MAP 556 Challenge 1

Peng-Wei Chen

La méthode utilisée consiste à discrétiser la loi gaussienne en utilisant les cellules de Voronoï. Afin de reconnaître les cellules, j'ai procédé ainsi :

1. Initiation avec 100 points obtenus par k-means sur 10^8 échantillons.

2. Itérer l'algorithme ci-dessous afin de calculer les «centres» de chaque cellule.

3. Finalement, pour l'estimation de l'espérance, cela se calcule à partir de la valeur en ces points avec un poids de la probabilité d'avoir le point dans la cellule. La probabilité est déterminée à l'aide de la méthode de Monte-Carlo. J'ai ainsi réalisé 10^9 fois de simulation et calculé le nombre de points dans chaque cellule pour la probabilité.

Pour $i \in \{1, \ldots, 100\}$, on peut itérer l'équation suivante pour trouver les centres de cellules de Voronoï.

$$X_i(k+1) = \frac{\sum_{m=1}^{M} Y(m) \mathbb{1}_{Y(m) \in C_i(\Gamma(k))}}{\sum_{m=1}^{M} \mathbb{1}_{Y(m) \in C_i(\Gamma(k))}}$$

où

$X_i(k)$ est le i-ième centre dans les cellules de Voronoï lors de la k-ième itération
M est le nombre de simulation pour chaque itération
$Y(m)$ est la m-ième point obtenu par la loi gaussienne centrée réduite de dimension 4
$C_i(\Gamma(k))$ est la i-ième cellule dans la répartion calculée à partir de $X_i(k)$

On peut montrer que cette itération tend vers les vrais centres de Voronoï grâce à la loi des grands nombres. En pratique, j'ai pris $M = 10^7$ et le nombre d'itération totale de l'ordre de 50.

Avec cette méthode là, j'ai réussi à avoir une loi discrétisée avec une distortion de l'ordre de 1. La distortion est définie comme suite :

$$D(X) = \int_{\mathbb{R}^4} |X_{\text{le plus proche de x}} - x|^2 dx, X = (X_i)_{i \in [1,100]} \text{ est l'ensemble des 100 points.}$$

Ainsi, nous avons une nouvelle distribution Q à valeurs dans 100 points avec les poids différents. On peut donc considérer que

$$\mathbb{E}_G\left[f(X)\right] = \mathbb{E}_Q\left[f(X)\right]$$

où $G = \mathcal{N}(0, I_n)$. Avec cette équation, on peut utiliser la variable de contrôle avec les polynomes d'Hermite et les transformation de Laplace, ce qui est expliqué dans le polycopié. En pratique, on sait que l'espérance est biaisée avec cette méthode. Ainsi, on utilise la méthode de rejet pour prendre un point random dans la cellule et lui donne le même poids que le centre. Cela suit le même principe de décalage aléatoire en quasi-Monte Carlo.