Chapitre 4 : Méthode de Monté Carlo - "Monte Carlo Method"

Dr. Nawel Arrar
3 ème année ENSIA

Décembre 2023

Description de la méthode

On cherche à évaluer une intégrale dans $\mathbb{R}^d$, plus générale, du type

$$I = \int_{\mathbb{R}^d} g(x) f(x) \, dx$$

Alors I s'écrit sous la forme $\mathbb{E}[g(X)]$ si X est une v.a. à valeurs dans $\mathbb{R}^d$ de loi $f(x)\,dx$. Soit une suite $(X_i, i \geq 1)$ de v.a. i.i.d selon une loi de X. L'estimation de I est donnée par

$$\widehat{I} \approx \frac{1}{n} \sum_{i=1}^{n} g(X_i).$$

Convergence et limite de la méthode

La loi forte des grands nombres qui permet de justifier la convergence de la méthode et le Théorème Central Limit qui précise la vitesse de convergence.

Convergence et limite de la méthode

Loi forte des grands nombres et méthode de Monté Carlo

Theorem

Soit $(X_n, n \geq 1)$ *une suite de v.a. indépendantes suivant toutes la même loi qu'une v.a.* X. *On suppose que* $\mathbb{E}\left[|X|\right] < +\infty$, *alors pour presque tout* $\omega \in \Omega$

$$\mathbb{E}\left[X\right] = \lim_{n \to +\infty} \frac{1}{n} \left(X_1 + X_2 + \ldots + X_n\right)(\omega).$$

Convergence et limite de la méthode

Théorème central limit et MMC

Theorem

Soit $(X_n, n \geq 1)$ *une suite de v.a. i.i.d. de carré intégrable* $(\mathbb{E}\left[X^2\right] < \infty)$.*On note* $m = \mathbb{E}\left[X\right]$ *et* $\sigma^2 = Var\left(X\right) : \sigma^2 = \mathbb{E}\left[X^2\right] - \mathbb{E}\left[X\right]^2$.
Alors la suite $\frac{\sqrt{n}}{\sigma}\left(m - \overline{X}_n\right) = \frac{\sqrt{n}}{\sigma}\epsilon_n$ *converge en loi vers* Z *une v.a. de loi gaussienne centrée réduite.*

Convergence et limite de la méthode

Théorème central limit et MMC

Une table de la fonction de répartition d'une loi gaussienne centrée réduite montre que si Z est $\mathcal{N}(0,1)$, $\mathbb{P}(Z \leq 1,96) = 0,95$. On en déduit que pour n assez grand,

$$\mathbb{P}\left(|\epsilon_n| \leq 1,96\frac{\sigma}{\sqrt{n}}\right) \backsim 0,95,$$

c'est à dire que l'on a un intervalle de confiance de $\mathbb{E}[X]$ à 95% en posant

$$\left[\overline{X}_n - 1,96\frac{\sigma}{\sqrt{n}}, \overline{X}_n + 1,96\frac{\sigma}{\sqrt{n}}\right].$$

Méthodes de réduction de la variance "Variance reduction"

Introduction

Nous avons vu que la vitesse de convergence de la méthode Monté Carlo est de l'ordre de $\frac{1}{\sqrt{n}}$. Pour améliorer cette méthode il existe de nombreuses techniques dites de réduction de variance, qui cherchent à diminuer la valeur de σ^2.
L'idée générale est de donner une autre représentation sous forme d'espérance de la quantité à calculer

$$\mathbb{E}\left[X\right] = \mathbb{E}\left[Y\right], \text{ telle que } V\left(Y\right) < V\left(X\right).$$

Echantillonnage préférentiel "Importance Sampling"

Supposons que l'on cherhce à calculer $\mathbb{E}\left[g\left(X\right)\right]$ avec X v.a. à valeurs dans $\mathbb{R}^d$ de densité de probabilité f sur $\mathbb{R}$, alors

$$\mathbb{E}\left[g\left(X\right)\right]=\int_{\mathbb{R}} g\left(x\right) f\left(x\right) dx.$$

Soit $\widetilde{f}$ la densité d'une autre variable Y telle que $\widetilde{f}>0$ et $\int_{\mathbb{R}} \widetilde{f}\left(x\right) dx=1,$ alors on peut écrire $\mathbb{E}\left[g\left(X\right)\right]$ comme suit :

$$\begin{aligned}\mathbb{E}\left[g\left(X\right)\right] &= \int_{\mathbb{R}^d} \frac{g\left(x\right) f\left(x\right)}{\widetilde{f}\left(x\right)} \widetilde{f}\left(x\right) dx \\ &= \mathbb{E}\left[\frac{g\left(Y\right) f\left(Y\right)}{\widetilde{f}\left(Y\right)}\right]\end{aligned}$$

où Y est une v.a. de loi de probabilité $\widetilde{f}$.

Echantillonnage préférentiel "Importance Sampling"

On a donc une autre méthode de calcul de $\mathbb{E}\left[g\left(X\right)\right]$, en utilisant n tirages de $\left(Y, Y_1, \cdots, Y_n\right)$ i.i.d. de densité de probabilité $\widetilde{f}$. On approxime $\mathbb{E}\left[g\left(X\right)\right]$, pour $\left(X_i\right)_{i=1}^{n}$ i.i.d. de même loi que X

$$\begin{aligned}\mathbb{E}\left[g\left(X\right)\right] &\approx \frac{1}{n}\sum_{i=1}^{n} g\left(x_i\right), \\ &\approx \frac{1}{n}\left(\frac{g\left(y_1\right)f\left(y_1\right)}{\widetilde{f}\left(y_1\right)} + \cdots + \frac{g\left(y_n\right)f\left(y_n\right)}{\widetilde{f}\left(y_n\right)}\right) \\ &= \frac{1}{n}\sum_{i=1}^{n}\frac{g\left(y_i\right)f\left(y_i\right)}{\widetilde{f}\left(y_i\right)}.\end{aligned}$$

Posons $Z = \frac{g(Y)f(Y)}{\widetilde{f}(Y)}$, si $Var\left(Z\right) << Var\left(g\left(X\right)\right)$ on aura amélioré l'algorithme.

Variable de contrôle "Control variables"

L'idée consiste à trouver une v.a. Y et une constante C telle que $\mathbb{E}[X] = \mathbb{E}[Y] + C$ avec $Var(Y) < Var(X)$.
On cherche à écrire $\mathbb{E}[f(X)]$ sous la forme

$$\mathbb{E}[f(X) - h(X)] + \mathbb{E}[h(X)]$$

avec $\mathbb{E}[h(X)]$ qui peut se calculer explicitement et

$$Var(f(X) - h(X)) < Var(f(X)).$$

On utilise alors une MMC pour le calcul de $\mathbb{E}[f(X) - h(X)]$

$$\mathbb{E}[f(X)] \approx \frac{1}{n}\sum_{i=1}^{n}(f(X_i) - h(X_i)) + \mathbb{E}[h(X)]$$

avec $(X_i)_{i=1}^{n}$ i.i.d de même loi que X.

Variables antithétiques "Antithetic variables"

Supposons que l'on cherche à calculer $I = \int_0^1 f(x)\,dx$, comme $x \longmapsto 1 - x$ laisse invariante la mesure dx sur $[0, 1]$, alors on peut écrire :

$$I = \int_0^1 \frac{[f(x) + f(1-x)]}{2} dx$$

On peut donc calculer I de la façon suivante : on tire n v.a $\sim \mathcal{U}$ $[0, 1]$ indépendantes $(U_1)_{i=1}^n$ et on approxime I par

$$\begin{aligned} I_{2n} &= \frac{1}{n}\left[\frac{1}{2}(f(u_1) + f(1-u_1)) + \cdots + \frac{1}{2}(f(u_n) + f(1-u_n))\right] \\ &= \frac{1}{2n}\sum_{i=1}^{n}[f(u_i) + f(1-u_i)]. \end{aligned}$$

La qualité de l'approximation s'améliore pourvu que $Var(I_{2n}) < Var(I)$ qui est vrai dès que $f(U)$ et $f(1-U)$ soient indépendantes.

Méthode de stratification "Stratified sampling"

Supposons que l'on cherche à calculer $I = \mathbb{E}\left[g\left(X\right)\right] = \int g\left(x\right) f\left(x\right) dx$ où X suit la loi de f définie sur un ensemble D. Ce qui nous donne l'idée d'une première méthode de M.C

$$I = \frac{g\left(X_1\right) + g\left(X_2\right) + ... + g\left(X_n\right)}{n}$$

avec $\left(X_i\right)_{i=1}^{n}$ i.i.d de même loi que $X \in D$. L'erreur commise

$$e_I = I - \frac{g\left(X_1\right) + g\left(X_2\right) + ... + g\left(X_n\right)}{n}.$$

Méthode de stratification "Stratified sampling"

Une nouvelle idée consiste à la décomposition de I. Supposons que D est fractionné en m sous ensembles $\left(D = \bigcup\limits_{i=1}^{m} D_i\right)$, alors on peut écrire

$$I = \sum_{i=1}^{m} p_i I_i$$

$$I_i = \mathbb{E}\left[g\left(X\right)/X \in D_i\right],$$

avec $p_i = \mathbb{P}\left(\left\{X \in D_i\right\}\right)$ et $\sum_{i=1}^{m} p_i = 1$. Donc $\forall i = \overline{1, m}$, on peut approximer I par la méthode de M.C à n tirage aléatoire $X_1^{(i)}, X_2^{(i)}, ..., X_n^{(i)}$ i.i.d de même loi que $X \in D_i$

$$I_i \approx \widehat{I_i} = \frac{g\left(X_1^{(i)}\right) + g\left(X_2^{(i)}\right) + ... + g\left(X_n^{(i)}\right)}{n_i}.$$