École Polytechnique de Montréal
Institut de génie biomédical

Cours GBM6103A
Traitement d'images et de signaux biomédicaux
Automne 2014

Travail pratique n° 3

Variables et signaux aléatoires

Instructions

- Les travaux pratiques sont effectués par équipes de deux.
- Le travail à remettre se compose :
  - d'un compte rendu qui doit comporter une réponse concise mais complète à chacune des questions, accompagnée au besoin des courbes, figures et images appropriées ;
  - des scripts et fonctions `matlab` correspondants, sous forme de fichiers texte (format `.m`).

  L'ensemble peut être composé soit à l'aide des fonctionnalités de publication de `matlab` (menu "File / Publish" de l'éditeur `matlab`), soit en utilisant un logiciel de traitement de texte indépendant. Dans tous les cas, seuls les formats html, pdf et texte (pour les fichiers `matlab`) sont acceptés. L'ensemble des fichiers doit être placé dans une unique archive `zip` et remis en utilisant l'outil approprié disponible sur le site web du cours.
- Le travail doit être remis par un seul des membres du groupe. Si tel n'est pas le cas, la version la plus récente du travail remis sera prise en compte.

1 Introduction

L'objet de cette séance est de se familiariser avec la manipulation de variables et de vecteurs aléatoires dont la principale utilité, dans ce cours, est de modéliser dans un cadre mathématique adéquat l'incertitude que l'on peut avoir sur une mesure expérimentale. On manipulera également quelques signaux aléatoires afin d'en mettre en évidence certaines caractéristiques importantes.

2 Loi de probabilité, histogramme (7 points)

Une expérience consiste à lancer deux dés à six faces non pipés, donnant les valeurs X_1 et X_2, et à noter leur somme $Z = X_1 + X_2$. On considère que X_1 et X_2 sont indépendantes.

Quelle est la loi de probabilité $\{p_X(i)\}_{i=1\ldots6}$ de X_1 ? Calculez la moyenne $E[X_1]$ et la variance $\mathrm{var}[X_1]$ de X_1. Montrez que l'on peut calculer la moyenne $E[Z]$ et la variance $\mathrm{var}[Z]$ de Z sans faire le calcul explicite de sa loi de probabilité p_Z. Évaluez $E[Z]$ et $\mathrm{var}[Z]$.

En vous reportant au besoin aux résultats fournis en annexe, faites le calcul explicite de la loi de probabilité $\{p_Z(j)\}_{j=2\ldots12}$ de Z ?

Nous allons maintenant vérifier que la loi de probabilité d'une variable aléatoire correspond bien à la fréquence empirique d'apparition d'une valeur lorsque l'on observe un échantillon de grande taille. Proposez une méthode `Matlab` (à l'aide de la fonction `rand`) qui génère une réalisation de X_1 et X_2, puis de Z. Générez un échantillon de N réalisations indépendantes de Z, $\{Z_1, Z_2, \ldots, Z_N\}$ pour $N = 10, 1\ 000, 1\ 000\ 000$. Calculez l'histogramme[1] $\{\hat{p}_j\}$ des valeurs des Z_i. Comparer $p_Z(j)$ et $\{\hat{p}_j\}$. Que se passe-t-il lorsque N croît ?

Compte-rendu

1. Loi de probabilités de X_1.
2. $E[X_1], \mathrm{var}[X_1]$.
3. $E[Z], \mathrm{var}[Z]$ (en utilisant les propriétés des moments).
4. Loi de probabilités de Z.
5. Histogramme (figure).
6. Comparaison de $p_Z(j)$ et de $\hat{p}_j$ avec discussion sur ce qui se passe quand N croît (un quart de page maximum).

3 Simulation de variables gaussiennes par la méthode de Box-Muller (5 points)

Considérons deux variables aléatoires X_1 et X_2, supposées indépendantes et de même loi uniforme sur $[0, 1]$. À partir de ces deux variables, on construit deux nouvelles variables Y_1 et Y_2 définies par :

$$Y_1 = \sqrt{-2 \ln X_1} \cos(2\pi X_2) \tag{1}$$
$$Y_2 = \sqrt{-2 \ln X_1} \sin(2\pi X_2) \tag{2}$$

Générez un échantillon de N valeurs de Y_1, $\{Y_1^{(1)}, \ldots, Y_1^{(N)}\}$. Calculez son histogramme pour N grand. Que remarquez-vous ?

Confirmez vos observations en calculant analytiquement la loi de Y_1 et celle de Y_2, en vous reportant au besoin aux résultats donnés en annexe.

Compte-rendu

1. Histogramme de Y_1, pour un grand nombre N d'échantillons.
2. Remarque sur l'histogramme (un quart de page maximum).
3. Détermination de la loi de probabilités de Y_1 et Y_2.

1. L'histogramme sera défini par $\hat{p}_j = \frac{1}{N} \times$ nombre d'échantillons Z_i ayant la valeur j

4 Processus du second ordre (8 points)

Cet exercice permettra d'illustrer les notions de loi instantanée, d'autocorrélation et de densité spectrale de puissance d'un processus du second ordre.

On se place dans la situation où l'on désire analyser la nature des incertitudes introduites par un système de mesure comprenant un étage d'amplification et un étage de filtrage. Plus précisément, on tente de déterminer si la principale source d'incertitudes se situe en amont ou en aval de l'étage de filtrage. Pour ce faire, on procède à une simulation numérique de ce système, en supposant d'abord que la source des incertitudes se situe en aval du filtre (hypothèse 1), puis en amont du filtre (hypothèse 2). On suppose que ces incertitudes, de source électronique, peuvent être modélisées par un bruit blanc gaussien centré. On suppose également que le filtre peut être représenté par la fonction de transfert :

$$H(z) = \frac{\sqrt{3}}{2}\,\frac{z}{z-1/2}\ ;\ |z| > 1/2 \quad . \tag{3}$$

En supposant qu'aucun signal « utile » n'est présent et en utilisant les résultats de la section précédente, générez $N = 1\,000$ échantillons de la sortie $\mathbf{Z} = [Z_1, ..., Z_N]^T$ du système sous l'hypothèse 1, puis N échantillons de la sortie $\mathbf{Y}$ du système sous l'hypothèse 2.

4.1 Propriétés du premier ordre (1 point)

Tracez l'histogramme de $\mathbf{Z}$, puis celui de $\mathbf{Y}$. Comparez ces deux histogrammes. Expliquez et commentez les résultats obtenus.

4.2 Propriétés du second ordre (7 points)

Quels sont les spectres théoriques de Z et de Y ?

Comme on le verra dans une prochaine leçon, on peut évaluer la fonction d'autocorrélation $R_{XX}(k)$ d'un signal discret aléatoire stationnaire X par la quantité :

$$\hat{R}_{XX}(k) = \frac{1}{N}\sum_{i=1}^{N-k} X_i X_{i+k}. \tag{4}$$

Programmez la fonction `autocorrel(X)` qui, à partir d'un vecteur $\mathbf{X}$ de taille N, retourne le vecteur $(\hat{R}_{XX}(0), \ldots, \hat{R}_{XX}(N-1))$.

Évaluez les fonctions d'autocorrélation $\hat{R}_{ZZ}(k)$ et $\hat{R}_{YY}(k)$. Déduisez-en les densités spectrales de puissance respectives de $\mathbf{Z}$ et $\mathbf{Y}$. Ces quantités permettent-elles de faire la distinction entre les hypothèses 1 et 2 ? Comparez les résultats obtenus aux densités spectrales de puissances théoriques de $\mathbf{Z}$ et $\mathbf{Y}$. Que concluez-vous ?

Chargez le signal `sortie_filtre` et analysez-le selon la procédure utilisée pour $\mathbf{Z}$ et $\mathbf{Y}$. Qu'en concluez-vous sur la position probable de la source des incertitudes ?

Compte-rendu

1. Histogramme de $\mathbf{Z}$ et de $\mathbf{Y}$.
2. Comparaison de ces histogrammes et commentaire (un tiers de page maximum).
3. Densité spectrale de puissance théorique de Z et de Y.

4. Représentation graphique des densités spectrales de puissance obtenues par calcul dans Matlab pour Z et pour Y.
5. Comparaison et conclusion sur la possibilité de distinguer entre les hypothèses 1 et 2 (un tiers de page maximum).
6. Analyse du signal `sortie_filtre` selon la procédure employée pour Z et Y ; conclusion sur la position probable des incertitudes (un tiers de page maximum).

Annexes

A Fonction de variable aléatoire

On considère une variable ou un vecteur aléatoire X, de loi $f_X(x)$. On définit une nouvelle variable (ou vecteur) aléatoire Y par :

$$Y = \phi(X)\,, \tag{5}$$

où ϕ est une fonction de dimension appropriée. Le problème posé est celui de la détermination de la loi de probabilité de Y $f_Y(y)$. L'expression de cette loi peut être établie relativement simplement si on suppose de plus que ϕ définit un changement de variable. En effet, dans ce cas, on a pour tout ensemble $\mathcal{X}$:

$$P(X \in \mathcal{X}) = P(Y \in \phi(\mathcal{X}))\,,$$

ce qui se traduit par :

$$\int_{\mathcal{X}} f_X(x)dx = \int_{\phi(\mathcal{X})} f_Y(y)dy\,. \tag{6}$$

Par ailleurs, pour toute fonction ψ la formule de changement de variable donne :

$$\int_{\phi(\mathcal{X})} \psi(y)dy = \int_{\mathcal{X}} \psi(\phi(x))|J(x)|dx\,,$$

où $J(x)$ désigne le jacobien du changement de variable, c'est-à-dire la matrice de terme général $\partial\phi_i/\partial x_j$. En prenant $\psi = f_Y$ et en reportant l'expression résultante dans (6), on obtient :

$$\int_{\mathcal{X}} f_Y(\phi(x))|J(x)|dx = \int_{\mathcal{X}} f_X(x)dx\,,$$

et cette égalité étant vraie pour tout $\mathcal{X}$, on aboutit finalement à :

$$f_Y(\phi(x))|J(x)| = f_X(x)\,.$$

Pour obtenir l'expression de $f_Y(y)$, il suffit de remplacer x par $\phi^{-1}(y)$ ce qui est toujours possible sous nos hypothèses, mais peut poser des problèmes calculatoires dans la pratique.

B Application : somme de deux variables aléatoires indépendantes

On considère deux variables aléatoires indépendantes X_1 et X_2 de loi respective $f_{X_1}(x_1)$ et $f_{X_2}(x_2)$. On désire obtenir la loi de la somme de X_1 et X_2. Pour cela, on introduit le changement de variable $\phi \ : \ (X_1, X_2) \longrightarrow (Y_1, Y_2)$ défini par :

$$\begin{aligned} Y_1 &= X_1 + X_2 \\ Y_2 &= X_2 \end{aligned}$$

De manière immédiate, le déterminant du jacobien de la transformation est constant et égal à un. D'après le résultat précédent et en raison de l'indépendance de X_1 et X_2, on a :

$$f_Y(y) = f_{X_1}(x_1) f_{X_2}(x_2)\,.$$

En exprimant x_1 et x_2 en fonction de y_1 et y_2, on obtient :

$$f_Y(y) = f_{X_1}(y_1 - y_2) f_{X_2}(y_2)$$

La loi de la somme de x_1 et x_2 s'obtient en marginalisant par rapport à y_2, c'est à dire en intégrant $f_Y(y)$ par rapport à y_2. On a donc :

$$f_{Y_1}(y_1) = \int f_{X_1}(y_1 - u) f_{X_2}(u) du\,,$$

d'où le résultat classique selon lequel la loi de la somme de deux variables aléatoires indépendantes est égale au produit de convolution de chacune des lois.