MACS 205 Projet 4 : Metropolis Hastings adaptatif

Rafael Brutti, Priam Cardouat, Caroline Xia

9 avril 2021

Table des matières

1 Metropolis Hastings classique

1.1 Description

Nous allons rapidement rappeler le principe de la méthode de Hasting Metropolis. Il s'agit d'une méthode de Monte Carlo par chaînes de Markov qui permet d'obtenir un échantillon à partir d'une densité f lorsque l'échantillonnage direct est compliqué.

On a une densité cible $f : R^d \to R_+$ et un noyau de Markov $q : R^d \to R^d$

Etape 1 : Tirage de Y : $Y \sim q(., X_{i-1})$
Etape 2 : 2 possibilités : Acceptation de Y et $X_i = Y$ ou rejet et $X_i = X_{i-1}$

L'acceptation pour le point X_i suit une loi de Bernoulli de paramètre :

$$p_i = min(1, \frac{f(Y)q(X_{i-1}, Y)}{f(X_{i-1})q(Y, X_{i-1})})$$

et $p_i = 1$ si $f(X_{i-1})q(Y, X_{i-1}) = 0$

1.2 Intérêt

1.3 Quelques exemples

1.4 Problèmes

1. Metropolis Hastings classique ne visite pas tout l'espace (exemple de la double gaussienne sur la fig 1.1.2)

2. Une possibilité pour le noyau est de prendre par exemple $X_i \sim \mathcal{N}(X_{i-1}, 1)$. L'algorithme ne s'adapte alors pas à la forme de la densité cible car la variance est constante (cf Haario et al.)

2 Metropolis Hastings adaptatif (Haario et al.)

2.1 Description

L'idée de l'algorithme de Metropolis adaptatif est de mettre à jour la distribution pour choisir le point suivant en utilisant toutes les connaissances ultérieures sur la densité cible. C'est-à-dire que la variance est calculée à chaque étape selon les tirages précédents. On ne travaille plus avec une chaîne de Markov. Sinon, la définition de l'algorithme est identique au processus habituel de Metropolis.

On a une densité cible $f : R^d \rightarrow R_+$ et à l'étape i, on a déjà tiré $X_0, X_1, ..., X_{i-1}$:

Etape 1 : Tirage de Y : $Y \sim q_i(.|X_0, X_1, ..., X_{i-1})$
Etape 2 : 2 possibilités : Acceptation de Y et $X_i = Y$ ou rejet et $X_i = X_{i-1}$

L'acceptation pour le point X_i suit une loi de Bernoulli de paramètre :

$$p_i = min(1, \frac{f(Y)}{f(X_{i-1})})$$

La densité $q_i(.|X_0, X_1, ..., X_{i-1})$ suit une distribution normale d'espérance X_{i-1} et de covariance C_i qui dépend de $X_0, X_1, ..., X_{i-1}$. Dans l'algorithme étudié :

$$C_i = \{ C_0 \, si \ t \leq t_0 s_d(Cov(X_0, X_1, ..., X_{i-1}) + \epsilon I_d) sinon.$$

2.2 Avantages de la méthode

2.3 Quelques exemples

3 Metropolis Hastings avec la méthode du noyau

3.1 Description

Par rapport aux travaux de Haario et al. nous avons cherché à étudier un nouveau type de méthode de Métropolis Hastings adaptatif avec une politique différente. Celle-ci est maintenant donnée par :

$$q_n(x) = \frac{1}{n} \sum_{i=1}^{n} K_h(x - X_i)$$

avec $K_h(u) = h^d K(\frac{u}{h}$ où $KR^d \rightarrow R_+$ une densité. L'algorithme reste fondamentalement le même par ailleurs.

3.2 Avantages de la méthode

3.3 Quelques exemples

4 Tests et comparaison

Nous avons chercher à comparer les différentes méthodes dans différents cas et à comparer leurs performances. Pour des raisons de puissance de calcul de nos ordinateurs à disposition, nous nous sommes limités à des densités en dimension 2 (même si les algorithmes implémentés fonctionnent pour des dimensions supérieures) et pour les valeurs de n suivantes : $n = \{500, 1000, 1500, 2000, 2500, 3000\}$. Pour réaliser les tests, nous avons choisi une densité cible gaussienne. L'erreur moyenne quadratique est donnée par :

$$MSE = ||\mu - \hat{\mu}_n||_2$$

avec μ la moyenne de la densité cible et $\hat{\mu}_n$ la moyenne estimée à partir des particules issues de la méthode de Métropolis Hastings.

4.1 Test avec une gaussienne et "bon" départ

Tout d'abord nous avons donc calculé l'erreur moyenne quadratique pour les différentes méthodes ci-dessus pour une densité cible qui est une gaussienne centrée réduite et pour une densité initiale q_0 qui est une gaussienne centrée réduite.

4.2 Test avec une gaussienne et un départ "à froid"

Dans ce test, nous avons cette fois considéré gaussienne de moyenne $\mu = [4, 4]$ de matrice de covariance I_2 comme densité cible test. La densité initiale est toujours une gaussienne centrée réduite.

4.3 Test avec une double gaussienne

La où le test nous a semblé le plus intéressant fût pour une densité cible qui était une double gaussienne. En effet, l'algorithme de Métroplis Hastings a du mal à explorer une densité "discontinue" en raison de la variance qui est fixée. Ce problème est quelque peu corrigé par les algorithmes de Métropolis Hastings adaptatifs comme celui proposé par Haario et al. En revanche sur le figure (? ? ?) nous pouvons voir que cet algorithme peut encontrer des difficultés. Enfin l'algorithme adaptatif que nous avons proposé s'adapte mieux pour de tels densités comme cela peut se voir sur la figure (? ? ? ? ?)

4.4 Comparaison

Il apparaît donc que les méthodes classique et adaptatives (comme celle proposée par Haario et al.) suffisent pour des densités cibles "simples", c'est-a-dire avec une variance relativement faible. En revanche, si la densité cible est plus "exotique", la méthode adaptative que nous proposons s'adapte mieux à ce genre de situation.