Weierstrass Institute for
Applied Analysis and Stochastics

Leibniz
Association

Computational finance – Lecture 3

Christian Bayer

Mohrenstrasse 39 · 10117 Berlin · Germany · Tel. +49 30 20372 0 · www.wias-berlin.de

April 30, 2021

Outline

1 Variance reduction

2 Quasi Monte Carlo

- $U \sim \mathcal{U}(]0,1[) \Rightarrow \widetilde{U} := 1 - U \sim \mathcal{U}(]0,1[)$, $W \sim \mathcal{N}(0, I_d) \Rightarrow \widetilde{W} := -W \sim \mathcal{N}(0, I_d)$.
  In an intuitive sense, in both cases X is anti-monotonic to $\widetilde{X}$.

Definition (Antithetic variates)

Given a r.v. X and a (deterministic) transformation $\widetilde{X}$ with the same distribution.

$$I_M^A[f; X, \widetilde{X}] := \frac{1}{M} \sum_{i=1}^{M} \frac{f(X_i) + f(\widetilde{X}_i)}{2}.$$

Antithetic variates

- $U \sim \mathcal{U}(]0,1[) \Rightarrow \widetilde{U} := 1 - U \sim \mathcal{U}(]0,1[)$, $W \sim \mathcal{N}(0, I_d) \Rightarrow \widetilde{W} := -W \sim \mathcal{N}(0, I_d)$. In an intuitive sense, in both cases X is anti-monotonic to $\widetilde{X}$.

Definition (Antithetic variates)

Given a r.v. X and a (deterministic) transformation $\widetilde{X}$ with the same distribution.

$$I_M^A[f; X, \widetilde{X}] := \frac{1}{M} \sum_{i=1}^{M} \frac{f(X_i) + f(\widetilde{X}_i)}{2}.$$

- Assume that $\mathrm{cost}(X, \widetilde{X}) \le 2\,\mathrm{cost}(X)$, i.e., cost of simulation of X and transformation to $\widetilde{X}$ is at most twice the cost of simulation of X.
- Then antithetic variates are more efficient than standard MC iff $\mathrm{cov}\left(f(X), f(\widetilde{X})\right) < 0$.

- Suppose we are given a r.v. Y and a function g s.t. $I[g;Y]$ is known. Then, for $\lambda \in \mathbb{R}$,

$$I[f;X] = E\left[f(X) - \lambda g(Y)\right] + \lambda I[g;Y].$$

Definition (Control variates)

$$I_M^{C,\lambda}[f;X,Y] := \frac{1}{M}\sum_{i=1}^{M}\left(f(X_i) - \lambda g(Y_i)\right) + \lambda I[g;Y].$$

- Suppose we are given a r.v. Y and a function g s.t. $I[g;Y]$ is known. Then, for $\lambda \in \mathbb{R}$,

$$I[f;X] = E\left[f(X) - \lambda g(Y)\right] + \lambda I[g;Y].$$

Definition (Control variates)

$$I_M^{C,\lambda}[f;X,Y] := \frac{1}{M}\sum_{i=1}^{M}\left(f(X_i) - \lambda g(Y_i)\right) + \lambda I[g;Y].$$

- Often $Y = X$, and the cost of $I_M^{C,\lambda}$ is only insignificantly higher than the cost of I_M.
- The MSE of $I_M^{C,\lambda}$ is minimized by $\lambda^* := \frac{\operatorname{cov}(f(X),g(Y))}{\operatorname{var}(g(Y))}$, the minimizer of $\operatorname{var}\left(f(X) - \lambda g(Y)\right)$.
- $\operatorname{MSE}\left[I_M^{C,\lambda^*}\right] = (1-\rho^2)\operatorname{MSE}\left[I_M\right]$, where $\rho := \operatorname{cor}\left(f(X), g(Y)\right) \Rightarrow |\rho|$ as large as possible!

- Suppose we are given a r.v. Y and a function g s.t. $I[g;Y]$ is known. Then, for $\lambda \in \mathbb{R}$,

$$I[f;X] = E\left[f(X) - \lambda g(Y)\right] + \lambda I[g;Y].$$

Definition (Control variates)

$$I_M^{C,\lambda}[f;X,Y] := \frac{1}{M}\sum_{i=1}^{M}\left(f(X_i) - \lambda g(Y_i)\right) + \lambda I[g;Y].$$

- Often $Y = X$, and the cost of $I_M^{C,\lambda}$ is only insignificantly higher than the cost of I_M.
- The MSE of $I_M^{C,\lambda}$ is minimized by $\lambda^* := \frac{\mathrm{cov}(f(X),g(Y))}{\mathrm{var}(g(Y))}$, the minimizer of $\mathrm{var}\left(f(X) - \lambda g(Y)\right)$.
- $\mathrm{MSE}\left[I_M^{C,\lambda^*}\right] = (1-\rho^2)\,\mathrm{MSE}\left[I_M\right]$, where $\rho := \mathrm{cor}\left(f(X), g(Y)\right) \Rightarrow |\rho|$ as large as possible!

Example

$E[S_T] = S_0$, hence $Y = S_T$, $g = \mathrm{id}$ can be used as control variate for option pricing.

- Black – Scholes model ($r = 0$ as usual): $S_t = S_0 \exp\left(\sigma W_t - \frac{1}{2}\sigma^2 t\right)$.
- Asian option: option depending on the time average of the stock price, e.g.,

$$f(X) := \left(\frac{1}{N}\sum_{n=1}^{N} X_n - K\right)^+, \quad X := (S_{t_1}, \dots, S_{t_N}), \quad 0 \le t_1 < \cdots < t_N \le T.$$

- Black – Scholes model ($r = 0$ as usual): $S_t = S_0 \exp\left(\sigma W_t - \frac{1}{2}\sigma^2 t\right)$.
- Asian option: option depending on the time average of the stock price, e.g.,

$$f(X) := \left(\frac{1}{N}\sum_{n=1}^{N} X_n - K\right)^+, \quad X := (S_{t_1}, \ldots, S_{t_N}), \quad 0 \leq t_1 < \cdots < t_N \leq T.$$

- No explicit price formula as sum of log-normal r.v.s are not log-normal.

- Black – Scholes model ($r = 0$ as usual): $S_t = S_0 \exp\left(\sigma W_t - \frac{1}{2}\sigma^2 t\right)$.
- Asian option: option depending on the time average of the stock price, e.g.,

$$f(X) := \left(\frac{1}{N}\sum_{n=1}^{N} X_n - K\right)^+, \quad X := (S_{t_1}, \ldots, S_{t_N}), \quad 0 \le t_1 < \cdots < t_N \le T.$$

- No explicit price formula as sum of log-normal r.v.s are not log-normal.
- MC simulation: For $Z \sim \mathcal{N}(0, I_N)$, set $W_{t_1} = \sqrt{t_1} Z_1$ and $W_{t_n} = W_{t_{n-1}} + \sqrt{t_n - t_{n-1}} Z_n$.

- Black – Scholes model ($r = 0$ as usual): $S_t = S_0 \exp\left(\sigma W_t - \frac{1}{2}\sigma^2 t\right)$.
- Asian option: option depending on the time average of the stock price, e.g.,

$$f(X) := \left(\frac{1}{N}\sum_{n=1}^{N} X_n - K\right)^+, \quad X := (S_{t_1}, \dots, S_{t_N}), \quad 0 \le t_1 < \dots < t_N \le T.$$

- No explicit price formula as sum of log-normal r.v.s are not log-normal.
- MC simulation: For $Z \sim \mathcal{N}(0, I_N)$, set $W_{t_1} = \sqrt{t_1} Z_1$ and $W_{t_n} = W_{t_{n-1}} + \sqrt{t_n - t_{n-1}} Z_n$.
- Observation: $\frac{1}{N}\sum_{n=1}^{N} x_n \approx \left(\prod_{n=1}^{N} x_n\right)^{1/N}$ (arithmetic vs. geometric mean), and

$$\left(\prod_{n=1}^{N} S_{t_n}\right)^{1/N} = S_0 \exp\left(\frac{\sigma}{N}\sum_{n=1}^{N} W_{t_n} - \frac{\sigma^2}{2N}\sum_{n=1}^{N} t_n\right) \sim \mathrm{LN}\left(\log S_0 - \frac{\sigma^2}{2N}\sum_{n=1}^{N} t_n, \ \frac{\sigma^2}{N^2}\sum_{n=1}^{N}(1 + 2(N-n))\, t_n\right).$$

- Black – Scholes model ($r = 0$ as usual): $S_t = S_0 \exp\left(\sigma W_t - \frac{1}{2}\sigma^2 t\right)$.
- Asian option: option depending on the time average of the stock price, e.g.,

$$f(X) := \left(\frac{1}{N}\sum_{n=1}^{N} X_n - K\right)^+, \quad X := (S_{t_1}, \dots, S_{t_N}), \quad 0 \le t_1 < \cdots < t_N \le T.$$

- No explicit price formula as sum of log-normal r.v.s are not log-normal.
- MC simulation: For $Z \sim \mathcal{N}(0, I_N)$, set $W_{t_1} = \sqrt{t_1} Z_1$ and $W_{t_n} = W_{t_{n-1}} + \sqrt{t_n - t_{n-1}} Z_n$.
- Observation: $\frac{1}{N}\sum_{n=1}^{N} x_n \approx \left(\prod_{n=1}^{N} x_n\right)^{1/N}$ (arithmetic vs. geometric mean), and

$$\left(\prod_{n=1}^{N} S_{t_n}\right)^{1/N} = S_0 \exp\left(\frac{\sigma}{N}\sum_{n=1}^{N} W_{t_n} - \frac{\sigma^2}{2N}\sum_{n=1}^{N} t_n\right) \sim \mathrm{LN}\left(\log S_0 - \frac{\sigma^2}{2N}\sum_{n=1}^{N} t_n, \ \frac{\sigma^2}{N^2}\sum_{n=1}^{N} (1 + 2(N-n))\, t_n\right).$$

- By the Black-Scholes formula, a control variate is given by the geometric Asian option

$$g(Y) := \left(\left(\prod_{n=1}^{N} Y_n\right)^{1/N} - K\right)^+, \quad Y := X = (S_{t_1}, \dots, S_{t_N}).$$

- Idea: Sample more where the “local variance” is higher.
- Setting: X has (d-dimensional) density p, additional r.v. Y with density q.

- Idea: Sample more where the “local variance” is higher.
- Setting: X has (d-dimensional) density p, additional r.v. Y with density q.

Definition (Importance sampling)

$$I_M^{IS}[f; X, Y] := \frac{1}{M} \sum_{i=1}^{M} f(Y_i) \frac{p(Y_i)}{q(Y_i)} = I_M\left[f\frac{p}{q}; Y \right].$$

- Idea: Sample more where the "local variance" is higher.
- Setting: X has (d-dimensional) density p, additional r.v. Y with density q.

Definition (Importance sampling)

$$I_M^{IS}[f;X,Y] := \frac{1}{M}\sum_{i=1}^{M} f(Y_i)\frac{p(Y_i)}{q(Y_i)} = I_M\left[f\frac{p}{q};Y\right].$$

- Mean squared error determined by

$$\operatorname{var}\left(f(Y)\frac{p(Y)}{q(Y)}\right) = E\left[f(Y)^2\frac{p(Y)^2}{q(Y)^2}\right] - I[f;X]^2 = E\left[f(X)^2\frac{p(X)}{q(X)}\right] - I[f;X]^2.$$

- Idea: Sample more where the “local variance” is higher.
- Setting: X has (d-dimensional) density p, additional r.v. Y with density q.

Definition (Importance sampling)

$$I_M^{IS}[f; X, Y] := \frac{1}{M} \sum_{i=1}^{M} f(Y_i) \frac{p(Y_i)}{q(Y_i)} = I_M\left[f \frac{p}{q}; Y \right].$$

- Mean squared error determined by

$$\operatorname{var}\left(f(Y) \frac{p(Y)}{q(Y)} \right) = E\left[f(Y)^2 \frac{p(Y)^2}{q(Y)^2} \right] - I[f; X]^2 = E\left[f(X)^2 \frac{p(X)}{q(X)} \right] - I[f; X]^2.$$

- Best possible speed-up: q proportional to $f \cdot p$, implying $f(Y)\frac{p(Y)}{q(Y)} \equiv 1$ – but requires knowledge of $I[f; X]$ for normalization.

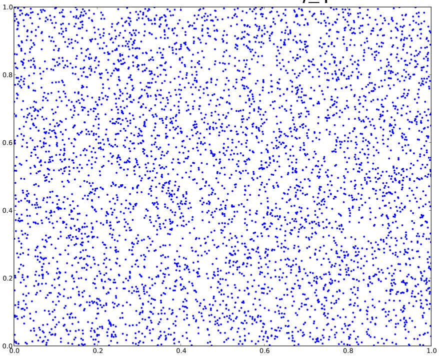
Outline

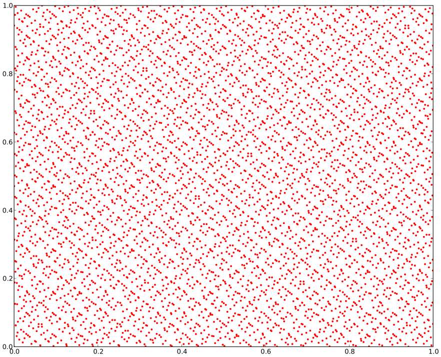
1 Variance reduction

2 Quasi Monte Carlo

$$I[f] := \int_{[0,1]^d} f(x)dx$$

- Quadrature: $J_M[f] := \frac{1}{M}\sum_{i=1}^{M} f(x_i),\ x_1, \ldots, x_M \in [0,1]^d.$





Discrepancy

- Intuition: discrepancy is the quadrature error for indocator functions of rectangles R.
- Consider point sets $(x_i)_{i=1}^{M} \subset (x_i)_{i \in \mathbb{N}}$ or $(x_i^M)_{i=1}^{M}$, $x_i \in [0,1]^d$.

- Intuition: discrepancy is the quadrature error for indocator functions of rectangles R.
- Consider point sets $(x_i)_{i=1}^M \subset (x_i)_{i\in\mathbb{N}}$ or $(x_i^M)_{i=1}^M$, $x_i \in [0,1]^d$.

Definition (Discrepancy)

The discrepancy D_M and star-discrepancy D_M^* for a point set $(x_i)_{i=1}^M$ are defined as

$$D_M = \sup_{\text{rectangles } R\subset[0,1]^d} \left| \frac{1}{M} \# \{ 1 \le i \le M : x_i \in R \} - \lambda(R) \right|,$$

$$D_M^* = \sup\left\{ \left| \frac{1}{M} \# \{ 1 \le i \le M : x_i \in R \} - \lambda(R) \right| : R = \bigtimes_{j=1}^d [0, b_j[, \ b_1, \dots, b_d \in [0,1] \right\}.$$

- Intuition: discrepancy is the quadrature error for indocator functions of rectangles R.
- Consider point sets $(x_i)_{i=1}^M \subset (x_i)_{i\in\mathbb{N}}$ or $(x_i^M)_{i=1}^M$, $x_i \in [0,1]^d$.

Definition (Discrepancy)

The discrepancy D_M and star-discrepancy D_M^* for a point set $(x_i)_{i=1}^M$ are defined as

$$D_M = \sup_{\text{rectangles } R\subset[0,1]^d} \left| \frac{1}{M} \#\{\, 1 \le i \le M : x_i \in R \,\} - \lambda(R) \right|,$$

$$D_M^* = \sup\left\{ \left| \frac{1}{M} \#\{\, 1 \le i \le M : x_i \in R \,\} - \lambda(R) \right| \;:\; R = \bigtimes_{j=1}^{d} [0, b_j[,\; b_1, \dots, b_d \in [0,1] \right\}.$$

Definition (Low discrepancy)

A sequence of point sets $(x_i^M)_{i=1}^M$, $M \in \mathbb{N}$, has low discrepancy iff $D_M^* \le c \frac{\log(M)^d}{M}$.

- Regularity of integrand f classically measured by variation.

Definition (Variation in the sense of Hardy-Krause)

For a one-dimensional function $f : [0, 1] \to \mathbb{R}$

$$V[f] := \int_0^1 \left| \frac{df}{dx}(x) \right| dx$$

and for $f : [0, 1]^d \to \mathbb{R}$

$$V[f] := \int_{[0,1]^d} \left| \frac{\partial^d f}{\partial x^1 \cdots \partial x^d}(x) \right| dx + \sum_{j=1}^{d} V[f_1^{(j)}],$$

where $f_1^{(j)}$ denotes the restriction of f to the boundary $x^j = 1$.

Theorem (Koksma-Hlawka inequality)

$$|I[f] - J_M[f]| \leq V[f] D_M^*.$$

Theorem (Koksma-Hlawka inequality)

$$|I[f] - J_M[f]| \leq V[f] D_M^*.$$

- Error factorization into term depending on regularity of integrand and uniformity of point set.
- In contrast to MC error, the Koksma-Hlawka inequality is deterministic, but usually far from sharp! (Note that it essentially require f to be C^d!)

Theorem (Koksma-Hlawka inequality)

$$|I[f] - J_M[f]| \le V[f] D_M^*.$$

- Error factorization into term depending on regularity of integrand and uniformity of point set.
- In contrast to MC error, the Koksma-Hlawka inequality is deterministic, but usually far from sharp! (Note that it essentially require f to be C^d!)

Example (Van der Corput sequence)

Let p prime and for $k \in \mathbb{N}_0$ let $(a_j(k))_{j\in\mathbb{N}_0}$ denote the p-ary expansion of k, i.e., $k = \sum_{j=0}^{\infty} a_j(k) p^j$. A one-dimensional low-discrepancy sequence is given by

$$x_i \coloneqq \psi_p(i) \coloneqq \sum_{j=0}^{\infty} \frac{a_j(i)}{p^{j+1}}, \quad i \in \mathbb{N}_0.$$

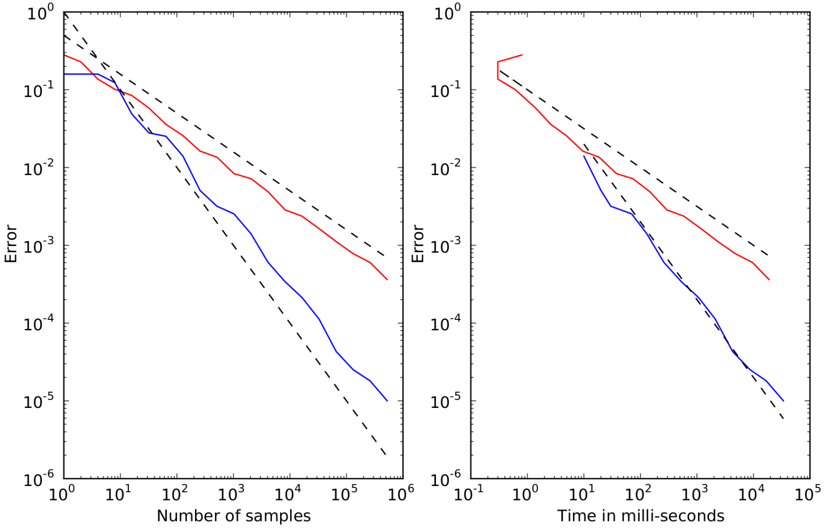


Figure: A call option in the Black-Scholes model using Monte Carlo and Quasi Monte Carlo simulation. Red: MC simulation, blue: QMC simulation, black: Reference lines proportional to $1/M$ and $1/\sqrt{M}$.

- For uniform (pseudo) r.v.s: If $U_i \sim \mathcal{U}([0,1])$, (U_i) independent, then $(U_{(i-1)d+1}, \ldots, U_{id})_{i\in\mathbb{N}}$ is a sequence of independent r.v.s from $\mathcal{U}([0,1]^d)$.
- In contrast, if $(x_i)_{i\in\mathbb{N}}$ is one-dimensional low-discrepancy sequence, then $(x_{(i-1)d+1}, \ldots, x_{id})_{i\in\mathbb{N}}$ is most likely not a d-dimensional low-discrepancy sequence.
- QMC convergence is asymptotically much faster than MC convergence, but note that $\log(M)^d/M \gg M^{-1/2}$ for all reasonably sized M even in fairly moderate dimensions d. E.g., for $d = 8$ roughly from $M \geq 1.8 \times 10^{29}$.

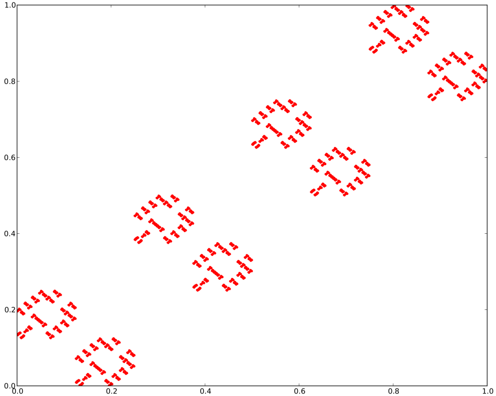


Figure: Pairs of one-dimensional Sobol numbers

Setting: $(x_i)_{i=1}^{M}$ low-discrepancy sequence in dim. d, $(U_l)_{l\in\mathbb{N}}$ i.i.d. sequence, $U_l \sim \mathcal{U}\left([0,1]^d\right)$.

Definition (Randomized quasi Monte Carlo simulation)

$$J^R_{M;m}[f] := \frac{1}{m}\sum_{l=1}^{m}\frac{1}{M}\sum_{i=1}^{M} f(x_i + U_l \pmod 1)$$

Setting: $(x_i)_{i=1}^M$ low-discrepancy sequence in dim. d, $(U_l)_{l\in\mathbb{N}}$ i.i.d. sequence, $U_l \sim \mathcal{U}\left([0,1]^d\right)$.

Definition (Randomized quasi Monte Carlo simulation)

$$J^R_{M;m}[f] \coloneqq \frac{1}{m}\sum_{l=1}^m \frac{1}{M}\sum_{i=1}^M f(x_i + U_l \pmod 1)$$

- Provides variance reduction for the Monte Carlo estimator $\frac{1}{m}\sum_{l=1}^m f(U_l)$
- Provides sharp, computable error control for the quasi Monte Carlo approximation $\frac{1}{M}\sum_{i=1}^M f(x_i)$:

$$E\left[\left(I[f] - J^R_{M;m}[f]\right)^2\right] = \frac{\operatorname{var}\left(\frac{1}{M}\sum_{i=1}^M f(x_i + U \pmod 1)\right)}{m}$$

- Typically $m \ll M$, as accuracy is more important than error control.