Quasi MC method

Note to approxi. $I(f) := \int_{[0,1]^d} f(x)dx$. We use $X_k \overset{i.i.d.}{\sim} U[0,1]^d$ and set $J_m(f) = m^{-1}\sum_1^m f(x_i)$ to apply MC method.

But the idea of QMC is to replace (X_k) by seq of deterministic number $(x_i) \subset [0,1]^d$ which distributes "evenly" to mimic $X_k \overset{i.i.d.}{\sim} [0,1]^d$.

(1) <u>Definition:</u>

<u>Def</u>: $\lambda = \mathcal{L}|_{[0,1]^d}$ restriction of Lebesgue measure on $[0,1]^d$. $R = \bigotimes_1^d [a_i, b_i)$ is rectangle subset in $[0,1]^d$.

i) Discrepancy D_m of $(x_i)_1^m$ is:

$$D_m := \sup_R |m^{-1} \#\{x_i \in R.\ 1 \le i \le m\} - \lambda(R)|$$

ii) $D_m^* := \sup\{ |m^{-1} \#\{x_i \in R.\ 1 \le i \le m\} - \lambda(R)| :$

$R = \bigotimes_1 [0, b_i) \}$.

iii) Hardy - Krause variation is recursively defined by: for $f: [0,1]^d \to \mathbb{R}^1$.

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

where $f_1^{(j)} = f|_{x_j = 1}$

And $V(g) = \int_0^1 |\partial g/\partial x| \, dx$. $g: [0,1] \to \mathbb{R}^1$.

Rmk: Note that J_m will depend on (X_k) and regularity of f.

$\Rightarrow$ we will require more on f.

Thm. (Koksma - Hlawka inequi.)

$\forall f \in L^1([0,1]^d; \mathbb{R}^1) \cap C^{(1,\cdots,1)}$. We have:

$|I(f) - J_m(f)| \leq V(f) D_m^*$

Rmk: i) It's a deterministic bound.

ii) Even we need regularity on f. it still works better practically.

Pf: For $d=1$. $f \in C^1$. Note that

$f(x) = f(1) - \int_0^1 f'(t) I_{(x,1]}(t) dt$.

So: $|I(f) - I_m(f)| =$

$$\left|\int_0^1 f'(t)\left(\frac{1}{m}\sum_1^m I_{(x_i,1]}(t)-\int_0^1 I_{(x,1]}(t)\,dx\right)dt\right|$$

$$\le \int_0^1 |f'(t)|\,\left|m^{-1}\sum_1^m I_{[0,t)}(x_i)-\int_0^1 I_{[0,t)}(x)\right|$$

$$\le V(f)\,D_m^*.$$

For d-dim :

Note $f(x_1\cdots x_d)=\int_0\cdots\int_0^1 \frac{\partial^d}{\partial x_1\cdots\partial x_d}f(t_1\cdots t_d)\,I_{[0,t_1]\times\cdots[0,t_d]}(x_1\cdots x_d)+\square\,dt_1\cdots dt_d.$

$\Rightarrow |I(f)-J_m(f)|\le V(f)\,D_m^*$

<u>Def</u>: $(x_i)_{\mathbb{N}}\subset[0,1]^d$ is low discrepancy if $D_m^*\le C(\log m)^d/m$.

<u>Rmk</u>: (x_i) is 1-dim low discrepancy $\not\Rightarrow$ $(x_{(i-1)d+1},\cdots x_{id})$ is also d-dim low discrepancy. (Although it's true as for "uniform")

<u>e.g.</u> $\psi_p(k)=\sum_{j=0} a_j(k)/p^{j+1}$ for $k=\sum_0^r a_j(k)\cdot p^j$ (p-adic expansion)

$x_i=\psi_p(i)$, $i\in\mathbb{N}$. Van der Corput seq. is low discrepancy.

Rmk: Halton seq. $X_i = (X_i^1, \cdots X_i^k)$. s.t. $x_i^j = \varphi_{p_j}(i)$. (P_j's are primes) is k-dim generalition.

Low dim.:

level of even dist. deteriorates in the dimension. e.g., First two coordinate (X_i^1, X_i^2) has better uniformity than last two coordinate (X_i^{k-1}, X_i^k).

And we require more regularity on f when dim $\uparrow$.

Rmk: Usually we expect $f(x_1, x_2 \cdots x_k)$ can be decomp. in: $\sum f^{(i_1 \cdots i_k)}(X^{i_1}, \cdots X^{i_k})$. So: the accuracy can be improved.

(2) Renormalized QMC:

Note that QMC has faster converge ($O(n^{1-\varepsilon})$) than MC ($O(n^{\frac{1}{2}})$). But it lacks of good error control.

Next, we see we can actually combine QMC and MC.

For (X_i) d-dim low discrepancy. U_ℓ i.i.d. $U[0,1]^d$ r.v. Set $X_\ell := (X_i + U_\ell \pmod 1)$

$J^k_{M,m}(f) = \frac{1}{m}\sum_1^m J_M(f, X_\ell)$. mean of QMC

Rmk: m is to compute the error estimate which bases on $Var(J_M(f, X_i))$. And M controls error of it.

(3) Pricing America option:

For f payoff func. We want to know NA price: $\sup\{\mathbb{E}(e^{-\tau r} f(X_\tau)) \mid \tau(\omega) \le T\}$.

But in reality, we can't exercise the option continuously but only on finite times: $t_1, \cdots t_m$.

Rmk: It's called Bermudan options.

Next, we assume:

i) interest rate $r = 0$.

ii) Under risk-neutral p.m. $\mathbb{P}$.

iii) The option can only be exercised at times $0 = t_0 < t_1 < t_2 \cdots < t_m = T$.

iv) Stock price $X_i := X_{t_i}$ is a $\tilde{\mathbb{P}}$-Markov Chain (dist. is given)

① Dynamic programming:

Denote $f_i(x)$ is payoff func. of option at time i given $X_i = x$ and $V_i(x)$ is value of option at time i given $X_i = x$

Rmk: We have $f_m(x) = f(x) =$

$$\mathbb{E}(e^{-r(m-m)} f(X_m) \mid X_m = x) = V_m(x)$$

$$(V_i(x) = \mathbb{E}(e^{-r(m-i)} f(X_m) \mid X_i = x))$$

By Bellman's equation: for $i < m$,

$V_i(x) = \max\{f_i(x),$ expected value at next step if keeping holding$\}$.

$$= \max\{f_i(x), \mathbb{E}(V_{i+1}(X_{i+1}) \mid X_i = x)\}.$$

$\Rightarrow$ Recursively, we can determine the price $V_0(x)$ of the option.

Denote $C_i(x) = \mathbb{E}(V_{i+1}(X_{i+1}) \mid X_i = x)$ continuation value which should be calculated at each time i. $\Rightarrow V_i(x) = f_i(x) \vee C_i(x)$

Set $\tau^* = \min\{i \in \{1, \cdots m\} \mid f_i(X_i) \geq V_i(X_i)\}$.

<u>Lem.</u> τ^* is the optimal stopping time. i.e.

$$V_0^{(\tau^*)}(X_0) = \mathbb{E}(f_{\tau^*}(X_{\tau^*})) = \sup_{\tau \in \mathcal{T}} \mathbb{E}(f_\tau(X_\tau)) = V_0(X_0)$$

<u>Rmk</u>: Replace $V_i(X_i)$ by $C_i(X_i)$ on τ^* above. It still works.

<u>Pf</u>: Note $(V_i(X_i))_{i \leq \tau^*}$ is mart.

since $V_i(X_i) > f_i(X_i)$, $\forall i < \tau^*$.

$$\Rightarrow \text{LHS} \overset{\text{def}}{=} \mathbb{E}(V_{\tau^*}(X_{\tau^*}))$$
$$= \mathbb{E}(\mathbb{E}(V_{\tau^*}(X_{\tau^*}) \mid X_{\tau^*-1}))$$
$$= \mathbb{E}(V_{\tau^*-1}(X_{\tau^*-1}))$$
$$= \cdots = V_0(X_0).$$

Assume we're given estimate price $\bar{V}_i(x)$

of $V_i(x)$ and define $\bar{\tau}$ as above.

So: $V_0^{(\bar{\tau})}(X_0) \leq V_0(X_0)$. ($\bar{\tau}$ may not optimal)

Let $\tilde{V}_i(x)$ is obtained by:

$\tilde{V}_m(x) = f_m(x)$.

$\tilde{V}_{i-1}(x) := \max\{f_{i-1}(x), I(V_i(X_i) \mid X_i = x)\}$

where $I(Y \mid X)$ is unbiased estimator st.

$\mathbb{E}(I(Y \mid X)) = \mathbb{E}(\mathbb{E}(Y \mid X)) = \mathbb{E}(Y)$

<u>Rmk</u>: $\tilde{V}_i(x)$ is r.v. unless I is truly conditional expectation.

<u>Lem.</u> $(\tilde{V}_i(x))$ bias higher. i.e. $i = 0, \cdots m$.

$$\mathbb{E}(\tilde{V}_i(X_i) \mid X_i) \geq V_i(X_i).$$

<u>Pf</u>: By backward induction, note that it holds if $i = m$.

$\mathbb{E}(\tilde{V}_{i-1}(X_{i-1}) \mid X_{i-1}) =$

$\mathbb{E}(\max\{f_{i-1}(X_{i-1}), I(\tilde{V}_i(X_i) \mid X_i)\} \mid X_{i-1})$

$\overset{\text{Jensen}}{\geq} \max\{f_{i-1}(X_{i-1}), \mathbb{E}(\mathbb{E}(\tilde{V}_i(X_i) \mid X_i))\}$

$= \max\{f_{i-1}(X_{i-1}), \mathbb{E}(\mathbb{E}(\tilde{V}_i(X_i) \mid X_i) \mid X_{i-1})\}$

$\overset{\text{hypo.}}{\geq} \max\{f_{i-1}(X_{i-1}), \mathbb{E}(V_i(X_i) \mid X_{i-1})\} = V_{i-1}(X_{i-1})$

Rmk: So using $\hat{V}_i(x)$ to construct τ implies $V_i^{(\tau)}$ will bias low (since τ isn't optimal!)

$\Rightarrow$ We use the estimate price $\in [V_0^{(\tau)}, \overline{V}_0]$.

③ Random tree:

Next, we want to approxi. the condition expectations in Bellman's equation by MC simulation. It leads to a random tree with b branchings:

i) Sample $(X_1^i)_{i=1,\dots b} \overset{\text{i.i.d.}}{\sim} X_1$

2) Sample $(X_2^{ij})_{j=1,\dots b} \overset{\text{i.i.d.}}{\sim} X_2 \mid X_1 = X_1^i$. $i=1,\dots b$.

3) Repeat above. We have $(X_m^{j_1\dots j_m})_{m,j_1,\dots j_m}$.

4) Set $V_m^{j_1\dots j_m} := f_m(X_m^{j_1\dots j_m})$ and recursively $V_i^{j_1\dots j_i} := \max\{f_i(X_i^{j_1\dots j_i}), \frac{1}{b}\sum_{j=1}^{b} V_{i+1}^{j_1\dots j_i j}\}$

$\Rightarrow$ We obtain $\overline{V}_0(X_0) = V_0^R$.

Rmk: i) We see now: this estimate also biased high by Lem. above is because it does depend on the future realizations ($\frac{1}{b}\sum_{j=1}^{b} V_{i+1}$-term) since $(V_i(X_i))$ is actually supermart. and we define $\bar{V}$ by estimating the conditional expectation

ii) Low-biased estimator can be constructed by:

$$V_{i,k}^{j_1\cdots j_i} = \begin{cases} f_i(X_i^{j_1\cdots j_i}), & \frac{1}{b-1}\sum_{j\neq k}^{b} V_{i+1}^{j_1\cdots j_i j} \le f_i(X_i^{j_1\cdots j_i}) \\ V_{i+1}^{j_1\cdots j_i k}, & \text{else.} \end{cases}$$

$$\Rightarrow V_i^{j_1\cdots j_i} := \frac{1}{b}\sum_{k=1}^{b} V_{i,k}^{j_1\cdots j_i}$$

We can show V_0^{λ} is biased low.

iii) $V_0^{\lambda} \to V_0(X_0)$ $(b\to\infty)$

② Pricing by regression:

Given basis func. $\psi_j: \mathbb{R}^n \to \mathbb{R}^1$ and unknown coeff. (β_{ij}). $\psi(x) = (\psi_1(x)\cdots\psi_r(x))^T$. $\beta_i = (\beta_{i1}\cdots\beta_{ir})$

Ansatz: $\mathbb{E}(V_{i+1}(X_{i+1}) \mid X_i = x) = \sum_{j=1}^{n} \beta_{ij} \psi_j(x)$

$\Rightarrow \mathbb{E}(\psi(X_i) V_{i+1}(X_{i+1})) \overset{\text{cond.}}{=} \mathbb{E}(\psi(X_i)\beta_i^T \psi(X_i))$

$= \mathbb{E}(\psi(X_i)\psi^T(X_i))\beta_i$

So: $\beta_i \overset{\Delta}{=} M_\psi^{-1}(X_i) M_{\psi V}(X_i)$. if $|M_\psi| \neq 0$.

Algorithm 2.40. *Simulate b independent paths of the Markov chain $X_1, \ldots, X_m$ starting from $X_0 = x$. We denote the simulated values by $X_i^{(j)}$, $i = 1, \ldots, m$, $j = 1, \ldots, b$. Set $V_{m,j} := f_m(X_m^{(j)})$ and proceed for $i = m-1, \ldots, 0$ (backwards in time):*

(i) Compute matrices $\hat{M}_\psi$ and $\hat{M}_{\psi V}$ by

$$(\hat{M}_\psi)_{l,k} := \frac{1}{b}\sum_{j=1}^{b} \psi_l(X_i^{(j)})\psi_k(X_i^{(j)}), \quad (\hat{M}_{\psi V})_k := \frac{1}{b}\sum_{j=1}^{b} \psi_k(X_i^{(j)}) V_{i+1,j}.$$

(ii) Set the regression coefficient $\hat{\beta}_i := \hat{M}_\psi^{-1}\hat{M}_{\psi V}$.

(iii) Obtain the new option price estimates by

$$V_{i,j} := \max\left(f_i(X_i^{(j)}), \hat{\beta}_i^T \psi(X_i^{(j)})\right). \tag{2.24}$$

Rmk: i) $V_0 = \frac{1}{b}\sum_{j}^{b} V_{0,j} \to V_0(X_0) \quad (b \to \infty)$

ii) It's less restrictive on memory and speed (mb v.s. m^b). And all $(X_i^{(j)})_{j=1\cdots b}$ are used to compute β_i. (RTM only use samples from each nodes

iii) $(\psi_i)_1^n$ generally choose from L^2-basis.