Les suites à discrépance faible

6-601-09 Simulation Monte Carlo

Geneviève Gauthier

HEC Montréal

Introduction

- Nous présentons une introduction à la simulation de Quasi Monte Carlo.
- C'est plutôt un survol historique présentant les idées maîtresses qu'une présentation détaillée des derniers développements en la matière.

Référence

- Boyle, P., M Broadie and P Glasserman, (1997). «Monte Carlo Methods for Security Pricing», *Journal of Economic Dynamics and Control*, 21, 1267-1321.
- Glasserman, P (2004) *Monte Carlo Methods in Financial Engineering*.
- Niederreiter, H (1992) *Random Number Generation and Quasi-Monte Carlo Methods*, Siam, 237 pages.

La problématique I

- Soit X une variable aléatoire dont la *fonction de répartition* est F_X, c'est-à-dire que

$$\forall x \in \mathbb{R}, \; F_X(x) \equiv \Pr[X \leq x].$$

- Considérons la fonction $g : \mathbb{R} \to \mathbb{R}$. Nous voulons estimer la quantité

$$\theta = \mathrm{E}[g(X)].$$

- Notons que si X est de loi continue et qu'elle possède une *fonction de densité* f_X,

$$\forall x \in \mathbb{R}, \; f_X(x) = \frac{d}{dx} F_X(x),$$

alors

$$\mathrm{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) \; dx.$$

La problématique II

- Il est possible de ramener le domaine d'intégration à l'intervalle $[0,1]$ par un changement de variable: si $u = \exp(x)$ et $v = \exp(-x)$, alors

$$
\begin{aligned}
& \mathrm{E}[g(X)] \\
= & \int_{-\infty}^{0} g(x) f_X(x) \, dx + \int_{0}^{\infty} g(x) f_X(x) \, dx \\
= & \int_{0}^{1} g(\ln u) f_X(\ln u) \, \frac{du}{u} - \int_{1}^{0} g\left(\ln \frac{1}{v}\right) f_X\left(\ln \frac{1}{v}\right) \frac{dv}{v} \\
= & \int_{0}^{1} \left(\frac{g(\ln u) f_X(\ln u)}{u} + \frac{g\left(\ln \frac{1}{u}\right) f_X\left(\ln \frac{1}{u}\right)}{u} \right) du.
\end{aligned}
$$

- L'estimation de $\mathrm{E}[g(X)]$ peut se ramener à un problème d'intégration numérique sur l'intervalle $[0,1]$.

Version multidimensionnelle I

- Soit les variables aléatoires $X_1, ..., X_d$ et la fonction $g : \mathbb{R}^d \longrightarrow \mathbb{R}$. On aimerait estimer

$$\mathrm{E}\left[g\left(X_1, ..., X_d\right)\right].$$

Version multidimensionnelle II

- Dans le cas de variables aléatoires de lois continues,

$$
\begin{aligned}
& \mathrm{E}\left[g\left(X_1, \ldots, X_d\right)\right] \\
= & \int_{\mathbb{R}^d} g\left(x_1, \ldots, x_d\right) f_{X_1, \ldots, X_d}\left(x_1, \ldots, x_d\right) \; dx_1 \ldots dx_d \\
= & \int_{\mathbb{R}^d} g(\mathbf{x}) f_{X_1, \ldots, X_d}(\mathbf{x}) \; d\mathbf{x}
\end{aligned}
$$

où $f_{X_1, \ldots, X_d}\left(x_1, \ldots, x_d\right)$ représente la *fonction de densité jointe* des variables aléatoires $X_1, \ldots, X_d$:

$$
\begin{aligned}
f_{X_1, \ldots, X_d}\left(x_1, \ldots, x_d\right) & = \frac{\partial^d}{\partial x_1 \ldots \partial x_d} F_{X_1, \ldots, X_d}\left(x_1, \ldots, x_d\right) \\
& = \frac{\partial^d}{\partial x_1 \ldots \partial x_d} \operatorname{Pr}\left[X_1 \leq x_1, \ldots, X_d \leq x_d\right].
\end{aligned}
$$

Version multidimensionnelle III

- En utilisant un changement de variables, on peut trouver une fonction f telle que

$$\begin{aligned} \mathrm{E}\left[g\left(X_1, ..., X_d\right)\right] &= \int_{\mathbb{R}^d} g\left(\mathbf{x}\right) f_{X_1,...,X_d}\left(\mathbf{x}\right) \; d\mathbf{x} \\ &= \int_{\mathcal{I}^d} f\left(\mathbf{u}\right) \; d\mathbf{u} \end{aligned}$$

où $\mathcal{I}^d$ est le cube unitaire en dimension d

$$\mathcal{I}^d = [0,1]^{\times d}.$$

La discrépance d'un ensemble de points I

- Nous cherchons à estimer une intégrale

$$\int_{\mathcal{I}^d} f\left(\mathbf{u}\right)\ d\mathbf{u}$$

dont le domaine est le cube unitaire en dimension d

$$\mathcal{I}^d \equiv [0,1]^{\times d}$$

et la fonction à intégrer est $f : \mathcal{I}^d \to \mathbb{R}$.

- L'approximation utilisée est

$$\frac{1}{N}\sum_{n=1}^{N} f\left(\mathbf{u}_n\right)$$

où $\mathbf{u}_1, ..., \mathbf{u}_N$ sont des points choisis dans $\mathcal{I}^d$.

La discrépance d'un ensemble de points II

Question. Avec quel ensemble de points $\{\mathbf{u}_1, ..., \mathbf{u}_N\} \subset \mathcal{I}^d$ allons-nous approcher l'intégrale

$$\int_{\mathcal{I}^d} f(\mathbf{u}) \ d\mathbf{u} \approx \frac{1}{N} \sum_{n=1}^{N} f(\mathbf{u}_n) \ ?$$

Minimalement, il faut que

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} f(\mathbf{u}_n) = \int_{\mathcal{I}^d} f(\mathbf{u}) \, d\mathbf{u}$$

pour une classe raisonnable de fonctions $f : \mathcal{I}^d \to \mathbb{R}$.

- Disons qu'il faudrait que l'égalité soit satisfaite pour l'ensemble $\mathcal{C}(\mathcal{I}^d, \mathbb{R})$ de toutes les fonctions f qui sont continues sur le domaine $\mathcal{I}^d$.

La discrépance d'un ensemble de points III

- Cette dernière condition,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f\left(\mathbf{u}_n\right) = \int_{\mathcal{I}^d} f\left(\mathbf{u}\right) d\mathbf{u},\ \forall f \in \mathcal{C}\left(\mathcal{I}^d, \mathbb{R}\right),$$

  est satisfaite si la suite de points $\{\mathbf{u}_n : n \in \mathbb{N}\}$ est uniformément distribuée sur le cube unitaire $\mathcal{I}^d$, c'est-à-dire que

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbf{1}_{\mathbf{u}_n \in \mathcal{J}} = \lambda\left(\mathcal{J}\right)$$

  pour tous les sous-intervalles $\mathcal{J}$ du cube unitaire $\mathcal{I}^d$ où
  - λ est la mesure de Lebesgue (le volume) de $\mathcal{J}$,
  - $\lambda\left(\mathcal{J}\right)$ est le volume de l'ensemble $\mathcal{J}$
  - $N^{-1} \sum_{n=1}^{N} \mathbf{1}_{\mathbf{u}_n \in \mathcal{J}}$ est la proportion des points de $\{\mathbf{u}_1, \ldots, \mathbf{u}_N\}$ qui se trouvent dans l'ensemble $\mathcal{J}$ et
  - $\mathbf{1}_{\mathbf{u}_n \in \mathcal{J}}$ est la fonction indicatrice. (Neiderreiter, p.14.).
- La condition sera alors valide pour toutes les fonctions intégrables au sens de Riemann.

La discrépance d'un ensemble de points IV

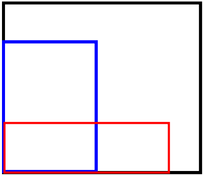
- Nous pouvons en conclure que les ensembles de points $\{\mathbf{u}_1, \ldots, \mathbf{u}_N\}$ désirables sont ceux dont la distribution empirique est proche de la distribution uniforme.
- Les diverses définitions de discrépance qui suivent se veulent des mesures de la distance entre la distribution empirique d'un ensemble de points donnés et la distribution uniforme.

*-Discrépance I

Définition. La $*-$*discrépance* (star discrepancy) $D_N^*\left(\mathbf{u}_1, \ldots, \mathbf{u}_N\right)$ d'un ensemble fini de points $\mathbf{u}_1, \ldots, \mathbf{u}_N$ est

$$D^*\left(\mathbf{u}_1, \ldots, \mathbf{u}_N\right) = \sup_{\mathcal{J}\in\mathcal{K}^*} \left| \frac{1}{N} \sum_{n=1}^{N} \mathbf{1}_{\mathbf{u}_n\in\mathcal{J}} - \lambda\left(\mathcal{J}\right) \right|$$

où $\mathcal{K}^* = \left\{ \prod_{j=1}^{d} [0, \alpha_j) : \forall j \in \{1, \ldots, d\}, 0 \leq \alpha_j \leq 1 \right\}$.



Définition.. Si $\mathcal{S} = \{\mathbf{u}_n : n \in \mathbb{N}\} \subset \mathcal{I}^d$ est une suite de points, alors nous définissons

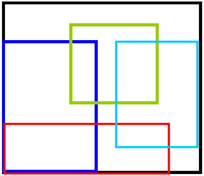
$$D_N^*\left(\mathcal{S}\right) = D^*\left(\mathbf{u}_1, \ldots, \mathbf{u}_N\right).$$

Discrépance extrême I

Définition. La *discrépance extrême* (extreme discrepancy) $D_N(\mathbf{u}_1, \ldots, \mathbf{u}_N)$ d'un ensemble fini de points est

$$D(\mathbf{u}_1, \ldots, \mathbf{u}_N) = \sup_{\mathcal{J} \in \mathcal{K}} \left| \frac{1}{N} \sum_{n=1}^{N} \mathbf{1}_{\mathbf{u}_n \in \mathcal{J}} - \lambda(\mathcal{J}) \right|$$

où $\mathcal{K} = \left\{ \prod_{j=1}^{d} \left[\alpha_j, \beta_j\right) : \forall j \in \{1, \ldots, d\}, 0 \leq \alpha_j \leq \beta_j \leq 1 \right\}$.



Définition. Si $\mathcal{S} = \{\mathbf{u}_n : n \in \mathbb{N}\} \subset \mathcal{I}^d$ est une suite de points, alors nous définissons

$$D_N(\mathcal{S}) = D(\mathbf{u}_1, \ldots, \mathbf{u}_N).$$

Discrépance d'une suite

Soit la suite de points $\mathcal{S} = \{\mathbf{u}_n : n \in \mathbb{N}\} \subset \mathcal{I}^d$. Alors les affirmations suivantes sont équivalentes:

1. La suite $\mathcal{S}$ est uniformément distribuée dans $\mathcal{I}^d$.
2. $\lim_{N\to\infty} D^*\left(\mathbf{u}_1, ..., \mathbf{u}_N\right) = 0$.
3. $\lim_{N\to\infty} D\left(\mathbf{u}_1, ..., \mathbf{u}_N\right) = 0$.

(Neiderreiter, p.17.).

Discrépance minimale I

- Dans le cas unidimensionnel $(d = 1)$, nous avons

$$\begin{aligned} & D^* (u_1, \ldots, u_N) \\ = & \sup_{\mathcal{J} \in \{[0,\alpha) : 0 \leq \alpha \leq 1\}} \left| \frac{1}{N} \sum_{n=1}^{N} \mathbf{1}_{u_n \in \mathcal{J}} - \lambda(\mathcal{J}) \right| \\ = & \sup_{0 \leq \alpha \leq 1} \left| \frac{1}{N} \sum_{n=1}^{N} \mathbf{1}_{u_n \in [0,\alpha)} - \alpha \right| . \end{aligned}$$

- **Question**. Pour un N fixé, quel est le plus petit D^* possible ?

Discrépance minimale II

Théorème.

$$D^*(u_1, \ldots, u_N) = \frac{1}{2N} + \max_{n \in \{1,\ldots,N\}} \left\{ \left| u_n - \frac{2n-1}{2N} \right| \right\}.$$

La démonstration se trouve aux dispositives suivantes (Niederreiter, th. 2.6).

Corollaire. Nous avons donc que

$$D^*(u_1, \ldots, u_N) \geq \frac{1}{2N}$$

et que la discrépance minimale est atteinte lorsque

$$u_n = \frac{2n-1}{2N},$$

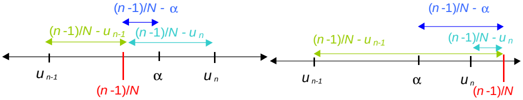
c'est-à-dire que

$$u_1 = \frac{1}{2N}, u_2 = \frac{3}{2N}, u_3 = \frac{5}{2N}, \ldots, u_N = \frac{2N-1}{2N}.$$

Discrépance minimale III

Preuve.

$$
\begin{aligned}
& D^{*}\left(u_1, \ldots, u_N\right) \\
= \; & \sup_{0 \le \alpha \le 1} \left| \frac{1}{N} \sum_{n=1}^{N} \mathbf{1}_{u_n \in [0,\alpha)} - \alpha \right| \text{ où } u_0 = 0 \text{ et } u_{N+1} = 1 \\
= \; & \max_{n \in \{1,\ldots,N+1\}} \sup_{u_{n-1} < \alpha \le u_n} \left| \frac{1}{N} \sum_{n=1}^{N} \mathbf{1}_{u_n \in [0,\alpha)} - \alpha \right| \\
= \; & \max_{n \in \{1,\ldots,N+1\}} \sup_{u_{n-1} \le \alpha \le u_n} \left| \frac{n-1}{N} - \alpha \right| \\
= \; & \max_{n \in \{1,\ldots,N+1\}} \max \left\{ \left| \frac{n-1}{N} - u_{n-1} \right|, \left| \frac{n-1}{N} - u_n \right| \right\}
\end{aligned}
$$



$$
= \max \left\{ \max_{n \in \{1,\ldots,N+1\}} \left\{ \left| \frac{n-1}{N} - u_{n-1} \right| \right\}, \max_{n \in \{1,\ldots,N+1\}} \left\{ \left| \frac{n-1}{N} - u_n \right| \right\} \right\}
$$

Discrépance minimale IV

$$
\begin{aligned}
&= \max\left\{\max_{n\in\{0,\ldots,N\}}\left\{\left|\frac{n}{N}-u_n\right|\right\}, \max_{n\in\{1,\ldots,N+1\}}\left\{\left|\frac{n-1}{N}-u_n\right|\right\}\right\} \\
&= \max\left\{\max_{n\in\{1,\ldots,N\}}\left\{\left|\frac{n}{N}-u_n\right|\right\}, \max_{n\in\{1,\ldots,N\}}\left\{\left|\frac{n-1}{N}-u_n\right|\right\}, 0\right\}
\end{aligned}
$$

parce ce que $u_0 = 0$ et $u_{N+1} = 1$

$$
\begin{aligned}
&= \max_{n\in\{1,\ldots,N\}}\left\{\max\left\{\left|\frac{n}{N}-u_n\right|, \left|\frac{n-1}{N}-u_n\right|\right\}\right\} \\
&= \max_{n\in\{1,\ldots,N\}}\left\{\max\left\{\left|\frac{n-\frac{1}{2}}{N}-u_n+\frac{1}{2N}\right|, \left|\frac{n-\frac{1}{2}}{N}-u_n-\frac{1}{2N}\right|\right\}\right\} \\
&= \frac{1}{2N}+\max_{n\in\{1,\ldots,N\}}\left\{\left|\frac{n-\frac{1}{2}}{N}-u_n\right|\right\}. \ \square
\end{aligned}
$$

Borne pour l'erreur

Dans ce qui suit, nous allons présenter deux bornes pour l'erreur

$$\int_{\mathcal{I}^d} f(\mathbf{u}) \; d\mathbf{u} - \frac{1}{N} \sum_{n=1}^{N} f(\mathbf{u}_n)$$

Elles sont respectivement basées sur :

1. la variation de la fonction à intégrer,
2. le module de continuité de la fonction f.

La variation I

Afin de présenter les résultats concernant les bornes pour l'erreur, nous avons besoin de quelques définitions.

- **Définition**. La *variation* $V(f)$ de la fonction $f : [0,1] \to \mathbb{R}$ est

$$V(f) \equiv \sup_{n \in \mathbb{N}} \sup_{0=x_0<x_1<\ldots<x_n=1} \sum_{i=1}^{n} \left| f(x_i) - f(x_{i-1}) \right| .$$

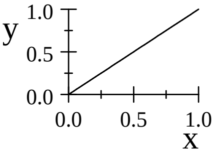
- La variation d'une fonction peut être interprétée comme étant la distance "verticale" parcourue par la fonction.
- Notons que si f admet une dérivée $f^{'}$, alors

$$V(f) = \int_0^1 \left| f^{'}(u) \right| du.$$

La variation II

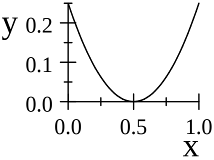
Exemples.

$f(x) = x$



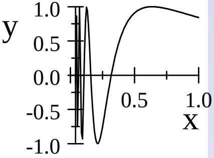
$V(f) = 1$

$f(x) = \left(x - \frac{1}{2}\right)^2$



$V(f) = \frac{1}{2}$

$f(x) = \sin\left(\frac{1}{x}\right)$



$V(f) = \infty$

L'inégalité de Koksma I

Théorème. *Si f est une fonction à variation bornée sur l'intervalle $[0,1]$, alors, pour tout ensemble de points $\{u_1, ..., u_N\} \subset [0,1]$,*

$$\left|\frac{1}{N}\sum_{n=1}^{N} f\left(u_n\right) - \int_0^1 f\left(u\right)\ du\right| \leq V\left(f\right) D^*\left(u_1, ..., u_N\right).$$

La preuve se trouvent dans les prochaines diapositives (Neiderreiter, p.18).

L'inégalité de Koksma II

Théorème. *Pour tout ensemble de points* $P = \{u_1, ..., u_N\} \subset [0,1]$ *et pour tout* $\varepsilon > 0$*, il existe une continue et bornée* $f_{\varepsilon,\mathcal{P}} : [0,1] \to R$ *telle que* $V(f_{\varepsilon,\mathcal{P}}) = 1$ *et*

$$\left| \frac{1}{N} \sum_{n=1}^{N} f(u_i) - \int_{\mathcal{I}^d} f(\mathbf{u}) \; du \right| > D^*(u_1, ..., u_N) - \varepsilon$$

où $\{u_1, ..., u_N\} \subset [0,1]$. (Niederreiter, p. 20).

Ce dernier théorème implique que la borne obtenue est la meilleure borne universelle (qui fonctionne pour toutes les fonctions et tous les ensembles de points) que l'on peut obtenir. Pour faire mieux, il faut choisir l'ensemble de points en fonction de la fonction f à intégrer.

L'inégalité de Koksma III

- **Preuve**. Nous pouvons supposer sans perte de généralité que $u_1 \leq ... \leq u_N$.
- Posons $u_0 = 0$ et $u_{N+1} = 1$.
- L'intégration par partie nous donne

$$\begin{aligned}\int_0^1 f(u)\ du &= 1 \times f(1) - 0 \times f(0) - \int_0^1 uf'(u)\ du \\ &= f(u_{N+1}) - \int_0^1 uf'(u)\ du.\end{aligned}$$

où $f'(u)$ dénote la dérivée de la fonction f.

  - Notons que la preuve originale se base sur l'intégrale de Stieltjes, qui ne requiert pas que f soit différentiable.

L'inégalité de Koksma IV

- Notons que

$$
\begin{aligned}
\frac{1}{N}\sum_{n=1}^{N} f\left(u_n\right) &= \frac{1}{N}\left[\sum_{n=1}^{N} nf\left(u_n\right) - \sum_{n=1}^{N}\left(n-1\right)f\left(u_n\right)\right] \\
&= \frac{1}{N}\left[\sum_{n=1}^{N} nf\left(u_n\right) - \sum_{m=0}^{N-1} mf\left(u_{m+1}\right)\right] \\
&= \frac{1}{N}\left[\sum_{n=0}^{N} nf\left(u_n\right) - \sum_{m=0}^{N} mf\left(u_{m+1}\right) + Nf\left(u_{N+1}\right)\right] \\
&= -\frac{1}{N}\sum_{n=0}^{N} n\left(f\left(u_{n+1}\right) - f\left(u_n\right)\right) + f\left(u_{N+1}\right) \\
&= -\sum_{n=0}^{N}\frac{n}{N}\int_{u_n}^{u_{n+1}} f^{'}\left(u\right)\, du + f\left(u_{N+1}\right).
\end{aligned}
$$

L'inégalité de Koksma V

- Rappelons que

$$D^*\left(u_1, \ldots, u_N\right) = \sup_{0 \leq \alpha \leq 1} \left| \frac{1}{N} \sum_{n=1}^{N} \mathbf{1}_{u_n \in [0,\alpha)} - \alpha \right|.$$

En posant $\alpha = u$ où $u_n < u \leq u_{n+1}$, nous avons

$$D^*\left(u_1, \ldots, u_N\right) \geq \left| \frac{1}{N} \sum_{n=1}^{N} \mathbf{1}_{u_n \in [0,u)} - u \right| = \left| \frac{n}{N} - u \right|.$$

L'inégalité de Koksma VI

- Par conséquent,

$$
\begin{aligned}
& \left| \frac{1}{N} \sum_{n=1}^{N} f\left(u_n\right) - \int_0^1 f\left(u\right)\ du \right| \\
= & \left| \sum_{n=0}^{N} \int_{u_n}^{u_{n+1}} \left(u - \frac{n}{N}\right) f'\left(u\right)\ du \right| \\
\leq & \sum_{n=0}^{N} \int_{u_n}^{u_{n+1}} \left|u - \frac{n}{N}\right| \left|f'\left(u\right)\right|\ du \\
\leq & \sum_{n=0}^{N} \int_{u_n}^{u_{n+1}} D^*\left(u_1, ..., u_N\right) \left|f'\left(u\right)\right|\ du \\
= & D^*\left(u_1, ..., u_N\right) \int_0^1 \left|f'\left(u\right)\right|\ du \\
= & D^*\left(u_1, ..., u_N\right) V\left(f\right). \ \square
\end{aligned}
$$

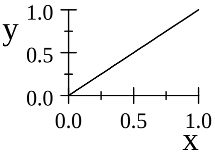
Module de continuité I

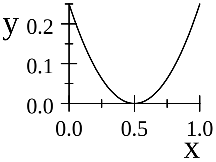
- **Définition**. Le *module de continuité* $\omega(f, \delta)$ de la fonction $f : [0,1] \to \mathbb{R}$ où $\delta \in [0,1]$ est
$$\omega(f, \delta) \equiv \sup_{\substack{u,v \in [0,1] \\ |u-v| \leq \delta}} |f(v) - f(u)|.$$
- **Rappel**. La fonction f est continue au point x_0 si, pour tout $\varepsilon > 0$, il existe un nombre $\delta(x_0, \varepsilon) > 0$ tel que $|x - x_0| \leq \delta(x_0, \varepsilon)$ implique que
$$|f(x) - f(x_0)| < \varepsilon.$$
- **Intuitivement**... Dans le cas du module de continuité, on ne fixe pas le x_0. Nous choisissons le δ et nous cherchons le "pire" (ou un des pires) x_0, c'est-à-dire l'endroit (ou les endroits) ou la distance entre $f(x)$ et $f(x_0)$ est la plus grande lorsque $|x - x_0| \leq \delta$.

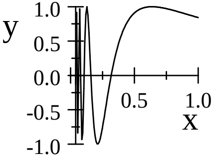
Module de continuité II

Exemples.

$$f(x) = x \qquad f(x) = \left(x - \frac{1}{2}\right)^2 \qquad f(x) = \sin\left(\frac{1}{x}\right)$$







$$\omega(f,\delta) = \delta \qquad \omega(f,\delta) = (1-\delta)\,\delta \qquad \omega(f,\delta) = 2$$

Les calculs sont à la page suivante.

Module de continuité III

$$\omega(f,\delta) = \sup_{\substack{u,v\in[0,1]\\|u-v|\le\delta}} |v-u| = \delta$$

$$\omega(f,\delta) = \sup_{\substack{u,v\in[0,1]\\|u-v|\le\delta}} \left|\left(v-\frac{1}{2}\right)^2 - \left(u-\frac{1}{2}\right)^2\right|$$

$$= \sup_{\substack{u,v\in[0,1]\\|u-v|\le\delta}} \left|\left(\left(v-\frac{1}{2}\right)+\left(u-\frac{1}{2}\right)\right)\left(\left(v-\frac{1}{2}\right)-\left(u-\frac{1}{2}\right)\right)\right|$$

$$= \sup_{\substack{u,v\in[0,1]\\|u-v|\le\delta}} |v+u-1|\,|v-u| = (1-\delta)\,\delta$$

$$\omega(f,\delta) = \sup_{\substack{u,v\in[0,1]\\|u-v|\le\delta}} \left|\sin\left(\frac{1}{v}\right)-\sin\left(\frac{1}{u}\right)\right| = 2.$$

Module de continuité IV

Théorème. *Si f est une fonction continue sur l'intervalle* $[0,1]$, *alors, pour tout ensemble de points* $\{u_1, ..., u_N\} \subset [0,1]$,

$$\left| \frac{1}{N} \sum_{n=1}^{N} f(u_n) - \int_0^1 f(u) \ du \right| \leq \omega\left(f, D_N^*(u_1, ..., u_N)\right).$$

(Neiderreiter, p.19).

Module de continuité V

- **Preuve**. Nous pouvons supposer sans perte de généralité que $u_1 \leq \ldots \leq u_N$.
- Posons $u_0 = 0$ et $u_{N+1} = 1$.
- La preuve se base sur le théorème de la moyenne pour les intégrales (mean-value theorem):
  - **Théorème**. *Si* $f : \mathbb{R} \to \mathbb{R}$ *est une fonction continue sur l'intervalle* $[a, b]$, *alors il existe une constante* $c \in (a, b)$ *telle que* $\int_a^b f(x)\, dx = (b - a) f(c)$.
- Nous en déduisons qu'il existe un nombre $t_n \in \left(\frac{n-1}{N}, \frac{n}{N}\right)$ tel que $\int_{\frac{n-1}{N}}^{\frac{n}{N}} f(u)\, du = \frac{1}{N} f(t_n)$. Ainsi,

$$\int_0^1 f(u)\, du = \sum_{n=1}^{N} \int_{\frac{n-1}{N}}^{\frac{n}{N}} f(u)\, du = \frac{1}{N} \sum_{n=1}^{N} f(t_n).$$

Module de continuité VI

- Par conséquent,

$$
\begin{aligned}
\left|\frac{1}{N}\sum_{n=1}^{N} f\left(u_n\right) - \int_0^1 f\left(u\right)\ du\right| &= \left|\frac{1}{N}\sum_{n=1}^{N} f\left(u_n\right) - \frac{1}{N}\sum_{n=1}^{N} f\left(t_n\right)\right| \\
&\leq \frac{1}{N}\sum_{n=1}^{N} \left|f\left(u_n\right) - f\left(t_n\right)\right| \\
&\leq \frac{1}{N}\sum_{n=1}^{N} \sup_{\substack{u,v\in[0,1] \\ |u-v|\leq|u_n-t_n|}} \left|f\left(v\right) - f\left(u\right)\right| \\
&\leq \sup_{\substack{u,v\in[0,1] \\ |u-v|\leq \max_{n\in\{1,\ldots,N\}} |u_n-t_n|}} \left|f\left(v\right) - f\left(u\right)\right| \\
&= \omega\left(f, \max_{n\in\{1,\ldots,N\}} \left|u_n - t_n\right|\right).
\end{aligned}
$$

Module de continuité VII

- Le module de continuité étant une fonction non-décroissante de δ, nous avons que $\omega\left(f, \max_{n\in\{1,\ldots,N\}} |u_n - t_n|\right) \leq \omega\left(f, D_N^*\left(u_1, \ldots, u_N\right)\right)$ puisque

$$\max_{n\in\{1,\ldots,N\}} \left\{|t_n - u_n|\right\} \leq D_N^*\left(u_1, \ldots, u_N\right).$$

- En effet,

$$\begin{aligned} & D_N^*\left(u_1, \ldots, u_N\right) \\ = \quad & \max_{n\in\{1,\ldots,N\}} \left\{\max\left\{\left|\frac{n}{N} - u_n\right|, \left|\frac{n-1}{N} - u_n\right|\right\}\right\} \\ \geq \quad & \max_{n\in\{1,\ldots,N\}} \left\{|t_n - u_n|\right\} \text{ car } t_n \in \left(\frac{n-1}{N}, \frac{n}{N}\right). \; \square \end{aligned}$$

Suite à discrépance faible I

1. Nous savons maintenant que pour tout ensemble de N points $\{u_1, \dots, u_N\} \subset [0,1]$, $D^*(u_1, \dots, u_N) \geq \frac{1}{2N}$ et qu'il y a égalité lorsque

$$\widetilde{u}_1 = \frac{1}{2N}, \widetilde{u}_2 = \frac{3}{2N}, \dots, \widetilde{u}_n = \frac{2n-1}{2N}, \dots, \widetilde{u}_N = \frac{2N-1}{2N}.$$

2. Le problème avec cet ensemble de points est que si on désire améliorer la précision de notre intégration numérique en ajoutant des points supplémentaires, on doit tout recommencer, c'est-à-dire que l'on ne peut pas conserver les N premiers points générés et conserver la discrépance minimale

$$D^*(\widetilde{u}_1, \dots, \widetilde{u}_N, u_{N+1}, \dots, u_M) > \frac{1}{2M}$$

sauf pour quelques rares $M \in \mathbb{N}$.

Suite à discrépance faible II

3. En fait, il a été démontré qu'il n'existe pas de suite $\mathcal{S} = \{u_n : n \in \mathbb{N}\} \subset [0,1]$ telle que

$$D_N^*(\mathcal{S}) = O\left(N^{-1}\right) \text{ pour tout } N \in \mathbb{N}.$$

   3.1 Rappelons qu'une fonction $g(N) = O\left(N^{-k}\right)$ si $\sup_{N\in\mathbb{N}} \left|N^k g(N)\right| < \infty$.

4. De plus, il existe une constante c telle que pour toute suite $\mathcal{S} \subset [0,1]$,

$$D_N^*(\mathcal{S}) \geq c\frac{\log N}{N}.$$

   (Neiderreiter, p.24).

5. Nous allons donc construire des suites dont la discrépance est de l'ordre de la borne ci-dessus.

Les bases I

- Nous utilisons généralement la base 10. La notation 10 533, 2 signifie la quantité

$$(10533,2)_{10} = 1\times 10^4 + 0\times 10^3 + 5\times 10^2 + 3\times 10^1 + 3\times 10^0 + 2\times 10^{-1}.$$

- Or la même notation en base 6 signifie une quantité différente :

$$\begin{aligned} & (10533,2)_6 \\ = \; & 1\times 6^4 + 0\times 6^3 + 5\times 6^2 + 3\times 6^1 + 3\times 6^0 + 2\times 6^{-1} \\ = \; & \left(1497.\,\overline{3}\right)_{10}. \end{aligned}$$

- Pour exprimer un nombre en base $b \in \{2,3,4,...\}$, on le fait avec les entiers $\{0,1,...,b-1\}$.

Van der Corput I

- Choisissons notre base $b \in \{2, 3, 4, ...\}$ préférée.
- Chaque entier $n \in \{0, 1, 2, ...\}$ a une représentation unique en terme de cette base :

$$n = \sum_{j=0}^{\infty} a_j\left(n, b\right) b^j$$

où, pour tout j, $a_j\left(n, b\right) \in \{0, 1, ..., b-1\}$ et $a_j\left(n, b\right) = 0$ pour tous les j suffisamment grands (c'est-à-dire que la somme ci-dessus est finie).

Exemples

n	$b=2$	$b=3$	$b=4$	n	$b=2$	$b=3$	$b=4$
0	$(0)_2$	$(0)_3$	$(0)_4$	$(5)_{10}$	$(101)_2$	$(12)_3$	$(11)_4$
$(1)_{10}$	$(1)_2$	$(1)_3$	$(1)_4$	$(6)_{10}$	$(110)_2$	$(20)_3$	$(12)_4$
$(2)_{10}$	$(10)_2$	$(2)_3$	$(2)_4$	$(7)_{10}$	$(111)_2$	$(21)_3$	$(13)_4$
$(3)_{10}$	$(11)_2$	$(10)_3$	$(3)_4$	$(8)_{10}$	$(1000)_2$	$(22)_3$	$(20)_4$
$(4)_{10}$	$(100)_2$	$(11)_3$	$(10)_4$	$(9)_{10}$	$(1001)_2$	$(100)_3$	$(21)_4$

Van der Corput II

- Rappelons que $n = \sum_{j=0}^{\infty} a_j(n,b)\, b^j$.
- **Définition**. Pour tout entier $b \in \{2,3,4,\ldots\}$, nous posons

$$\phi_b(n) \equiv \sum_{j=0}^{\infty} a_j(n,b)\, b^{-j-1}.$$

  La *suite de van der Corput* en base b est $\mathcal{C}_b = \{\phi_b(n) : n \in \{0,1,2,\ldots\}\}$.
- **Remarque**. Le cas $b = 2$ date de 1935.

Van der Corput III

Exemple. Rappelons que $\phi_b(n) \equiv \sum_{j=0}^{\infty} a_j(n,b)\, b^{-j-1}$.

n	$b=2$	$\phi_b(n)$
0	$(0)_2$	$0\times 2^{-1} = 0,0000$
1	$(1)_2$	$1\times 2^{-1} = 0,5000$
2	$(10)_2$	$1\times 2^{-2} + 0\times 2^{-1} = 0,2500$
3	$(11)_2$	$1\times 2^{-2} + 0\times 2^{-1} = 0,7500$
4	$(100)_2$	$1\times 2^{-3} + 0\times 2^{-2} + 0\times 2^{-1} = 0,1250$
5	$(101)_2$	$1\times 2^{-3} + 0\times 2^{-2} + 1\times 2^{-1} = 0,6250$
6	$(110)_2$	$1\times 2^{-3} + 1\times 2^{-2} + 0\times 2^{-1} = 0,3750$
7	$(111)_2$	$1\times 2^{-3} + 1\times 2^{-2} + 1\times 2^{-1} = 0,8750$
8	$(1000)_2$	$1\times 2^{-4} + 0\times 2^{-3} + 0\times 2^{-2} + 0\times 2^{-1} = 0,0625$

Van der Corput IV

- On peut rajouter des points mais, pour conserver l'uniformité, il vaut mieux s'en tenir à une puissance de la base. Cependant, pour les très grandes valeurs de N cela aura plus ou moins d'impact.

Van der Corput V

Exemple. Rappelons que $\phi_b(n) \equiv \sum_{j=0}^{\infty} a_j(n,b)\, b^{-j-1}$.

n	$b=3$	$\phi_b(n)$
0	$(0)_3$	$0 \times 3^{-1} = 0,\overline{0}$
1	$(1)_3$	$1 \times 3^{-1} = 0,\overline{3}$
2	$(2)_3$	$2 \times 3^{-1} = 0,\overline{6}$
3	$(10)_3$	$1 \times 3^{-2} + 0 \times 3^{-1} = 0,\overline{1}$
4	$(11)_3$	$1 \times 3^{-2} + 1 \times 3^{-1} = 0,\overline{4}$
5	$(12)_3$	$1 \times 3^{-2} + 2 \times 3^{-1} = 0,\overline{7}$
6	$(20)_3$	$2 \times 3^{-2} + 0 \times 3^{-1} = 0,\overline{2}$
7	$(21)_3$	$2 \times 3^{-2} + 1 \times 3^{-1} = 0,\overline{5}$
8	$(22)_3$	$2 \times 3^{-2} + 2 \times 3^{-1} = 0,\overline{8}$

Van der Corput VI

Les résultats. Si $\mathcal{C}_b$ dénote la suite de van der Corput en base b, alors :

1. $D_N^*(\mathcal{C}_b) = O\left(\frac{\log N}{N}\right)$.
2. $\lim_{N\to\infty} \frac{N}{\log N} D_N^*(\mathcal{C}_b) = c(b) = \begin{cases} \frac{b^2}{4(b+1)\log b} & \text{si } b \text{ est pair} \\ \frac{b-1}{4\log b} & \text{si } b \text{ est impair} \end{cases}$

- Le point 1 nous signale que $\mathcal{C}_b$ est une suite à discrépance faible alors que le point 2 nous donne la valeur asymptotique de la constante universelle.
- Le cas $b = 3$ est asymptotiquement optimal.

$$\{c(b) : b \in \{2, 3, ...\}\} \approx \{0,481; 0,455; 0,577; 0,621; 0,718; 0,771; \ldots\}$$

Rappelons que $D_N^*(\mathcal{S}) \geq c\frac{\log N}{N}$ pour toute suite $\mathcal{S} \subset [0, 1]$. (Neiderreiter, p.25).

Généralisation de Van der Corput I

- Il est possible de faire asymptotiquement mieux en généralisant la construction des suites de Van der Corput.
- **Définition**. Pour tout entier $b \in \{2, 3, 4, ...\}$, nous posons

$$\phi_{b,\sigma}(n) \equiv \sum_{j=0}^{\infty} \sigma(a_j(n,b))\, b^{-j-1}$$

  où σ est une permutation de $\{0, 1, ..., b-1\}$.
- La *suite de Van der Corput généralisée* en base b est $\mathcal{C}_b = \{\phi_{b,\sigma}(n) : n \in \{0, 1, 2, ...\}\}$.

Généralisation de Van der Corput II

Exemple. Posons $b = 3$. La fonction de permutation $\sigma(0) = 0, \sigma(1) = 2, \sigma(2) = 1$.

n	$b = 3$	$\phi_b(n)$	$\phi_{b,\sigma}(n)$
0	$(0)_3$	$0 \times 3^{-1} = 0,\overline{0}$	$0 \times 3^{-1} = 0,\overline{0}$
1	$(1)_3$	$1 \times 3^{-1} = 0,\overline{3}$	$2 \times 3^{-1} = 0,\overline{6}$
2	$(2)_3$	$2 \times 3^{-1} = 0,\overline{6}$	$1 \times 3^{-1} = 0,\overline{3}$
3	$(10)_3$	$1 \times 3^{-2} + 0 \times 3^{-1} = 0,\overline{1}$	$2 \times 3^{-2} + 0 \times 3^{-1} = 0,\overline{2}$
4	$(11)_3$	$1 \times 3^{-2} + 1 \times 3^{-1} = 0,\overline{4}$	$2 \times 3^{-2} + 2 \times 3^{-1} = 0,\overline{8}$
5	$(12)_3$	$1 \times 3^{-2} + 2 \times 3^{-1} = 0,\overline{7}$	$2 \times 3^{-2} + 1 \times 3^{-1} = 0,\overline{5}$
6	$(20)_3$	$2 \times 3^{-2} + 0 \times 3^{-1} = 0,\overline{2}$	$1 \times 3^{-2} + 0 \times 3^{-1} = 0,\overline{1}$
7	$(21)_3$	$2 \times 3^{-2} + 1 \times 3^{-1} = 0,\overline{5}$	$1 \times 3^{-2} + 2 \times 3^{-1} = 0,\overline{7}$
8	$(22)_3$	$2 \times 3^{-2} + 2 \times 3^{-1} = 0,\overline{8}$	$1 \times 3^{-2} + 1 \times 3^{-1} = 0,\overline{4}$

Généralisation de Van der Corput III

- En 1991, le meilleur choix de paramètres étaient ceux trouvés par Faure en prenant $b = 12$ et une permutation spécifique de $\{0, 1, ..., 11\}$. Il a obtenu $\lim_{N\to\infty} \frac{N}{\log N} D_N^* (\mathcal{C}_b) = 0,223...$ (Neiderreiter, p.26).

Suites basées sur des nombres irrationnels I

- **Définition**. Pour tout nombre réel z, nous notons $\lfloor z \rfloor$ sa partie entière et
$$\langle z \rangle = z - \lfloor z \rfloor$$
sa partie fractionnaire.
- **Exemple**.
$$\begin{array}{rcl} \lfloor \pi \rfloor & = & 3 \\ \langle \pi \rangle & = & 0.141\,592\,654... \end{array}$$
- **Définition**. Pour tout nombre irrationnel z, nous pouvons construire la suite $S(z) = \{u_n : n \in \{0, 1, 2, ...\}\}$ où
$$u_n = \langle nz \rangle .$$

Suites basées sur des nombres irrationnels II

Exemple.

$$
\begin{array}{rlrl}
 & & \langle 0\pi \rangle &= 0.000\,000\,000\ldots = u_0 \\
\pi &= 3.141\,592\,654\ldots & \langle \pi \rangle &= 0.141\,592\,654\ldots = u_1 \\
2\pi &= 6.283\,185\,307\ldots & \langle 2\pi \rangle &= 0.283\,185\,307\ldots = u_2 \\
3\pi &= 9.424\,777\,961\ldots & \langle 3\pi \rangle &= 0.424\,777\,961\ldots = u_3 \\
4\pi &= 12.566\,370\,614\ldots & \langle 4\pi \rangle &= 0.566\,370\,614\ldots = u_4
\end{array}
$$

Suites basées sur des nombres irrationnels III

- La théorie permettant de déterminer quels sont les nombres irrationnels qui engendrent des suites à faible discrépance est basée sur les fractions continues.
- Comme cela est techniquement lourd pour un cours d'introduction, nous nous contenterons de mentionner qu'il existe des critères nous permettant de choisir ces nombres réels.

Cas multidimensionnel

Nous considérons maintenant l'intégration numérique sur l'hypercube de dimension d :

$$\int_{\mathcal{I}^d} f(\mathbf{u})\, d\mathbf{u} \approx \frac{1}{N} \sum_{n=1}^{N} f(\mathbf{u}_n)$$

Cas multidimensionnel

Il existe plusieurs bornes pour l'erreur $\int_{\mathcal{I}^d} f(\mathbf{u})\, d\mathbf{u} - \frac{1}{N}\sum_{n=1}^{N} f(\mathbf{u}_n)$ commise en réalisant l'intégration numériquement.

Nous présentons brièvement et sans entrer dans les détails la borne basée sur la variation de la fonction et celle basée sur son module de continuité.

Variation au sens de Hardy et Krause I

1. Considérons la fonction f sur l'hyper-cube unitaire $\mathcal{I}^d$ et le sous-intervalle $\mathcal{J} \subseteq \mathcal{I}^d$.
2. Soit $\Delta(f; J)$ la somme alternée des valeurs de f aux sommets de $\mathcal{J}$ (c'est-à-dire que les valeurs de la fonction à des sommets adjacents ont des signes opposés).
3. **Définition**. La *variation de f sur $\mathcal{I}^d$ au sens de Vitali* est définie par

$$V^{(d)}(f) = \sup_{\mathcal{P}} \sum_{J \in \mathcal{P}} |\Delta(f; J)|$$

où le supremum est pris sur l'ensemble de toutes les partitions $\mathcal{P}$ de $\mathcal{I}^d$ en sous-intervalles.

Variation au sens de Hardy et Krause II

3.1 Dans le cas où les dérivées partielles existent et sont continues sur $\mathcal{I}^d$, nous avons

$$V^{(d)}(f) = \int_0^1 \dots \int_0^1 \left| \frac{\partial^d f}{\partial u_1 \dots \partial u_d} \right| du_1 \dots du_d.$$

4. **Définition**. Pour $1 \leq k \leq d$ et $1 \leq i_1 < i_2 < \dots < i_k \leq d$, $V^{(k)}(f; i_1, \dots, i_k)$ est la variation au sens de Vitali de la restriction de la fonction f au sous-espace de dimension k

$$\left\{ (u_1, \dots, u_d) \in \mathcal{I}^d : u_j = 1 \text{ for } j \neq i_1, \dots, i_k \right\}.$$

Alors

$$V(f) = \sum_{k=1}^{d} \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq d} V^{(k)}(f; i_1, \dots, i_k)$$

Variation au sens de Hardy et Krause III

est la *variation de f sur $\mathcal{I}^d$ au sens de Hardy and Krause*, et f est à variation bornée au sens de Hardy et Krause si $V(f)$ est fini. (Niederreiter, p. 19).

Borne pour l'erreur basée sur la variation de la fonction I

Théorème. *Si la fonction $f : I^d \to R$ est à variation bornée $(V(f) < \infty)$ au sens de Hardy et Krause alors*

$$\left| \frac{1}{N} \sum_{n=1}^{N} f(\mathbf{u}_i) - \int_{\mathcal{I}^d} f(\mathbf{u}) \; d\mathbf{u} \right| \leq V(f) \, D^* (\mathbf{u}_1, ..., \mathbf{u}_N)$$

où $\{\mathbf{u}_1, ..., \mathbf{u}_N\} \subset \mathcal{I}^d$. (Niederreiter, p. 20).

Borne pour l'erreur basée sur la variation de la fonction II

Théorème. *Pour tout ensemble de points* $P = \{\mathbf{u}_1, ..., \mathbf{u}_N\} \subset I^d$ *et pour tout* $\varepsilon > 0$*, il existe une fonction continue et bornée* $f_{\varepsilon,\mathcal{P}} : I^d \to R$ *telle que* $V\left(f_{\varepsilon,\mathcal{P}}\right) = 1$ *et*

$$\left| \frac{1}{N} \sum_{n=1}^{N} f\left(\mathbf{u}_i\right) - \int_{\mathcal{I}^d} f\left(\mathbf{u}\right) \ d\mathbf{u} \right| > D^*\left(\mathbf{u}_1, ..., \mathbf{u}_N\right) - \varepsilon$$

où $\{\mathbf{u}_1, ..., \mathbf{u}_N\} \subset \mathcal{I}^d$. (Niederreiter, p. 20).

Ce dernier théorème implique que la borne obtenue est la meilleure borne universelle (qui fonctionne pour toutes les fonctions et tous les ensembles de points) que l'on peut obtenir. Pour faire mieux, il faut choisir l'ensemble de points en fonction de la fonction f à intégrer.

Borne pour l'erreur basée sur le module de continuité I

- Une version multidimensionnelle du module de continuité.
- **Définition**. Le *module de continuité* $\omega(f, \delta)$ de la fonction continue $f : \mathcal{I}^d \to \mathbb{R}$ est

$$\omega(f, \delta) = \sup_{\substack{\mathbf{u}, \mathbf{v} \in \mathcal{I}^d \\ \|\mathbf{u} - \mathbf{v}\| \leq \delta}} |f(\mathbf{u}) - f(\mathbf{v})|$$

où $\delta > 0$ et

$$\|\mathbf{u}\| \equiv \max_{i \in \{1, \ldots, d\}} |u_i| .$$

(Niederreiter, p. 20).

Borne pour l'erreur basée sur le module de continuité II

- **Théorème**. *Si la fonction $f : I^d \to R$ est continue sur I^d, alors*

$$\left| \frac{1}{N} \sum_{n=1}^{N} f\left(\mathbf{u}_i\right) - \int_{\mathcal{I}^d} f\left(\mathbf{u}\right) \ d\mathbf{u} \right| \leq 4\omega\left(f, \left[D^*\left(\mathbf{u}_1, ..., \mathbf{u}_N\right)\right]^{\frac{1}{d}}\right).$$

(Niederreiter, p. 21).

Erreur d'intégration I

- Suite aux résultats précédents, nous concluons que l'erreur liée à l'intégration numérique dépend en grande partie de la discrépance de la suite choisie.
- Il nous faut donc construire des suites en dimensions multiples et mesurer leur discrépance.

Suites à discrépance faible

Dans ce qui suit, nous présentons

1. les suites de Halton,
2. les suites de Hammersley.

Halton I

- Les suites de Halton sont une extension naturelle des suites de Van der Corput.
- Rappelons que

$$n = \sum_{j=0}^{\infty} a_j\left(n, b\right) b^j .$$

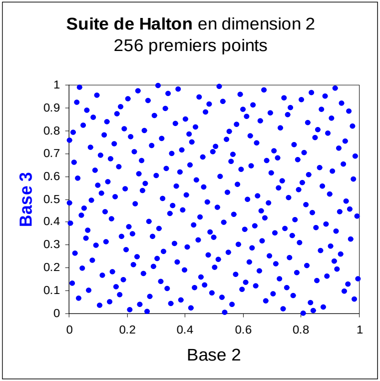
et que pour tout entier $b \in \{2, 3, 4, \ldots\}$, nous posons

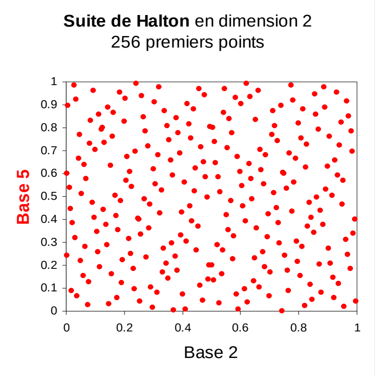
$$\phi_b\left(n\right) \equiv \sum_{j=0}^{\infty} a_j\left(n, b\right) b^{-j-1} .$$

- **Définition**. (1960) Soit $b_1, \ldots, b_d \in \{2, 3, 4, \ldots\}$. La *suite de Halton* en dimension d est $\mathcal{H}_{b_1,\ldots,b_d} = \{\mathbf{u}_n : n \in \{0, 1, 2, \ldots\}\}$ où

$$\mathbf{u}_n = \left(\phi_{b_1}\left(n\right), \ldots, \phi_{b_d}\left(n\right)\right).$$

Halton II


Suite de Halton en dimension 2
256 premiers points
Base 3
Base 2
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.2
0.4
0.6
0.8


Suite de Halton en dimension 2
256 premiers points
Base 5
Base 2
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.2
0.4
0.6
0.8

Halton III

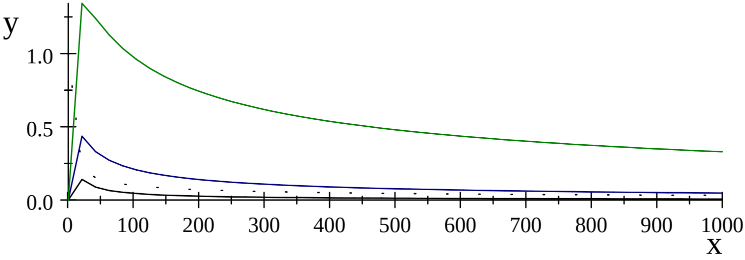
Théorème. *Si* $H_{b_1,\ldots,b_d} = \{\mathbf{u}_n : n \in \{0, 1, 2, \ldots\}\}$ *est une suite de Halton telle que les bases* $b_1, \ldots, b_d$ *sont relativement premières entre elles, alors pour tout* $N \in \{2, 3, \ldots\}$,

$$\begin{aligned} D_N^*\left(\mathcal{H}_{b_1,\ldots,b_d}\right) &\leq \frac{d}{N} + \frac{1}{N}\prod_{i=1}^{d}\left(\frac{b_i - 1}{2\log b_i}\log N + \frac{b_i + 1}{2}\right) \\ &= \kappa\left(b_1, \ldots, b_d\right) O\left(\frac{\left(\log N\right)^d}{N}\right) + O\left(\frac{1}{N}\right) \end{aligned}$$

où $\kappa\left(b_1, \ldots, b_d\right)$ *est une constante qui dépend du choix des bases.* (Niederreiter, p. 29-31).

Halton IV

Étude de la vitesse de convergence de la borne...



Noir : $N^{-1}(\log N)^1$

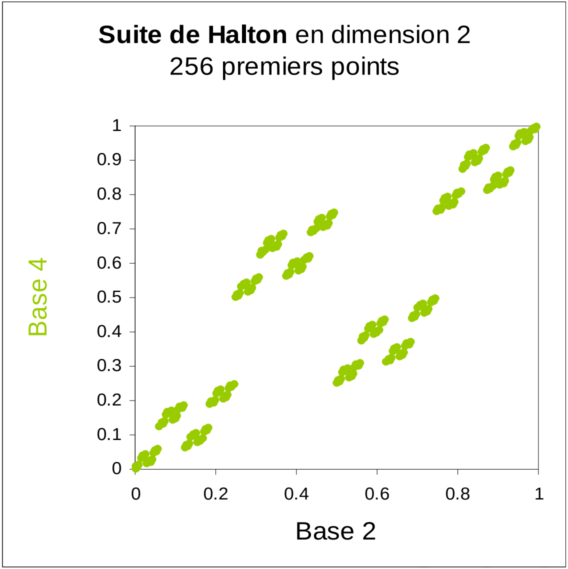
Bleu : $N^{-1}(\log N)^2$

Vert : $N^{-1}(\log N)^3$

Noir, pointillé : $\frac{1}{\sqrt{N}}$

Halton V

N'oubliez pas que les bases doivent être relativement premières...



Hammersley I

- Rappelons que

$$n = \sum_{j=0}^{\infty} a_j(n, b)\, b^j.$$

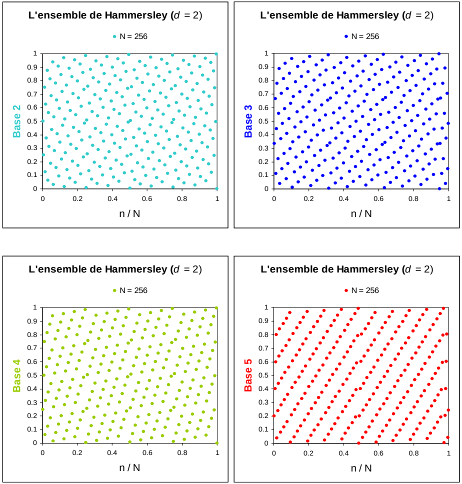
et que pour tout entier $b \in \{2, 3, 4, ...\}$, nous posons

$$\phi_b(n) \equiv \sum_{j=0}^{\infty} a_j(n, b)\, b^{-j-1}.$$

- **Définition**. (1960) Soit $b_1, ..., b_{d-1} \in \{2, 3, 4, ...\}$. L'*ensemble de N points de Hammersley* en dimension $d > 2$ est $\mathcal{A}_{b_1,...,b_{d-1}} = \{\mathbf{u}_n : n \in \{0, 1, 2, ...\}\}$ où

$$\mathbf{u}_n = \left(\frac{n}{N}, \phi_{b_1}(n), ..., \phi_{b_{d-1}}(n)\right).$$

Hammersley II



Hammersley III

- **Théorème**. *Si* $A_{b_1,\ldots,b_{d-1}} = \{\mathbf{u}_n : n \in \{0, 1, 2, \ldots, N\}\} \subset I^d$ *est un ensemble de* N *points de Hammersley tel que les bases* $b_1, \ldots, b_{d-1}$ *sont relativement premières entre elles, alors*

$$D_N^*\left(\mathcal{H}_{b_1,\ldots,b_d}\right) \leq \frac{d}{N} + \frac{1}{N}\prod_{i=1}^{d-1}\left(\frac{b_i - 1}{2\log b_i}\log N + \frac{b_i+1}{2}\right).$$

  (Niederreiter, p. 29-31).
- L'avantage des ensembles de Hammersley sur les suites de Halton est le taux de convergence de la borne pour la $*-$discrépance est un peu plus rapide (elle dépend de $d-1$ base au lieu de d).
- L'inconvénient est qu'il faut déterminer à l'avance le nombre de points utilisés.

Les (k,m,d)-treillis et les (k,d)-suites I

- Rappelons que si $\mathcal{H}_{b_1,\ldots,b_d} = \{\mathbf{u}_n : n \in \{0,1,2,\ldots\}\}$ est une suite de Halton telle que les bases $b_1, \ldots, b_d$ sont relativement premières, alors pour tout $N \in \{2,3,\ldots\}$,

$$D_N^*\left(\mathcal{H}_{b_1,\ldots,b_d}\right) \leq \underbrace{\left(\prod_{i=1}^{d} \frac{b_i - 1}{2\log b_i}\right)}_{=\kappa(b_1,\ldots,b_d)} \frac{(\log N)^d}{N} + \frac{1}{N}\left(d + \prod_{i=1}^{d} \frac{b_i+1}{2}\right).$$

- Or, on peut montrer que

$$\lim_{d\to\infty} \frac{\log \kappa\left(b_1,\ldots,b_d\right)}{d \log d} = \lim_{d\to\infty} \frac{\sum_{i=1}^{d} \log\left(b_i - 1\right) - \log\left(\log b_i^2\right)}{d \log d} = 1,$$

c'est-à-dire que la constante $\kappa\left(b_1,\ldots,b_d\right) \uparrow \infty$ "superexponentiellement". (Niederreiter, p. 47).

- On retrouve le même genre d'inégalité pour les suites de Hammersley.

Les (k,m,d)-treillis et les (k,d)-suites II

- Ceci implique qu'en pratique, les bornes pour les erreurs commises ne sont utilisables que pour les problèmes de petites dimensions.
- C'est pourquoi des suites de points ayant des constantes beaucoup plus petites ont été construites.

Les (k,m,d)-treillis et les (k,d)-suites III

- **Définition**. Un sous-intervalle E de $\mathcal{I}^d$ de la forme

$$E = \prod_{j=1}^{d} \left(\frac{a_j}{b^{c_j}}, \frac{a_j + 1}{b^{c_j}} \right], \; c_j \in \{0, 1, 2, ...\}, \; a_j \in \{0, 1, ..., b^{c_j} - 1\}$$

  est appelé un *intervalle élémentaire en base b*.
- **Exemple**. Si la dimension $d = 3$, la base $b = 3$, $c_1 = 1$, $c_2 = 1$ et $c_3 = 2$ alors les deux premières dimensions sont subdivisées par les intervalles $\left(0, \frac{1}{3}\right], \left(\frac{1}{3}, \frac{2}{3}\right], \left(\frac{2}{3}, 1\right]$ tandis que la dernière dimension est séparée par les intervalles $\left(0, \frac{1}{9}\right], \left(\frac{1}{9}, \frac{2}{9}\right], ..., \left(\frac{8}{9}, 1\right]$. Il y a donc $3 \times 3 \times 9 = 81$ intervalles élémentaires possibles dans ce cas.

Les (k,m,d)-treillis et les (k,d)-suites IV

- **Définition**. Soit $m \in \{0, 1, 2, ...\}$ et $k \in \{0, 1, ..., m\}$. Un (k, m, d) *−treillis en base b* est un ensemble $P = \{\mathbf{u}_1, ..., \mathbf{u}_{b^m}\}$ de b^m points dans $\mathcal{I}^d$ tel que

$$\sum_{i=1}^{b^m} \mathbf{1}_{\{\mathbf{u}_i \in E\}} = b^k$$

pour tout intervalle élémentaire E tel que

$$\lambda(E) = \frac{1}{b^{m-k}},$$

c'est-à-dire que chaque intervalle élémentaire ayant un volume de $\frac{1}{b^{m-k}}$ contient exactement b^k points de notre ensemble P.

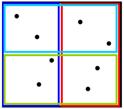
- Intuitivement, pour les intervalles élémentaires ayant un volume assez grand $\left(\geq 1/b^{m-k}\right)$ la proportion de points contenus dans l'intervalle correspond à son volume.

Les (k,m,d)-treillis et les (k,d)-suites V

Exemple. Supposons que la dimension est $d = 2$, la base est $b = 2$ et que $m = 3$, c'est-à-dire que notre ensemble de points P en contient $2^3 = 8$.
Si $k = 2$ alors on considère que les intervalles élémentaires de volume $\frac{1}{b^{3-2}} = \frac{1}{2}$. Or, il n'y a que quatre de ces intervalles:

$$\left(0, \frac{1}{2}\right] \times (0,1]\,, \left(\frac{1}{2}, 1\right] \times (0,1]\,, (0,1] \times \left(0, \frac{1}{2}\right]\,, (0,1] \times \left(\frac{1}{2}, 1\right]$$

Notre ensemble de points P sera un $(2,3,2)$ $-$treillis en base 2, si E_1, E_2, E_3 et E_4 contiennent chacun exactement $2^2 = 4$ points.

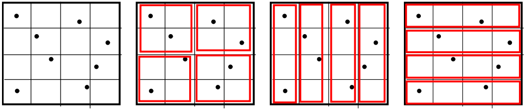


Les (k,m,d)-treillis et les (k,d)-suites VI

Exemple (suite). Si $k = 1$ alors on considère que les intervalles élémentaires de volume $\frac{1}{b^{3-1}} = \frac{1}{4}$. Dans ce cas, il y a 12 intervalles élémentaires:

$$\left\{ \begin{array}{c} \left(0, \frac{1}{4}\right] \times (0,1], \left(\frac{1}{4}, \frac{2}{4}\right] \times (0,1], \left(\frac{2}{4}, \frac{3}{4}\right] \times (0,1], \left(\frac{3}{4}, 1\right] \times (0,1], \\ (0,1] \times \left(0, \frac{1}{4}\right], (0,1] \times \left(\frac{1}{4}, \frac{2}{4}\right], (0,1] \times \left(\frac{2}{4}, \frac{3}{4}\right], (0,1] \times \left(\frac{3}{4}, 1\right] \\ \left(0, \frac{1}{2}\right] \times \left(0, \frac{1}{2}\right], \left(0, \frac{1}{2}\right] \times \left(\frac{1}{2}, 1\right], \left(\frac{1}{2}, 1\right] \times \left(0, \frac{1}{2}\right], \left(\frac{1}{2}, 1\right] \times \left(\frac{1}{2}, 1\right] \end{array} \right\}$$

Notre ensemble de points P sera un $(1,3,2)$ $-$treillis en base 2, si chacun des intervalles élémentaires E_1, E_2, ..., E_{12} contiennent chacun exactement $2^1 = 2$ points.



Les (k,m,d)-treillis et les (k,d)-suites VII

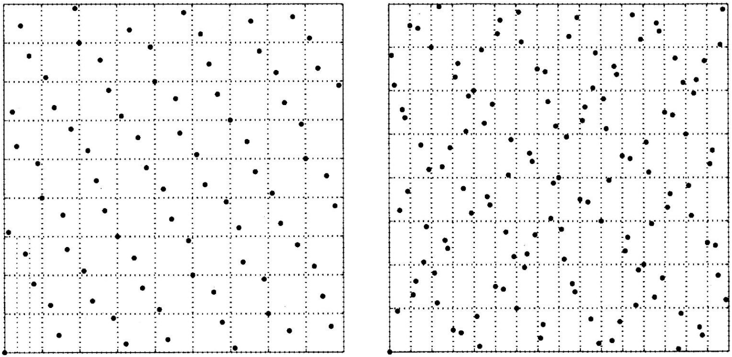


Fig. 5.3. Left panel shows 81 points comprising a $(0, 4, 2)$-net in base 3. Right panel shows 128 points comprising a $(1, 7, 2)$-net in base 2. Both include a point at the origin.

Les (k,m,d)-treillis et les (k,d)-suites VIII

Définition. Soit $k \in \{0, 1, 2, ...\}$. Une suite de points $\{\mathbf{u}_i : i \in \{0, 1, ...\}\}$ dans $\mathcal{I}^d$ est une (k, d) *−suite en base b* si pour tout entier $\kappa \in \{0, 1, 2, ...\}$ et pour tout entier $m \in \{k + 1, k + 2, ...\}$, l'ensemble de points $\left\{\mathbf{u}_{\kappa b^m}, \mathbf{u}_{\kappa b^m+1}, ..., \mathbf{u}_{(\kappa+1)b^m-1}\right\}$ est un (k, m, d) −treillis en base b (c'est-à-dire qu'il y a k points par intervalle élémentaire de volume $1/2^{m-k}$) :

$$\underbrace{\mathbf{u}_0, \mathbf{u}_1 ..., \mathbf{u}_{b^m-1}}_{(k,m,d)-\text{treillis}}\ \underbrace{\mathbf{u}_{b^m}, \mathbf{u}_{b^m+1} ..., \mathbf{u}_{2b^m-1}}_{(k,m,d)-\text{treillis}}\ \underbrace{\mathbf{u}_{2b^m}, \mathbf{u}_{2b^m+1} ..., \mathbf{u}_{3b^m-1}}_{(k,m,d)-\text{treillis}} ...,$$

$$m \in \{k + 1, k + 2, ...\}.$$

Les (k,m,d)-treillis et les (k,d)-suites IX

Plus explicitement,

$$\underbrace{\mathbf{u}_0, \mathbf{u}_1 \ldots, \mathbf{u}_{b^{k+1}-1}}_{(k,k+1,d)-\text{treillis}} \underbrace{\mathbf{u}_{b^{k+1}}, \mathbf{u}_{b^{k+1}+1} \ldots, \mathbf{u}_{2b^{k+1}-1}}_{(k,k+1,d)-\text{treillis}} \underbrace{\mathbf{u}_{2b^{k+1}}, \mathbf{u}_{2b^{k+1}+1} \ldots, \mathbf{u}_{3b^{k+1}-1}}_{(k,k+1,d)-\text{treillis}} \cdots$$

k points par intervalle élémentaire de volume $1/b$

$$\underbrace{\mathbf{u}_0, \mathbf{u}_1 \ldots, \mathbf{u}_{b^{k+2}-1}}_{(k,k+2,d)-\text{treillis}} \underbrace{\mathbf{u}_{b^{k+2}}, \mathbf{u}_{b^{k+2}+1} \ldots, \mathbf{u}_{2b^{k+2}-1}}_{(k,k+2,d)-\text{treillis}} \underbrace{\mathbf{u}_{2b^{k+2}}, \mathbf{u}_{2b^{k+2}+1} \ldots, \mathbf{u}_{3b^{k+2}-1}}_{(k,k+2,d)-\text{treillis}} \cdots$$

k points par intervalle élémentaire de volume $1/b^2$

etc.

Les (k,m,d)-treillis et les (k,d)-suites X

- **Remarque**. Un (k, m, d) $-$treillis en base b est aussi un (K, m, d) $-$treillis pour tout $K \in \{k, k+1, ..., m\}$.
- En effet, notons qu'un (k, m, d) $-$treillis en base b place b^k points dans tous les intervalles élémentaires de volume $1/b^{m-k}$. Notons aussi que, puisque $K \in \{k, k+1, ..., m\}$, un (K, m, d) $-$treillis en base b place plus de points $\left(b^K\right)$ dans tous les intervalles élémentaires de volume plus grand $\left(1/b^{m-K}\right)$.
- Choisissons arbitrairement un de ces "grands" intervalles élémentaires de volume $1/b^{m-K}$. Il est lui-même constitué de "petits" intervalles élémentaires de volume plus petit, c'est-à-dire qu'un intervalle élémentaire de volume $1/b^{m-K}$ peut être représenté comme une union disjointe de b^{K-k} intervalles élémentaires de volume $1/b^{m-k}$.

Les (k,m,d)-treillis et les (k,d)-suites XI

- Un (k, m, d) $-$treillis en base b place b^k points dans chacun des intervalles élémentaires disjoints de volume $1/b^{m-k}$. Il en placera donc $b^{K-k} \times b^k = b^K$ dans le "grand intervalle". $\square$.
(Niederreiter, p. 48)

Les (k,m,d)-treillis et les (k,d)-suites XII

- **Remarque**. Une (k, d) −suite en base b est aussi une (K, d) −suite pour tout $K \in \{k, k+1, ...\}$.

En effet, $\{\mathbf{u}_i : i \in \{0, 1, ...\}\}$ est une (k, d) −suite en base b si

$$\underbrace{\mathbf{u}_0, \mathbf{u}_1 ..., \mathbf{u}_{b^m-1}}_{(k,m,d)-\text{treillis}} \underbrace{\mathbf{u}_{b^m}, \mathbf{u}_{b^m+1} ..., \mathbf{u}_{2b^m-1}}_{(k,m,d)-\text{treillis}} \underbrace{\mathbf{u}_{2b^m}, \mathbf{u}_{2b^m+1} ..., \mathbf{u}_{3b^m-1}}_{(k,m,d)-\text{treillis}} \cdots$$

$$m \in \{k+1, k+2, ...\} .$$

et sera une (K, d) −suite en base b si

$$\underbrace{\mathbf{u}_0, \mathbf{u}_1 ..., \mathbf{u}_{b^m-1}}_{(K,m,d)-\text{treillis}} \underbrace{\mathbf{u}_{b^m}, \mathbf{u}_{b^m+1} ..., \mathbf{u}_{2b^m-1}}_{(K,m,d)-\text{treillis}} \underbrace{\mathbf{u}_{2b^m}, \mathbf{u}_{2b^m+1} ..., \mathbf{u}_{3b^m-1}}_{(K,m,d)-\text{treillis}} \cdots,$$

$$m \in \{K+1, K+2, ...\} .$$

Or, la dernière condition est satisfaite puisqu'un (k, m, d) −treillis en base b est aussi un (K, m, d) −treillis pour tout $K \in \{k, k+1, ..., m\}$. $\square$

Les (k,m,d)-treillis et les (k,d)-suites XIII

- **Conclusion**. Les deux dernières remarques impliquent que plus k est petit, plus les propriétés de régularité sont fortes. (Niederreiter, p. 48)
- **Remarque**. La suite de Van der Corput en base b est une $(0,1)$ $-$suite en base b. (Niederreiter, p. 48)

Bornes pour la *-discrépance des (k,d)-suites

- **Théorème**. Soit P un (k, m, d) $-$treillis en base b. Si $b \in \{3, 4, 5, ...\}$ alors

$$D_N^*(P) \leq \underbrace{\frac{1}{N} b^k \sum_{j=1}^{d-1} \binom{d-1}{j} \binom{m-k}{j} \left\lfloor \frac{b}{2} \right\rfloor^j}_{\kappa(k,m,d)}.$$

  (Niederreiter, p.49, th. 4.5).
- Les pages 49 à 60 de Niederreiter (1992) contiennent une succession de bornes pour la $*-$discrépance des (k, m, d) $-$treillis selon les diverses valeurs de k, m, d et b.

Existence

- Évidemment, tout le monde souhaiterait avoir des $(0, m, d)-$treillis et les $(0, d)-$suites puisque ce sont ces derniers qui offrent le plus de régularité. Or cela n'existe pas toujours. (Niederreiter, p. 60-62)
- **Théorème**. Soit $m \in \{2, 3, 4, ...\}$. Un $(0, m, d)$ $-$treillis en base b peut exister seulement si $d \in \{2, 3, ..., b + 1\}$. (Niederreiter, p. 62, cor. 4.21)
- **Théorème**. Une $(0, d)$ $-$suite en base b peut exister seulement si $d \in \{2, 3, ..., b\}$. (Niederreiter, p. 62, cor. 4.24)

Sobol et Faure

- Les suites de Sobol sont des (k, d) −suites en base 2.
  - L'implémentation est décrite dans Glasserman (2004), section 5.2.3.
  - Selon Galsserman (section 5.5 et 5.6), ce sont les suites de Sobol qui performent le mieux dans les applications financières.
- Celles de Faure sont des $(0, d)$ −suites en base b où b est un nombre premier plus grand ou égal à la dimension d.
  - L'implémentation est décrite dans Glasserman (2004), section 5.2.2.
- Les suites de Faure sont donc "plus régulières" mais sa base est plus grande. Un petite base implique souvent l'utilisation de moins de points pour obtenir l'uniformité (Boyle, Broadie and Glasserman, 1997, p.1295).
  - Rappelons que l'on suggère de prendre un nombre de points correspondant à une puissance de la base

Heuristique

- Comme les bornes pour les erreurs sont généralement beaucoup plus grandes que l'erreur effectivement commise, plusieurs heuristiques ont été développées pour décider quand on a suffisamment de points. (Boyle, Broadie and Glasserman, 1997, p.1295).

Problème de la dimension I

- Les suites à discrépance faible sont souvent très bien uniformément distribuées sur les premières composantes d'un problème en grande dimension..

Problème de la dimension II

- Cependant, plusieurs de ces suites couvrent beaucoup moins bien les dernières composantes.

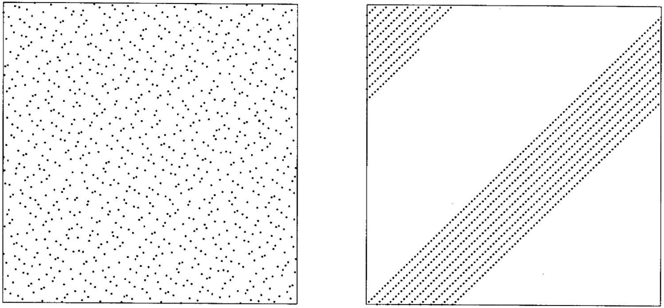


Fig. 5.5. First 1000 points of the Halton sequence in dimension 30. Left panel shows projection onto first two coordinates (bases 2 and 3); right panel shows projection onto last two coordinates (bases 109 and 113).

Problème de la dimension III

- Il devient important de gérer ce problème de dimensionnalité.
  - Par exemple, en utilisant la décomposition en composantes principales de sorte que les facteurs responsables des plus grandes variations soient générés avec les premières coordonnées.
  - Il est aussi possible d'utiliser le pont brownien.

Quasi-Monte-Carlo randomisé

- L'évaluation numérique d'une intégrale basée sur les suites à discrépance faible n'est pas une méthode aléatoire.
  - Par conséquent, il n'y a pas d'intervalle de confiance autour de l'estimation comme cela est le cas pour l'estimateur de Monte Carlo.
  - Au mieux, il y a une borne pour l'erreur mais l'erreur réellement commise peut être beaucoup plus petite.
  - La marge d'erreur associée à l'intervalle de confiance (Monte Carlo) et la borne pour l'erreur (Quasi Monte Carlo) ne sont pas des quantités comparables.
- Afin de pouvoir calculer une marge d'erreur dans le cas Quasi Monte Carlo, il est possible de randomiser les suites à discrépance faible.
- Il existe différents procédés pour randomiser la suite. Nous n'en présentons qu'un seul.

Quasi-Monte-Carlo randomisé

Algorithme

1. Générer les N premiers points d'une suite $\mathcal{S} = \{\mathbf{u}_n : n \in \mathbb{N}\}$ à discrépance faible sur l'hypercube $\mathcal{I}^d$ de dimension d.
2. Générer un variable aléatoire $\mathbf{V}_i$ uniformément distribuée sur $\mathcal{I}^d$
3. Translater la suite $\mathcal{S}$ à l'aide de $\mathbf{V}_i$.
   3.1 $\mathbf{u}_n^{(i)} = \mathbf{u}_n + \mathbf{V}_i$ modulo 1 (afin d'éviter que le nouveau point soit à l'extérieur de l'hypercube unitaire)
4. Évaluer numériquement l'intégrale : $I_i = \frac{1}{N} \sum_{n=1}^{N} f\left(\mathbf{u}_n^{(i)}\right)$.
5. Répéter m fois les étapes 2 à 4 afin d'obtenir un échantillon $I_1, ..., I_m$ constitué de variables aléatoires indépendantes et identiquement distribuées.
   5.1 $\overline{I} = \frac{1}{m} \sum_{i=1}^{n} I_i$ est un estimateur de $\int_{\mathcal{I}^d} f(\mathbf{u})\, d\mathbf{u}$.
   5.2 Parce que l'échantillon est composé de v.a. iid, il est possible d'utiliser les méthodes traditionnelles afin de calculer un intervalle de confiance.

Quasi-Monte-Carlo

Réduction de la variance

Il est possible d'améliorer la précision de nos estimations en utilisant certaines techniques de réduction de la variance. Par exemple, les variables de contrôle peuvent facilement être ajoutées.

Travail pratique I

Le but est de comparer l'efficacité d'une suite à discrépance faible et celle de la simulation de Monte Carlo lors de l'évaluation d'options d'achat. Afin de pouvoir comparer les résultats numériques à des étalons, nous nous plaçons dans le contexte où le titre sous-jacent suit un mouvement brownien géométrique : $S_0 = 100$, $r = 5\%$, $\sigma = 0.3$, $T = 1$, $K \in \{90, 100, 110\}$

1. Déterminez analytiquement les valeurs des options.
2. En utilisant la suite de Van der Corput, évaluez l'option en faisant varier le nombre de points utilisés et la base. Tracez le graphe de l'erreur de tarification (abscisse) et du temps de calcul (ordonnée). Commentez vos résultats.

Travail pratique II

3. Comparez vos résultats à ce que vous obtenez dans le cas de la simulation de Monte Carlo. Discutez des avantages et inconvénients associés aux différentes méthodes. Expliquez quelles sont les principales différences entre les mesures de précision des estimateurs de Monte Carlo et de Quasi Monte Carlo.
4. Est-ce pertinent d'introduire les techniques de réduction de la variance à la simulation de Quasi Monte Carlo? Si oui, donnez un exemple et examinez numériquement les performances. Sinon, expliquez pourquoi.
5. Utilisez le Quasi Monte Carlo randomisé pour estimer les prix d'options. Faites le graphe de la marge d'erreur en fonction du temps de calcul en faisant varier le nombre de points utilisés ainsi que le nombre de trajectoires simulées. Comparez vos résultats à ce que vous obtenez dans le cas de la simulation de Monte Carlo.

Travail pratique III

6. Vaut-il la peine d'introduire une technique de réduction de la variance avec la simulation de Quasi Monte Carlo randomisée? Justifiez votre réponse.