maxima de la DSP et est donc moins restrictif que le modèle AR.

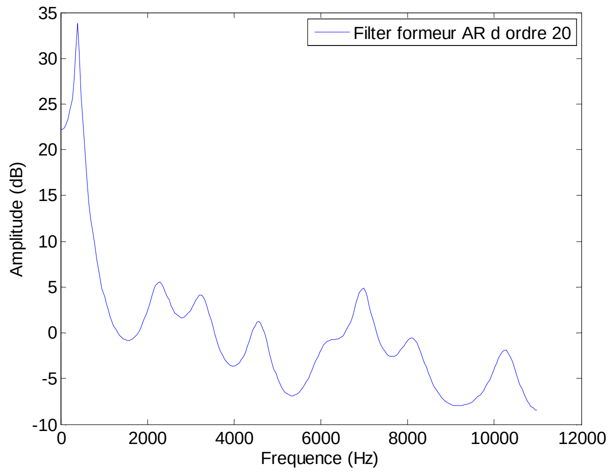
Les application des modèles AR, MA, ARMA sont nombreuses, entre autres :

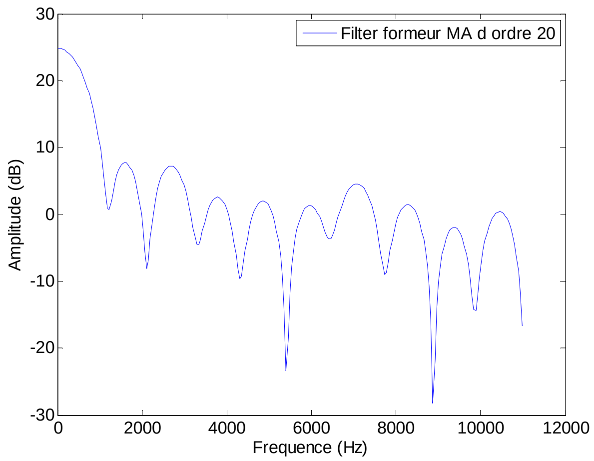
- la modélisation et la prédiction de série temporelle dite séries chronologiques Une série chronologique est une suite formée d'observations au cours du temps que l'on cherche à modéliser pour la prédiction de données futures. Ainsi, en finance, cela permet de modéliser le cours des devises ou du pétrole. Alors qu'en météorologie, les cela permet de faire des prévisions sur la température ou les précipitations. Dans chacun des cas, on essaiera à partir d' un échantillon de données de construire le meilleur modèle qui s'ajuste ces données.

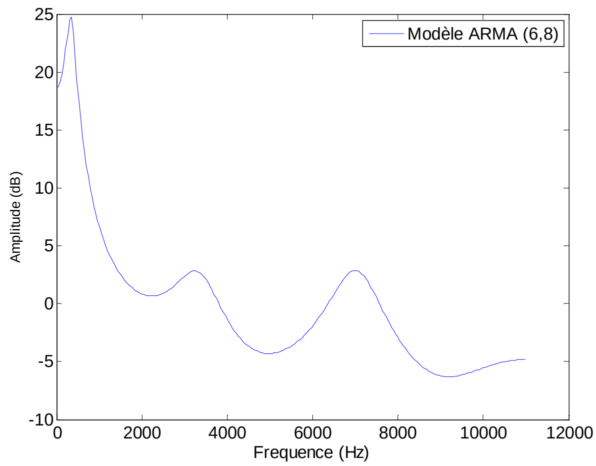
- l'estimation du spectre d'un signal aléatoire, etc. Cette dernière application est basée sur l'identification des paramètres du modèle considéré: Le modèle AR est bien adapté aux signaux composés de raies pures dans du bruit blanc. Alors que le modèle MA est bien adapté aux signaux dont la puissance est nulle dans certaines bandes de fréquences.

Exemple : Reprenons à nouveau l'exemple de parole, pour lequel, nous avons testé les trois variantes









Nous n'irons pas plus loin, dans le cadre de ce cours, le but n'étant que d'introduire les notions de modélisation et prédiction linéaire.

TD n°4 : Modèles AR, MA, ARMA

1. Soit un signal aléatoire y(n) SSL et ergodique dont l'autocorrélation temporelle $\overline{R}_y(k) = \alpha^{|k|}$ avec $0<\alpha<1$

- Identifier le modèle linéaire adéquat (AR ou MA) pour y(n).
- En supposant que le système h(n) est filtre purement récursif, donner le schéma du modèle en définissant l'entrée, le système, et la sortie.
- Rappeler les hypothèses nécessaires liées à l'emploi de ce modèle.
- On considère que le modèle est d'ordre 1, déterminer ses paramètres.

2. On considère le filtre linéaire à temps discret défini par $y(n) = x(n) + b_1 x(n-1) + b_2 x(n-2)$.
où X(n) et Y (n) désignent respectivement les processus aléatoires réels d'entrée et de sortie du filtre où b1 et b2 sont 2 coefficients réels. On suppose que x(n) est une suite de variables aléatoires centrées**,** indépendantes et de varianceσ^2**.**

- Donner l'expression de Rx(k) et Sx(f) puis donner l'expression de Ry(k) et tracer la pour $b_1=1$ et $b_2=-1$.
- Connaissant la DSP du signal y(n), sur quoi se base-t-on pour le choix du modèle ?

3. Soit un processus AR défini par : $y(n)=-a_1\, y(n-1) - a_2\, y(n-2) + x(n)$ où x(n) bb décorrélé de variance 1

- Calculer $\mu_y(n)$ puis sans calcul, expliquer pourquoi y(n) est SSL.
- Montrer que pour k>0, $Ry(k)=-a_1\, Ry(k-1)-a_2\, Ry(k-2)$
- Déterminer $a_1$et a_2

4. Soit un filtre formeur dont l'équation aux différences est y(n)= 0.25 y(n-1) -0.25 y(n-2) +x(n)
- Identifier ce modèle linéaire AR ou MA (Justifier)
- Déterminer la moyenne de y(n) et donner l'expression de $R_{yy}(k)$.
- Calculer et tracer $R_{yy}(k)$ (Prendre $R_{xx}(0)=\sigma^2=1$).

Solutions
1. Modèle AR ($R_y(k)\neq 0$)
x(n) entrée bb, y(n) signal aléatoire à modéliser h(n) filtre formeur (modèle math)
entrée bb + ergodisme $a_1= - \alpha$ $\sigma_x^2=1-\alpha^2$
2.$R_x(k)=\sigma_x^2\delta(k)$ $S_x(f)=\sigma_x^2$ MA d'ordre 2 $R_y(0)=1+b_1^2+b_2^2 R_y(1)=b_1(1+b_2)$ $R_y(2)= b_2 R_y(k)= 0$ pour $k\geq 0$
3. $\mu_y(n)=0$ entrée bb SSL $\Rightarrow$sortie SSL voir cours (page 41)

Exercices supplémentaires
1.Les signaux x(n) et y(n) ont été obtenus en filtrant, au moyen d'un filtre à réponse impulsionnelle finie, un bruit blanc b(n) gaussien centré de variance σ^2.

A]L'équation de filtrage est la suivante : x(n) = 2.b(n) + 0.5.b(n -1) -0.2.b(n - 2) + 0.1.b(n - 3)
- Calculez, en fonction de σ^2, les coefficients d'autocorrélation d'ordre 0,1,2,3 du signal x(n) .
- On notera Rxx (0), Rxx (1), Rxx (2), Rxx (3) ces coefficients.
- La répartition des niveaux d'amplitude du signal x(n) est elle gaussienne (sans justifier)?

B]y(n) = b(n) - b(n - 2)
- Donnez la TZ de la réponse impulsionnelle du filtre qui a permis d'obtenir y(n) à partir de b(n).
- Placez les zéros de ce filtre sur un cercle unité, quelles sont la ou les fréquence(s) coupées par ce filtre ?
- Tracez approximativement sa réponse en fréquence.

C]On considère maintenant le coefficient d'intercorrélation Rxy(k), entre les signaux x(n) et y(n) donné par $Rxy(k) = E[x(n)\, y(n - k)^*]$.
Calculez Rxy (0), Rxy (1), Rxy (2), Rxy(-1), Rxy(-2)

2.On considère un signal aléatoire stationnaire x(n) et l'on suppose connu ses coefficients d'autocorrélation :
$R(0) = 3\sigma^2$, $R(1) = 2\sigma^2$, $R(2) = \sigma^2$, $R(3) = 0$
- On cherche le filtre MA d'ordre 3 de ce signal. Identifiez les paramètres du filtre.
- Si le signal x(n) a été obtenu par filtrage d'un bruit blanc gaussien de variance σ^2 par un filtre à réponse impulsionnelle finie de fonction de transfert $H(z) = 1+ a.z^{-1} + b.z^{-2}$, en déduire les valeurs de a et b.

3. Soit un filtre formeur dont l'équation aux différences est $y(n)=0.5\,(x(n)+x(n-1)+x(n-2)+x(n-3))$
- Expliquer la notion de filtre formeur
- Identifier ce modèle linéaire AR ou MA
- Donner la moyenne, l'autocorrélation et la DSP de son entrée *x(n).*
- Calculer et tracer $R_{yy}(k)$ puis déterminer $S_{yy}(f)$

4. Les signaux x(n) et y(n) ont été obtenus en filtrant, au moyen d'un filtre à réponse impulsionnelle finie, un bruit blanc b(n) gaussien centré de variance σ^2.
$x(n) = 2.b(n) + 0.5.b(n -1) -0.2.b(n - 2) + 0.1.b(n - 3)$
$y(n) = b(n) - b(n - 2)$
- Identifier les 2 modèles (justifier)
- Pour chacun, calculer et tracer les coefficients d'autocorrélation
- x(n) et y(n) sont-ils Gaussiens (justifier)
- Calculer les intercorrélations Rxy(k), et Ryx(k) et commenter
- Donner quelques applications des modèles AR,MA, ARMA

5. Soit un filtre formeur dont l'équation aux différences est $y(n)= - \alpha\, y(n-1) - \beta\, y(n-2) + x(n)$
- Identifier l'ordre du modèle linéaire AR.
- Déterminer les paramètres du modèles, on suppose que $R_{yy}(k)=2*0.5^{|k|}$

6. Soit un filtre formeur dont l'équation aux différences est $y(n)= \alpha\, y(n-1) + x(n)$
- Identifier l'ordre du modèle linéaire AR.
- Déterminer la moyenne de y(n) et montrer que $R_{yy}(k)=\alpha^k R_{yy}(0)$, déduire une condition sur α.
- Montrer que $R_{yy}(0)=R_{xx}(0)/(1-\alpha^2)$
- Tracer $R_{yy}(k)$ (Prendre $R_{xx}(0)=\sigma^2=1$).
- Citer 2 applications concrètes des modèles AR.

7. Soit un filtre formeur dont l'équation aux différences est $y(n)= 0.25\, y(n-1) - 0.25\, y(n-2) + x(n)$
- Identifier ce modèle linéaire AR ou MA (Justifier)
- Déterminer la moyenne de y(n) et donner l'expression de $R_{yy}(k)$.
- Calculer et tracer $R_{yy}(k)$ (Prendre $R_{xx}(0)=\sigma^2=1$).

Solutions

1.$Rx(0)=(b_0^2+b_1^2+b_2^2+b_3^2).\sigma^2$ $Rx(1)=(b_0b_1+b_1b_2+b_2b_3).\sigma^2$ $Rx(2)=(b_0\, b_2+ b_1\, b_3).\sigma^2$ $Rx(3)=(b_0\, b_3).\sigma^2 Rx(k\geq4)=0$
$H(z)=(z^2-1)/z^2$ $Rxy(0)=2.2$ $Rxy(1)=0.4$ $Rxy(2)=-0.2$ $Rxy(3)=0.1$ $Rxy(-1)=-0.5$ $Rxy(-2)=-2$
2. $b_0=1$ $b_1=1$ et $b_2=1$ etpar identification, on trouve a=1 et b=1
3. MA, $\mu x=0$ et $Sx(f)=\sigma^2$
$Ry(0)=\sigma^2$, $Ry(1)=Ry(-1)=0.75\,\sigma^2$, $Ry(-2)=Ry(2)=0.5\,\sigma^2$, $Ry(-3)=Ry(3)=0.25\,\sigma^2$, $Ry(k>3)=0$.
$Sy(f)=\sigma^2\,(1+1.5\cos(4\pi f)+ \cos(6\pi f) + 0.5\cos(8\pi f))$

TP n° 4 : Modèles AR et MA

Ce TP a pour objectif de modéliser un signal aléatoire par les modèles AR et MA (par approximation AR) et d'en extraire les informations utiles. Puis, une application d'identification sera envisagée.

Exercice 1 :Télécharger le fichier 'vous avez du courrier en attente.wav' et le placer dans le même répertoire que votre programme commençant come suit :

```
clc; clear all; close all;
nom_fich = uigetfile('*.wav', 'Selectionner le fichier son');
% Lire, écouter et afficher le son complet
[x,fe]=wavread(nom_fich);sound(x,fe); t=(0:length(x)-1)/fe; subplot(2,1,1);plot(t,x);
legend('Son');xlabel('Temps (s)');ylabel('Amplitude');
% lire, écouter et afficher une partie du son qui correspond à une voyelle
N1=21000;N2=21500;[y,fe]=wavread(nom_fich,[N1 N2]);sound(y,fe)
N=length(y);t=(N1:N2)/fe; subplot(2,1,2);plot(t,y);
legend('Voyelle sur 25 ms');xlabel('Temps (s)');ylabel('Amplitude');
% Afficher le spectre
Sy=fft(y); f=(0:fe/N:fe/2-1/N);figure;plot(f,20*log10(abs(Sy(1:length(f)))+eps));
legend('Spectre'); xlabel('Frequence (Hz)'); ylabel('Amplitude (dB)');
```

Nous allons commencer par déterminer les paramètres du modèle soit la modélisation

1. Ecrire les instruction nécessaires pour résoudre les équations de Yule-Walker en suivant ces étapes:

a. Montrer qu'on peut les formuler comme suit :

$$\overbrace{\begin{bmatrix} R_{yy}(0) & R_{yy}(1) & \ldots\ldots & R_{yy}(P-1) \\ R_{yy}(1) & R_{yy}(0) & \ldots\ldots & R_{yy}(P-2) \\ & & \ldots\ldots & \\ R_{yy}(P-1) & R_{yy}(P-2) & \ldots\ldots & R_{yy}(0) \end{bmatrix}}^{C} \cdot \begin{bmatrix} a_1 \\ a_2 \\ . \\ a_P \end{bmatrix} = \overbrace{\begin{bmatrix} -R_{yy}(1) \\ -R_{yy}(2) \\ . \\ -R_{yy}(P) \end{bmatrix}}^{B} \quad \sigma^2 = R_{yy}(0) + \sum_{i=1}^{P} a_i R_{yy}(i)$$

b. Pour l'ordre P=20, utiliser la fonctionautocorrpour calculer les valeurs de R_{yy} sur P valeurs.

Pour la matrice C, utiliser la fonction de Toeplitz (Faire attention aux indices).

Par inversion matricielle (inv(C)*B), déterminer les coefficients 'a' (rajouter a_0=1 aux coefficients trouvés) puis déterminer σ^2.

2. Rajouter les instructions permettant de visualiser le filtre formeur en db superposé au spectre de 'y'.

3. L'autocorrélation R_{yy} intervenant dans les équations de Yule-Walker est-elle statistique ou temporelle ? Qu'en est-il de celle calculée dans ce programme? Commenter.

Entamons la partie synthèse en ayant comme données de départ les paramètres du modèle (a_i et sigma)

3. A partir de σ^2 et des coefficients AR 'a' trouvés, synthétiser et écouter le son 'zz' (utiliser randn et filter).

4. Visualiser le son modélisé et le son synthétisé ainsi que leurs spectres en db sur le même graphe.

5. Modifier N et commenter les graphes précédents.

6. Rajouter les lignes suivantes et donner une application possible.

```
r=roots(a);                     % Racines de du dénominateur
r=r(imag(r)>0.01);              % Recherche des pôles positifs (freq>0)
freq_form=sort(atan2(imag(r),real(r))*fe/(2*pi));  % convert en Hz et trier
for i=1:length(freq_form)
fprintf('Formant %d Frequence %.1f\n',i,freq_form(i));
end
```

7. Pourquoi utilise-t-on le modèle AR pour trouver les fréquences principales au lieu du signal 'y'?

Exercice 2 :

Sachant qu'un modèle MA peut être obtenu en identifiant le modèle MA d'ordre M avec un modèle AR d'ordre P>>M :

$$\sum_{i=0}^{M} b_i z^{-i} = 1/\sum_{i=0}^{\infty} a_i z^{-i}$$

1. Modifier le programme précédent en rajoutant l'instruction suivante au bon endroit:`bb=impz(1,a,P);`
2. Expliquer comment on s'est servi de cette instruction pour obtenir les b_i.
3. Visualiser les b_i pour P=60 puis en prendre la moitié puis le tiers et reprendre les questions de 2 à 4 de l'exercice 1 en modifiant les instructions nécessaires.

Exercice 3

Vous avez deux fichiers audio contenant l'enregistrement des 5 voyelles.
1. Lire le premier fichier audio et identifier le début et la fin de chaque voyelle et mettre les signaux obtenus dans une matrice de 5 lignes.
2. Ecrire une boucle, qui pour chaque voyelle, permet de :
- écouter la voyelle, la centrer et la visualiser
- visualiser son spectre
- déterminer les paramètres du modèle AR d'ordre 20 par la commande `[a,sigma]=lpc(x,20);`
- Déterminer les formants à partir du modèle
3. Prendre le deuxième fichier son et lui faire subir les mêmes traitements.
4. Comparer par visualisation les deux signaux correspondant à la même voyelle ainsi que leur spectre et commenter
5. Comparer les formants de la voyelle "a" du premier fichier à ceux de toutes les voyelles du deuxième fichier son. L'identification est-elle concluante?