Simulation stochastique et méthodes bayésiennes pour le traitement du signal - 2018

TP1 : Estimation bayésienne
Yann Traonmilin

On donne un fichier regroupant toutes les fonctions R utiles à ce TP et quelques conseils, utiliser la fonction `help(fun)` pour en déterminer l'utilisation.

Télécharger dans le dossier courant le fichier source :
https://drive.google.com/open?id=18Pm20C6fW9dpZQF3KBG1XjxR7EUumm2t

1 Un exemple simple d'estimation

On pourra s'inspirer du code des TP précédents pour cette partie. On se fixe σ. On s'attachera à observer et analyser les résultats pour différentes valeurs des paramètres. On considère le modèle bayésien suivant : $(X_i)_{i=1,n}$ iid à valeur dans $\mathbb{R}$ avec $X_i \sim \mathcal{N}(\theta, \sigma^2)$ et $\theta \sim \mathcal{N}(\mu, \tau^2)$.

1. Générer un jeu de données suivant une loi $\mathcal{N}(\theta_0, \sigma^2)$.
2. Rappeler (cours) l'expression de la loi a posteriori sur $\pi(\theta|x)$. Représenter cette loi pour différentes valeur de n.
3. On considère la fonction de coût $C(u,v) = |u-v|^2$. Rappeler l'expression de l'estimateur de Bayes pour cette fonction de coût.
4. Représenter sur un même graphique l'estimateur de Bayes et l'estimateur du maximum de vraisemblance en fonction du nombre d'observations.
5. Représenter sur un même graphique l'erreur quadratique de l'estimateur de Bayes et de l'estimateur du maximum de vraisemblance en fonction du nombre d'observations. Faire varier les paramètres du modèle bayésien et commenter.
6. Pour 1000 expériences donner, sur un exemple, un intervalle de confiance de l'estimateur de Bayes à 95 % .

2 Étude d'un modèle polynomial

Dans cette section, on cherche à déterminer une relation polynomiale entre deux suites d'échantillons $y = (y_1, ...y_n)$ et $z = (z_1, ..., z_n)$. On utilise le modèle bayésien linéaire suivant

$$y_i = \theta_0 + \theta_1 z_i + \theta_2 z_i^2 + \epsilon_i,$$

où les ϵ_i sont des variables aléatoires iid de loi $\mathcal{N}(0, \sigma^2)$ de variance connue, et les paramètres $\theta = (\theta_0, \theta_1, \theta_2)$ suivent la loi *a priori* suivante :

$$\pi(\theta) \sim \mathcal{N}(\mu, \tau^2 I) \tag{1}$$

1. Générer un jeux de données en utilisant un modèle aléatoire pour z et en choisissant une valeur au choix θ_v et σ^2 pour les paramètres.

2. Rappeler (cours) l'estimateur du maximum de vraisemblance de θ_v. Rappeler l'estimateur de Bayes de θ_v (confondu ici avec le maximum a posteriori) pour le coût quadratique.

3. Représenter sur un même graphique l'erreur quadratique de l'estimateur de Bayes et de l'estimateur du maximum de vraisemblance en fonction du nombre d'observations. On pourra se limiter aux cas où la matrice du modèle linéaire est sur-déterminée.

4. Même question que précédemment en moyennant sur plusieurs réalisations du bruit. Faire varier les paramètres du modèle bayésien et commenter.