Travaux de Laboratoire
Année 2008-2009

Ecole Centrale Marseille

# Transformée de Fourier
# Application en statistique non-paramétrique

**Partie 1 : TFD, FFT et phénomène de Gibbs** L'abréviation FFT est utilisé pour "Fast Fourier Transform", en français "Transformée de Fourier Rapide". L'algorithme FFT permet de calculer la transformée de Fourier discrète d'un signal (TFD) en un temps beaucoup plus court que par un calcul direct. Cette première partie propose d'examiner les performances de l'algorithme FFT implémenté sous *Scilab* et les phénomènes qui lui sont liés.

Soit $f$ une fonction périodique de période 1. On définit la transformée de Fourier de $f$ par

$$\hat{f}(\omega) = \int_0^1 f(t)e^{-2i\pi\omega x}dx.$$

La transformée de Fourier discrète (TFD) associe au vecteur $(f(\frac{k}{N}))_{0\leq k\leq N-1}$ le vecteur $(\tilde{f}_k)_{0\leq k\leq N-1}$ défini par

$$\tilde{f}_k = \sum_{j=0}^{N-1} f(\frac{j}{N})e^{-\frac{2i\pi kj}{N}}.$$

Deux programmes permettent de calculer la TFD sous *Scilab* : *dft* dans le cas général, et *fft* lorsque $N$ est une puissance de 2.

1. Comparaison de *fft* et *dft*
   (a) Créer une fonction qui, à partir d'une fonction $f$ et d'un entier $N$, retourne le vecteur d'échantillonage de taille $N$ de $f$ sur [0,1].
   (b) Utiliser *dft* et *fft* pour calculer la TFD des fonctions $\sin(2\pi x)$ et $\cos(2\pi x)$. Faites le lien avec les coefficients de Fourier.
   (c) En utilisant un échantillonnage de la fonction "chapeau"

$$f(x) = \begin{cases} 2x & si \;\; x \in [0,1/2] \\ 2-2x & si \;\; x \in ]1/2,1] \end{cases}$$

   sur $N$ points, vérifier que le nombre d'opérations de la *dft* est proportionnel à $N^2$ alors que celui de la *fft* est de l'ordre de $N\log(N)$ (on utilisera la fonction *timer()*).
2. La fonction *fft*
   Dans toute la suite, on utilise la fonction *fft* pour calculer la TFD. On prendra donc un échantillon dont la longueur est une puissance de 2.
   (a) Soit $f$ définie sur [0,1] qui vaut 1 sur l'intervalle [1/4,3/4] et 0 sinon. Montrer que $f \in L^1$ et calculer sa transformée de Fourier.
   (b) Créer un vecteur $u$ de taille $2^8$ qui échantillone $f$ sur [0,1]. Visualisez $v = abs(fft(u))$ et $v1 = fftshift(v)$. Que remarquez-vous ?
   (c) Répétez le procédé en faisant varier la taille de l'échantillon.
3. Phénomène de Gibbs
   (a) Créer un vecteur $u$ de taille $2^{10}$ qui échantillone $f$ sur [0,1]. Visualisez $abs(fftshift(fft(u)))$. Si on considère qu'il représente les valeurs de $\hat{f}$ sur un intervalle $[-\frac{T}{2}, \frac{T}{2}]$, où se trouvent les valeurs correspondant à l'intervalle $[-\frac{T}{2}, -\frac{T}{4}]$ et $[\frac{T}{4}, \frac{T}{2}]$ ?
   (b) Même question pour $abs(fft(u))$.

(c) Mettre à zéro les points de cet intervalle, soit directement dans $v = fft(u)$, soit dans $v1 = fftshift(fft(u))$.

(d) Effectuer l'opération d'inversion de Fourier en appliquant la fonction $ifft$. Visualiser $abs(ifft(v1)$.

(e) Que se passe-t'il lorsqu'on augmente la taille $N$ de l'échantillon ?

(f) Etudier le phénomène pour les fonctions $g_1 = 2f$, $g_2 = 5f$ et $g_3 = 10f$. Que remarquez-vous ?

*Le phénomène que vous observez est appelé phénomène de Gibbs. Il résulte de la troncature brutale des hautes fréquences. Il apparaît autour des points de discontinuités et illustre la convergence simple de la série de Fourier vers la fonction $f$, convergence qui n'est pas uniforme.*

**Partie 2 : Regression non-paramétrique** Le but de cette deuxième partie est d'estimer une fonction $f \in L^2[0,1]$ à partir d'observations bruitées de $f$ sur l'intervalle [0,1]. On considère le modèle de régression suivant :

$$y_k = f(\frac{k}{n}) + \sigma\epsilon_k, k = 1, \dots n, \tag{1}$$

où la suite des $\epsilon_k$ est indépendante identiquement distribuée (i.i.d.) de loi $\mathcal{N}(0,1)$. Les observations $y_k$ représentent les valeurs de $f$ sur l'intervalle [0,1] perturbées par un bruit gaussien. Dans ce qui suit, on se propose d'estimer la $f(x) = \sin(2\pi x + 1)^2$.

1. Ecrire une fonction $f$, ayant $\sigma$ et $n$ en argument et qui génère une suite d'observations $y_k, k = 1 \dots n$ vérifiant (1).

2. Représenter dans le cas $n = 2^8$ et $\sigma = 5$ :
   - la fonction $f$ et la version bruitée, notée $f_\epsilon$ ;
   - $\hat{f}$ la TFD de $f$ et $\hat{f}_\epsilon$ la TFD de $f_\epsilon$.

   Que remarquez-vous ?
3. Pour estimer $f$ à partir de la suite $(y_k)$, on propose de tronquer la TFD du signal bruité en ne gardant que $M$ coefficients correspondant aux basses fréquences. On note $\hat{f}_{\epsilon,M}$ la TFD tronquée, et $T_{n,M}$ l'estimateur de $f$ obtenu en appliquant la TFD inverse à $\hat{f}_{\epsilon,M}$.
   Ecrire une fonction qui, ayant $M$ comme argument, retourne l'estimateur $T_{n,M}$.
4. Dans le cas $n = 2^8$ et $\sigma = 5$, étudier l'effet du nombre $M$ de coefficients conservés sur la forme de l'estimateur.
5. On veut évaluer la qualité de cette procédure d'estimation. Ecrire une fonction qui évalue $\|T_{n,M} - f\|^2_{L^2}$, l'erreur $L^2$ de reconstruction.
6. En déduire un choix optimal de $M$ dans le cas $n = 2^8$ et $\sigma = 5$ pour estimer $f$.
7. Etudier l'effet de $\sigma$ sur le choix optimal de $N$.