Traitement statistique du signal

TP 3 : Estimateur ELMQ, reconstruction sonore, biais et variances d'un estimateur

Université Paris 13, Institut Galilée
Master Signal, Images et Multimédia - 2ème année

2013-2014

A préparer à la maison

On considère N échantillons d'un signal sonore

$$\mathbf{X} = [X(1), \ldots, X(N)]^T$$

On suppose que ce signal est stationnaire au sens large, et que sa fonction d'autocorrélation $R_{XX}(k)$ est connue. On suppose que tous les échantillons dont les indices sont compris entre n_1 et n_2 ont été perdus ou corrompus. Le but de cet exercice va être de les reconstruire à partir des échantillons restants et de la fonction d'autocorrélation.

$$\underbrace{X(1), \ldots, X(n_1-1)}_{\text{presents}}\underbrace{X(n_1), \ldots, X(n_2)}_{\text{absents}}\underbrace{X(n_2+1), \ldots, X(N)}_{\text{presents}}$$

On notera $\boldsymbol{\theta}$ l'ensemble des échantillons manquants

$$\boldsymbol{\theta} = [X(n_1), \ldots, X(n_2)]^T$$

et $\mathbf{Y}$ l'ensemble des observations (échantillons présents)

$$\mathbf{Y} = [X(1), \ldots, X(n_1-1), X(n_2+1), \ldots, X(N)]^T$$

On notera respectivement N_θ et N_Y le nombre d'éléments de $\boldsymbol{\theta}$ et $\mathbf{Y}$ (On a $N_\theta = n_2 - n_1 + 1$ et $N_Y = N - n_2 + n_1 - 1$). On notera respectivement $\mathbf{i}_\theta$ et $\mathbf{i}_Y$ les indices des éléments de $X(n)$ stockés dans les vecteurs $\boldsymbol{\theta}$ et $\mathbf{Y}$.

$$\mathbf{i}_\theta = [n_1, \ldots, n_2] \quad \theta_j = X(\mathbf{i}_\theta(j))$$
$$\mathbf{i}_Y = [1, \ldots, n_1-1, n_2+1, \ldots, N] \quad Y_j = X(\mathbf{i}_Y(j))$$

Notre but est d'estimer $\boldsymbol{\theta}$ à partir de $\mathbf{Y}$. Pour cela on va considérer un estimateur linéaire minimisant l'erreur moyenne quadratique sous la forme

$$\hat{\boldsymbol{\theta}} = \mathbf{HY}$$

où $\mathbf{H}$ est une matrice de taille $N_\theta \times N_Y$ à déterminer.

1. En appliquant le principe d'orthogonalité, montrer que

$$\mathbf{H} = \mathbf{R}_{\theta\mathbf{Y}}\mathbf{R}_{\mathbf{YY}}{}^{-1}$$

   où $\mathbf{R}_{\mathbf{YY}} = \{E\{Y_iY_j\}\}_{\substack{i=1,\ldots,N_Y\\ j=1,\ldots,N_Y}}$ et $\mathbf{R}_{\theta\mathbf{Y}} = \{E\{\theta_iY_j\}\}_{\substack{i=1,\ldots,N_\theta\\ j=1,\ldots,N_Y}}$

2. En utilisant l'hypothèse de stationnarité au sens large, montrer que

$$E\{Y_iY_j\} = R_{XX}(|\mathbf{i}_Y(i) - \mathbf{i}_Y(j)|)$$

$$E\{\theta_iY_j\} = R_{XX}(|\mathbf{i}_\theta(i) - \mathbf{i}_Y(j)|)$$

1 Exercice 1

Les notations utilisées dans cet exercice sont les mêmes que celles de la préparation à faire à la maison. Le but est de reconstruire des échantillons audio perdus à partir des échantillons présents et de la fonction d'autocorrélation.

1. Charger le fichier audio `murder.wav` ainsi que la fonction d'autocorrélation `autocor.mat` (le premier élement $R_{XX}(0)$ correspond à l'indice 1).
2. Ecouter et visualiser le signal et déterminer n_1, n_2, N, N_θ et N_Y.
3. Calculer les matrices $\mathbf{R_{YY}}$ et $\mathbf{R_{\theta Y}}$ et en déduire une reconstruction $\hat{\boldsymbol{\theta}}$ des échantillons manquants.
4. Ecouter le résultat et commenter.

2 Exercice 2

On considère N observations $X_1, \dots, X_N$ supposées i.i.d et suivant chacune une loi exponentielle de paramètre θ. On définit deux estimateurs $\hat{\theta}_N^{(1)}$ et $\hat{\theta}_N^{(2)}$ du paramètre θ, définis par:

$$\hat{\theta}_N^{(1)} = \frac{1}{N}\sum_{k=1}^{N} X_N \qquad \hat{\theta}_N^{(2)} = \left(\frac{1}{2N}\sum_{k=1}^{N} X_k^2\right)^{\frac{1}{2}}$$

1. On suppose $\theta = 50$. Générer $M = 1$ réalisations pour $N = 1000$ observations grâce à la fonction `rand_exp` fournie et comparer les valeurs des deux estimateurs.
2. Tracer le biais, la variance et l'erreur quadratique moyenne de chaque estimateur en fonction de N pour $N = 1, \dots, 100$. (Les espérances seront calculées sur $M = 1000$ réalisations).
3. Comparer les résultats pour différentes valeurs de θ.