Importance Sampling and MM-Algorithms with Applications to Options Pricing

Thorsten Sauder

Bankgesellschaft Berlin

Introduction

- Aim of the talk

  We present a method that improves the calculation of expectation values $E(X)$ by Monte Carlo simulation.

Introduction

- Aim of the talk

  We present a method that improves the calculation of expectation values $E(X)$ by Monte Carlo simulation.
- Let $(X_i)_{i\geq 1}$ be i.i.d with $E(X_i) = \mu$ and $Var(X_i) = \sigma^2$.

Introduction

- Aim of the talk

  We present a method that improves the calculation of expectation values $E(X)$ by Monte Carlo simulation.
- Let $(X_i)_{i\geq 1}$ be i.i.d with $E(X_i) = \mu$ and $Var(X_i) = \sigma^2$.
- Strong Law of Large Numbers

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} X_i = \mu \quad a.s.$$

Introduction

- Aim of the talk

  We present a method that improves the calculation of expectation values $E(X)$ by Monte Carlo simulation.
- Let $(X_i)_{i\geq 1}$ be i.i.d with $E(X_i) = \mu$ and $Var(X_i) = \sigma^2$.
- Strong Law of Large Numbers

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} X_i = \mu \quad a.s.$$

- Central Limit Theorem

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right) \xrightarrow{\mathcal{D}} N(0,\sigma)$$

Basic Idea

- Find a non-trivial family $(Z_\gamma)_{\gamma\in\Gamma}$ of random variables with

$$X = Z_{\gamma_0} \quad \text{and} \quad E(Z_\gamma) = E(X) \text{ for all } \gamma \in \Gamma,$$

Basic Idea

- Find a non-trivial family $(Z_\gamma)_{\gamma\in\Gamma}$ of random variables with

$$X = Z_{\gamma_0} \quad \text{and} \quad E(Z_\gamma) = E(X) \text{ for all } \gamma \in \Gamma,$$

- so that one can solve the minimization problem

$$\bar\gamma \in \arg\min \{Var(Z_\gamma) \,:\, \gamma \in \Gamma\}\,.$$

It follows

$$Var(Z_{\bar\gamma}) \le Var(X)\,.$$

Then use i.i.d random variables that have the same distribuition as $Z_{\bar\gamma}$ and estimate $E(X)$ by Monte Carlo simulation.

Notation

- $\left(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t\in[0,T]}\right)$ is a filtered probability space

Notation

- $\left(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t\in[0,T]}\right)$ is a filtered probability space
- $(W(t))_{t\in[0,T]}$ is a d-dimensional Brownian motion

Notation

- $\left(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t\in[0,T]}\right)$ is a filtered probability space
- $(W(t))_{t\in[0,T]}$ is a d-dimensional Brownian motion
- $L^2(\mathcal{F}_T, P)$ is the Hilbert space of $\mathcal{F}_T$-measureable square-integrable real-valued random variables

Notation

- $\left(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t\in[0,T]}\right)$ is a filtered probability space
- $(W(t))_{t\in[0,T]}$ is a d-dimensional Brownian motion
- $L^2(\mathcal{F}_T, P)$ is the Hilbert space of $\mathcal{F}_T$-measureable square-integrable real-valued random variables
- $L^2(\mathcal{B}_T, \lambda)$ is the Hilbert space of $\mathcal{B}_T$-measureable square-integrable $\mathbf{R}^d$-valued functions with respect to the Lebesque measure λ

Notation

- $\left(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t\in[0,T]}\right)$ is a filtered probability space
- $(W(t))_{t\in[0,T]}$ is a d-dimensional Brownian motion
- $L^2(\mathcal{F}_T, P)$ is the Hilbert space of $\mathcal{F}_T$-measureable square-integrable real-valued random variables
- $L^2(\mathcal{B}_T, \lambda)$ is the Hilbert space of $\mathcal{B}_T$-measureable square-integrable $\mathbf{R}^d$-valued functions with respect to the Lebesque measure λ
- for $X, Y \in L^2(\mathcal{F}_T, P)$ we set

$$(X, Y)_P = E_P(XY) \quad \text{and} \quad \|X\|_P = \sqrt{E_P(X^2)}$$

Notation

- for $\gamma, \eta \in L^2(\mathcal{B}_T, \lambda)$ we set

$$(\gamma\,,\eta)_\lambda = \int_0^T (\gamma(t)\;,\eta(t))\,dt \quad \text{and} \quad \|\gamma\|_\lambda = \sqrt{\int_0^T \|\gamma(t)\|^2 dt}$$

Notation

- for $\gamma, \eta \in L^2(\mathcal{B}_T, \lambda)$ we set

$$(\gamma\,,\eta)_\lambda = \int_0^T (\gamma(t)\ ,\eta(t))\,dt \quad \text{and} \quad \|\gamma\|_\lambda = \sqrt{\int_0^T \|\gamma(t)\|^2 dt}$$

- for $\gamma \in L^2(\mathcal{B}_T, \lambda)$ we set

$$\mathcal{E}(W(\gamma)) = \exp\left(-\frac{1}{2}\|\gamma\|_\lambda^2 + W(\gamma)\right)$$

with

$$W(\gamma) = \int_0^T \gamma(u)\,dW(u)$$

Girsanov-Transformation

For each $\gamma \in L^2(\mathcal{B}_T, \lambda)$ a probability measure P^γ is defined by

$$\frac{dP^\gamma}{dP} = \mathcal{E}(W(\gamma))$$

that is equivalent to P. The process W^γ defined by

$$W^\gamma(t) = W(t) - \int_0^t \gamma(u)\, du$$

is a P^γ-Brownian motion.

Theorem 1

For each $\gamma \in L^2(\mathcal{B}_T, \lambda)$, $p > 2$, there is an isometry

$$G^\gamma : L^p(\mathcal{F}_T, P) \to L^p(\mathcal{F}_T, P^\gamma)$$

such that for each $X = F(W(t_1), \cdots, W(t_n))$ we have

$$G^\gamma(X) = F(W^\gamma(t_1), \cdots, W^\gamma(t_n)),$$

where $F : \mathbf{R}^{nd} \to \mathbf{R}$, $t_1 < \cdots < t_n$ with $t_i \in [0, T]$ and $n > 0$.

Theorem 1

For each $\gamma \in L^2(\mathcal{B}_T, \lambda)$, $p > 2$, there is an isometry

$$G^\gamma : L^p(\mathcal{F}_T, P) \to L^p(\mathcal{F}_T, P^\gamma)$$

such that for each $X = F(W(t_1), \cdots, W(t_n))$ we have

$$G^\gamma(X) = F(W^\gamma(t_1), \cdots, W^\gamma(t_n)),$$

where $F : \mathbf{R}^{nd} \to \mathbf{R}$, $t_1 < \cdots < t_n$ with $t_i \in [0, T]$ and $n > 0$. Moreover, for each $X \in L^p(\mathcal{F}_T, P)$ we have

$$E_P(G^\gamma(X)\mathcal{E}(W(\gamma))) = E_P(X),$$

$$Var_P(G^\gamma(X)\mathcal{E}(W(\gamma))) = E_P\left(X^2\mathcal{E}(W(\gamma))\exp\left(\|\gamma\|_\lambda^2\right)\right) - E_P(X)^2$$

Theorem 1

$$\begin{aligned} & E_P\left(F\left(W\left(t_1\right),\cdots,W\left(t_n\right)\right)\right) \\ = \;& E_{P^\gamma}\left(F\left(W^\gamma\left(t_1\right),\cdots,W^\gamma\left(t_n\right)\right)\right) \\ = \;& E_P\left(F\left(W^\gamma\left(t_1\right),\cdots,W^\gamma\left(t_n\right)\right)\mathcal{E}\left(W\left(\gamma\right)\right)\right) \end{aligned}$$

Theorem 1

$$\begin{aligned} & E_P\left(F\left(W\left(t_1\right),\cdots,W\left(t_n\right)\right)\right) \\ = \quad & E_{P^\gamma}\left(F\left(W^\gamma\left(t_1\right),\cdots,W^\gamma\left(t_n\right)\right)\right) \\ = \quad & E_P\left(F\left(W^\gamma\left(t_1\right),\cdots,W^\gamma\left(t_n\right)\right)\mathcal{E}\left(W\left(\gamma\right)\right)\right) \end{aligned}$$

Since $\mathcal{E}\left(W\left(\gamma\right)\right)=\mathcal{E}\left(W^\gamma\left(\gamma\right)\right)\exp\left(\|\gamma\|_\lambda^2\right)$, we get

$$\begin{aligned} & E_P\left(\left(F\left(W^\gamma\left(t_1\right),\cdots,W^\gamma\left(t_n\right)\right)\mathcal{E}\left(W\left(\gamma\right)\right)\right)^2\right) \\ = \quad & E_{P^\gamma}\left(F\left(W^\gamma\left(t_1\right),\cdots,W^\gamma\left(t_n\right)\right)^2\mathcal{E}\left(W\left(\gamma\right)\right)\right) \\ = \quad & E_{P^\gamma}\left(F\left(W^\gamma\left(t_1\right),\cdots,W^\gamma\left(t_n\right)\right)^2\mathcal{E}\left(W^\gamma\left(\gamma\right)\right)\right)\exp\left(\|\gamma\|_\lambda^2\right) \\ = \quad & E_P\left(F\left(W\left(t_1\right),\cdots,W\left(t_n\right)\right)^2\mathcal{E}\left(W\left(\gamma\right)\right)\right)\exp\left(\|\gamma\|_\lambda^2\right) \end{aligned}$$

Minimization Problem

Infinite-dimensional Minimization Problem

- Let $L(\gamma) = \exp\left(\|\gamma\|_\lambda^2\right)\left(X^2, \mathcal{E}\left(W(\gamma)\right)\right)_P$ and find

$$\bar{\gamma} \in \arg\min\left\{L(\gamma) \,:\, \gamma \in L^2\left(\mathcal{B}_T, \lambda\right)\right\}.$$

Minimization Problem

Infinite-dimensional Minimization Problem

- Let $L(\gamma) = \exp\left(\|\gamma\|_\lambda^2\right)\left(X^2, \mathcal{E}(W(\gamma))\right)_P$ and find

$$\bar{\gamma} \in \arg\min\left\{L(\gamma) \,:\, \gamma \in L^2(\mathcal{B}_T, \lambda)\right\}.$$

Finite-dimensional Minimization Problem

- choose $\{\phi_1, ..., \phi_n\}$ *real*-valued, $\mathcal{B}_T$-measurable square-integrable functions of unit norm which are linearly independent
- set $\mathcal{D} = \sum_{i=1}^n \mathcal{D}_i$, where $\mathcal{D}_i = \left\{\gamma : \gamma = \alpha\,\phi_i, \alpha \in \mathbf{R}^d\right\}$
- find $\bar{\gamma} \in \arg\min\{L(\gamma) \,:\, \gamma \in \mathcal{D}\}$.

Theorem 2

Suppose that $X \in L^p(\mathcal{F}_T, P)$ with $p > 2$ and $X \neq 0$. Then there is $\alpha > 0$ with

$$D^2L(\gamma) \cdot (\delta, \delta) \geq \alpha \|\delta\|^2,$$

i.e. L is strictly convex on $\mathcal{D}$. Furthermore, L is coercive in the sense that the set $\{\eta \in \mathcal{D} : L(\eta) \leq L(\gamma)\}$ is compact for every $\gamma \in \mathcal{D}$. The finite-dimensional minimization problem has a unique minimum point.

MM-Algorithm

- let $\mathcal{C} = \bigtimes_{i=1}^{n} \mathcal{C}_i$ with open subsets $\mathcal{C}_i$ of $\mathbf{R}^d$

MM-Algorithm

- let $\mathcal{C} = \times_{i=1}^{n} \mathcal{C}_i$ with open subsets $\mathcal{C}_i$ of $\mathbf{R}^d$
- for $x = (x_1, \cdots, x_n) \in \mathcal{C}$, $y \in \mathcal{C}_i$ define $\hat{x}_i(y) \in \mathcal{C}$ by

$$\hat{x}_i(y)_j = \begin{cases} x_j & : \quad j \neq i \\ y & : \quad j = i \end{cases}$$

MM-Algorithm

- let $\mathcal{C} = \bigtimes_{i=1}^{n} \mathcal{C}_i$ with open subsets $\mathcal{C}_i$ of $\mathbf{R}^d$
- for $x = (x_1, \cdots, x_n) \in \mathcal{C}$, $y \in \mathcal{C}_i$ define $\hat{x}_i(y) \in \mathcal{C}$ by

$$\hat{x}_i(y)_j = \begin{cases} x_j & : \quad j \neq i \\ y & : \quad j = i \end{cases}$$

- let $f : \mathcal{C} \to \mathbf{R}$ be an objective function and set $f_i : \mathcal{C}_i \times \mathcal{C} \to \mathbf{R}$ by $f_i(y, x) = f(\hat{x}_i(y))$

MM-Algorithm

- let $\mathcal{C} = \bigtimes_{i=1}^{n} \mathcal{C}_i$ with open subsets $\mathcal{C}_i$ of $\mathbf{R}^d$
- for $x = (x_1, \cdots, x_n) \in \mathcal{C}$, $y \in \mathcal{C}_i$ define $\hat{x}_i(y) \in \mathcal{C}$ by

$$\hat{x}_i(y)_j = \begin{cases} x_j & : \quad j \neq i \\ y & : \quad j = i \end{cases}$$

- let $f : \mathcal{C} \to \mathbf{R}$ be an objective function and set $f_i : \mathcal{C}_i \times \mathcal{C} \to \mathbf{R}$ by $f_i(y, x) = f(\hat{x}_i(y))$
- $g_i : \mathcal{C}_i \times \mathcal{C}_i \times \mathcal{C} \to \mathbf{R}$, $1 \leq i \leq n$, such that

$$\begin{aligned} i)\ f_i(y, x) &= g_i(y, y, x), \\ ii)\ f_i(y, x) &\leq g_i(y, z, x) \text{ for all } z \in \mathcal{C}_i. \end{aligned}$$

MM-Algorithm

- i) choose a starting point $x^0 \in \mathcal{C}$,

MM-Algorithm

- i) choose a starting point $x^0 \in \mathcal{C}$,
- ii) set $x^{k,0} = x^k$ and do iii) for $j = 1, \cdots, n$,

MM-Algorithm

- i) choose a starting point $x^0 \in \mathcal{C}$,
- ii) set $x^{k,0} = x^k$ and do iii) for $j = 1, \cdots, n$,
- iii) find $z^j = \arg\min \left\{ g_j \left(y, x_j^k, x^{k,j-1} \right) : y \in \mathcal{C}_j \right\}$ and $x^{k,j} = x^{k,\hat{j}-1}{}_j \left(z^j \right)$,

MM-Algorithm

- i) choose a starting point $x^0 \in \mathcal{C}$,
- ii) set $x^{k,0} = x^k$ and do iii) for $j = 1, \cdots, n$,
- iii) find $z^j = \arg\min \left\{ g_j \left(y, x_j^k, x^{k,j-1} \right) : y \in \mathcal{C}_j \right\}$ and $x^{k,j} = x^{k,\hat{j}-1}{}_j \left(z^j \right)$,
- iV) set $x^{k+1} = x^{k,n}$ and continue with ii).

MM-Algorithm

The MM-algorithm generates a non-increasing sequence, since

$$f\left(x^{k,j-1}\right) = f_j\left(x_j^k, x^{k,j-1}\right) = g_j\left(x_j^k, x_j^k, x^{k,j-1}\right)$$

and

$$f\left(x^{k,j}\right) = f_j\left(z^j, x^{k,j-1}\right) \leq g_j\left(z^j, x_j^k, x^{k,j-1}\right),$$

so that property iii) implies $f\left(x^{k+1}\right) \leq f\left(x^k\right)$.

ω-limit set

- $(x^k)_{k\geq 0}$ a sequence generated by the MM-algorithm with starting point x^0

ω-limit set

- $\left(x^k\right)_{k\geq 0}$ a sequence generated by the MM-algorithm with starting point x^0
- $\omega\left(x^0\right)$-limit set is defined by

$$\omega\left(x^0\right) = \left\{u \in \mathcal{C} : \text{increasing } (k_i)_{i\geq 1} \text{ with } \lim_{i\to\infty} x^{k_i} = u\right\}.$$

ω-limit set

- $\left(x^k\right)_{k\geq 0}$ a sequence generated by the MM-algorithm with starting point x^0
- $\omega\left(x^0\right)$-limit set is defined by

$$\omega\left(x^0\right) = \left\{u \in \mathcal{C} : \text{increasing } \left(k_i\right)_{i\geq 1} \text{ with } \lim_{i\to\infty} x^{k_i} = u\right\}.$$

- $\Sigma\left(f\right) = \{x \in \mathcal{C} : Df\left(x\right) = 0\}$ is called the set of stationary points of f.

Theorem 3

Suppose that f is coercive and continuously differentiable. Furthermore, we assume that $g_i, D_1 g_i, D_1^2 g_i$ are continuous maps for all $1 \leq i \leq n$. Additionally, if $n > 1$, then we assume that $D_1^2 g_i$ is positive definite. Then $\omega\left(x^0\right) \neq \emptyset$ and $\omega\left(x^0\right) \subset \Sigma(f)$ for all $x^0 \in \mathcal{C}$. If f is strictly convex, then every sequence generated by the MM-algorithm converges to the unique minimum point of f.

$$D_j f(x) = 0 \Longleftrightarrow D_1 g_j(x_j, x_j, x) = 0 \text{ for all } 1 \leq j \leq n.$$

Construction of the MM-Algorithm

- $\mathcal{C} = \bigtimes_{i=1}^{n} \mathcal{C}_i$ with $\mathcal{C}_i = \mathbf{R}^d$

Construction of the MM-Algorithm

- $\mathcal{C} = \bigtimes_{i=1}^{n} \mathcal{C}_i$ with $\mathcal{C}_i = \mathbf{R}^d$
- $f(x) = L(\Phi(x))$ with $\Phi(x) = \sum_{i=1}^{n} x_i \phi_i$

Construction of the MM-Algorithm

- $\mathcal{C} = \bigtimes_{i=1}^{n} \mathcal{C}_i$ with $\mathcal{C}_i = \mathbf{R}^d$
- $f(x) = L(\Phi(x))$ with $\Phi(x) = \sum_{i=1}^{n} x_i \phi_i$
- with $\mathcal{G}(\gamma) = \exp\left(-\|\gamma\|_\lambda^2 + W(\gamma)\right)$ we have

$$\|\mathcal{G}(\gamma)\|_P = 1$$

and

$$L(\gamma) = \exp\left(\frac{3}{2}\|\gamma\|_\lambda^2\right)\left(X^2, \mathcal{G}(\gamma)\right)_P$$

Construction of the MM-Algorithm

- $\mathcal{C} = \bigtimes_{i=1}^{n} \mathcal{C}_i$ with $\mathcal{C}_i = \mathbf{R}^d$
- $f(x) = L(\Phi(x))$ with $\Phi(x) = \sum_{i=1}^{n} x_i \phi_i$
- with $\mathcal{G}(\gamma) = \exp\left(-\|\gamma\|_\lambda^2 + W(\gamma)\right)$ we have

$$\|\mathcal{G}(\gamma)\|_P = 1$$

and

$$L(\gamma) = \exp\left(\frac{3}{2}\|\gamma\|_\lambda^2\right)\left(X^2, \mathcal{G}(\gamma)\right)_P$$

- $|f(x)| \leq \exp\left(\frac{3}{2}\|\Phi(x)\|_\lambda^2\right)\|X^2\|_P$

Construction of the MM-Algorithm

- $g_i(y,z,x) = \exp\left(q_i(y,z,x)\right)$, where q_i is a quadratic function with respect to y

Construction of the MM-Algorithm

- $g_i(y,z,x) = \exp(q_i(y,z,x))$, where q_i is a quadratic function with respect to y
- defined by $\tilde{f}(x) = \ln(X^2, \mathcal{G}(\Phi(x)))_P$

Construction of the MM-Algorithm

- $g_i(y,z,x) = \exp(q_i(y,z,x))$, where q_i is a quadratic function with respect to y
- defined by $\tilde{f}(x) = \ln(X^2, \mathcal{G}(\Phi(x)))_P$
- note $\ln(\alpha + x) \leq \ln \alpha + x/\alpha$

Construction of the MM-Algorithm

- $g_i\left(y,z,x\right) = \exp\left(q_i\left(y,z,x\right)\right)$, where q_i is a quadratic function with respect to y
- defined by $\tilde{f}\left(x\right) = \ln(X^2, \mathcal{G}\left(\Phi\left(x\right)\right))_P$
- note $\ln\left(\alpha + x\right) \leq \ln\alpha + x/\alpha$

$$\tilde{f}\left(\hat{x}_i\left(y\right)\right) = \ln\left(\left(X^2, \mathcal{G}\left(\Phi\left(\hat{x}_i\left(y\right)\right)\right) - \mathcal{G}\left(\Phi\left(\hat{x}_i\left(z\right)\right)\right) + \mathcal{G}\left(\Phi\left(\hat{x}_i\left(z\right)\right)\right)\right)\right)$$

$$\leq \tilde{f}\left(\hat{x}_i\left(z\right)\right) + (X^2, \mathcal{G}\left(\Phi\left(\hat{x}_i\left(y\right)\right)\right) - \mathcal{G}\left(\Phi\left(\hat{x}_i\left(z\right)\right)\right))/(X^2, \mathcal{G}\left(\Phi\left(\hat{x}_i\left(z\right)\right)\right))$$

Construction of the MM-Algorithm

- find surrogate function for $(X^2, \mathcal{G}(\Phi(\hat{x}_i(y))) - \mathcal{G}(\Phi(\hat{x}_i(z))))_P$

Construction of the MM-Algorithm

- find surrogate function for $\left(X^2, \mathcal{G}\left(\Phi\left(\hat{x}_i\left(y\right)\right)\right) - \mathcal{G}\left(\Phi\left(\hat{x}_i\left(z\right)\right)\right)\right)_P$
- $\left\|\Phi\left(\hat{x}_i\left(y\right)\right)\right\|_\lambda^2 = G_{0,i}\left(x\right) + \left(G_{1,i}\left(x\right), y\right) + \|y\|^2$

Construction of the MM-Algorithm

- find surrogate function for $\left(X^2, \mathcal{G}\left(\Phi\left(\hat{x}_i\left(y\right)\right)\right) - \mathcal{G}\left(\Phi\left(\hat{x}_i\left(z\right)\right)\right)\right)_P$
- $\|\Phi\left(\hat{x}_i\left(y\right)\right)\|_\lambda^2 = G_{0,i}\left(x\right) + \left(G_{1,i}\left(x\right), y\right) + \|y\|^2$
- $W\left(\Phi\left(\hat{x}_i\left(y\right)\right)\right) = G_{2,i}\left(x\right) + \left(y, W\left(\phi_i\right)\right)$

Construction of the MM-Algorithm

- find surrogate function for $(X^2, \mathcal{G}(\Phi(\hat{x}_i(y))) - \mathcal{G}(\Phi(\hat{x}_i(z))))_P$
- $\|\Phi(\hat{x}_i(y))\|_\lambda^2 = G_{0,i}(x) + (G_{1,i}(x)\ , y) + \|y\|^2$
- $W(\Phi(\hat{x}_i(y))) = G_{2,i}(x) + (y, W(\phi_i))$
- $G_{0,i}(x) = \sum_{k,l\in \mathsf{N}_n \setminus \{i\}} (x_k, x_l)\rho_{kl}$,
  $G_{1,i}(x) = 2\sum_{k\in \mathsf{N}_n \setminus \{i\}} x_k\rho_{ki}$

Construction of the MM-Algorithm

- find surrogate function for $\left(X^2, \mathcal{G}\left(\Phi\left(\hat{x}_i\left(y\right)\right)\right) - \mathcal{G}\left(\Phi\left(\hat{x}_i\left(z\right)\right)\right)\right)_P$
- $\left\|\Phi\left(\hat{x}_i\left(y\right)\right)\right\|_\lambda^2 = G_{0,i}\left(x\right) + \left(G_{1,i}\left(x\right), y\right) + \|y\|^2$
- $W\left(\Phi\left(\hat{x}_i\left(y\right)\right)\right) = G_{2,i}\left(x\right) + \left(y, W\left(\phi_i\right)\right)$
- $G_{0,i}\left(x\right) = \sum_{k,l \in \mathbf{N}_n \setminus \{i\}} \left(x_k, x_l\right) \rho_{kl}$, $G_{1,i}\left(x\right) = 2 \sum_{k \in \mathbf{N}_n \setminus \{i\}} x_k \rho_{ki}$
- $G_{2,i}\left(x\right) = \sum_{k \in \mathbf{N}_n \setminus \{i\}} \left(x_k, W\left(\phi_k\right)\right)$, $W\left(\phi_k\right) = \left(W_1\left(\phi_k\right), \cdots, W_d\left(\phi_k\right)\right)$

Construction of the MM-Algorithm

- $\mathcal{G}\left(\Phi\left(\hat{x}_i\left(y\right)\right)\right) = \Gamma_i\left(x\right)\exp\left(-\|y\|^2 + \left(y\,, W\left(\phi_i\right) - G_{1,i}\left(x\right)\right)\right)$

Construction of the MM-Algorithm

- $\mathcal{G}\left(\Phi\left(\hat{x}_i\left(y\right)\right)\right) = \Gamma_i\left(x\right)\exp\left(-\|y\|^2 + \left(y\,, W\left(\phi_i\right) - G_{1,i}\left(x\right)\right)\right)$
- $h\left(x, y, z\right) = e^{-\|x\|^2 + \left(x\,,z\right)} - e^{-\|y\|^2 + \left(y\,,z\right)}, \quad x, y, z \in \mathbf{R}^d$

Construction of the MM-Algorithm

- $\mathcal{G}\left(\Phi\left(\hat{x}_i\left(y\right)\right)\right) = \Gamma_i\left(x\right) \exp\left(-\|y\|^2 + \left(y\,,W\left(\phi_i\right) - G_{1,i}\left(x\right)\right)\right)$
- $h\left(x,y,z\right) = e^{-\|x\|^2+\left(x\,,z\right)} - e^{-\|y\|^2+\left(y\,,z\right)}, \quad x,y,z \in \mathbf{R}^d$

$$\left(X^2\,,\mathcal{G}\left(\Phi\left(\hat{x}_i\left(y\right)\right)\right) - \mathcal{G}\left(\Phi\left(\hat{x}_i\left(z\right)\right)\right)\right)_P$$

$$= \int_\Omega X^2 \Gamma_i\left(x\right) h\left(y,z,W\left(\phi_i\right) - G_{1,i}\left(x\right)\right) dP$$

Construction of the MM-Algorithm

Let

$$\begin{aligned} p_+(x,y,z) &= e^{-\|y\|^2+(y\,,z)}(z-2y\,,x-y)+2e^{\|z\|^2/4-3/2}\|x-y\|^2, \\ p_-(x,y,z) &= e^{-\|y\|^2+(y\,,z)}(z-2y\,,x-y)-e^{\|z\|^2/4}\|x-y\|^2. \end{aligned}$$

Construction of the MM-Algorithm

Let

$$\begin{aligned} p_+(x,y,z) &= e^{-\|y\|^2+(y\,,z)}(z-2y\,,x-y)+2e^{\|z\|^2/4-3/2}\|x-y\|^2, \\ p_-(x,y,z) &= e^{-\|y\|^2+(y\,,z)}(z-2y\,,x-y)-e^{\|z\|^2/4}\|x-y\|^2. \end{aligned}$$

Then $p_- \leq h \leq p_+$ and $h(x,x,z)=p_\pm(x,x,z)=0$ for all $x,z \in \mathbf{R}^d$.

Lemma 2

Let $X \in L^p(\mathcal{F}_T, P)$ with $p > 2$. Then we have

$$\left(X^{\pm}, \mathcal{G}\left(\Phi\left(\hat{x}_i\left(y\right)\right)\right) - \mathcal{G}\left(\Phi\left(\hat{x}_i\left(z\right)\right)\right)\right)_P$$

$$\leq \alpha^{\pm} a\left(X^{\pm}, x, z, i\right) \|y - z\|^2 + \left(b\left(X^{\pm}, x, z, i\right), y - z\right)$$

with $X^+ = X \vee 0$, $X^- = X \wedge 0$, $\alpha^+ = 2e^{-3/2}$, $\alpha^- = -1$ and

$$\begin{aligned} a &= \int_{\Omega} X \mathcal{G}\left(\Phi\left(\hat{x}_i\left(z\right)\right)\right) \exp\left(\frac{1}{4}\|W\left(\phi_i\right) - G_{1,i}\left(x\right) - 2z\|^2\right) dP, \\ b &= \int_{\Omega} X \mathcal{G}\left(\Phi\left(\hat{x}_i\left(z\right)\right)\right)\left(W\left(\phi_i\right) - G_{1,i}\left(x\right) - 2z\right) dP. \end{aligned}$$

Surrogate-Functions

Defined $g_i : \mathcal{C}_i \times \mathcal{C}_i \times \mathcal{C} \longrightarrow$ **R** by

$$g_i\left(y, z, x\right) = f_i\left(z, x\right) \exp\left(\alpha_2\left(z, x, i\right) \|y - z\|^2 + \left(\alpha_1\left(z, x, i\right)\, , y - z\right)\right)$$

with

$$\begin{aligned}
\alpha_2\left(z, x, i\right) &= 1.5 + 2\, e^{-3/2} a\left(X^2, x, z, i\right) / \left(X^2\, , \mathcal{G}\left(\Phi\left(\hat{x}_i\left(z\right)\right)\right)\right), \\
\alpha_1\left(z, x, i\right) &= 3\, z + 1.5\, G_{1,i}\left(x\right) + b\left(X^2, x, z, i\right) / \left(X^2\, , \mathcal{G}\left(\Phi\left(\hat{x}_i\left(z\right)\right)\right)\right)
\end{aligned}$$

Theorem 4

Suppose $X \in L^p(\mathcal{F}_T, P)$ with $p > 4$ and $X \neq 0$. Then the MM-algorithm defined by $f, g_1, \cdots, g_n$ generates for every starting point $x^0 \in \mathcal{C}$ a sequence converging to the unique minimum point of f.

M-Estimator

- $l\,(\xi, \eta, \theta) = \xi^2 \exp\left((\eta\,,\theta) + \frac{1}{2}(\theta\,,\rho^{(d)}\theta)\right),\ (\xi, \eta, \theta) \in \mathbf{R}\times\mathcal{C}\times\mathcal{C}$

M-Estimator

- $l(\xi,\eta,\theta)=\xi^2\exp\left((\eta\,,\theta)+\frac{1}{2}(\theta\,,\rho^{(d)}\theta)\right),\ (\xi,\eta,\theta)\in \mathbf{R}\times\mathcal{C}\times\mathcal{C}$
- with $W(\Phi)=(W(\phi_1),\cdots,W(\phi_n))$ we get

$$E_P(l(X,W(\Phi),\theta))=f(\theta)$$

M-Estimator

- $l\left(\xi,\eta,\theta\right)=\xi^{2}\exp\left(\left(\eta\,,\theta\right)+\frac{1}{2}\left(\theta\,,\rho^{(d)}\theta\right)\right),\ \left(\xi,\eta,\theta\right)\in \mathbf{R}\times\mathcal{C}\times\mathcal{C}$
- with $W\left(\Phi\right)=\left(W\left(\phi_{1}\right),\cdots,W\left(\phi_{n}\right)\right)$ we get

$$E_{P}\left(l\left(X,W\left(\Phi\right),\theta\right)\right)=f\left(\theta\right)$$

- Let $\left(\left(X_{i},Y_{i}\right)\right)_{i\geq 1}$ be a sequence of $i.i.d$ random variables, where $\left(X_{1},Y_{1}\right)$ has the same distribution as $\left(X,W\left(\Phi\right)\right)$, and define

$$l_{N}\left(\theta\right)=N^{-1}\sum_{i=1}^{N}l\left(X_{i},Y_{i},\theta\right).$$

Then $\lim_{N\rightarrow\infty}l_{N}\left(\theta\right)=f\left(\theta\right)$ a.s. for all $\theta\in\mathcal{C}$.

M-Estimator

- Definition of M-Estimator $\hat{\Theta}_N$

$$\hat{\Theta}_N \in \arg\min \{ l_N(\theta) : \theta \in \mathcal{C} \} \ a.s.$$

M-Estimator

- Definition of M-Estimator $\hat{\Theta}_N$

$$\hat{\Theta}_N \in \arg\min \{ l_N(\theta) : \theta \in \mathcal{C} \} \ a.s.$$

- Consistency of $\hat{\Theta}_N$

  Suppose that $X \in L^p(\mathcal{F}_T, P)$ with $p > 2$. Then $\hat{\Theta}_N$, $N \geq 1$, is consistent, i.e. we have $\lim_{N \to \infty} \hat{\Theta}_N = \theta_0$ a.s.

M-Estimator

- Definition of M-Estimator $\hat{\Theta}_N$

$$\hat{\Theta}_N \in \arg\min \left\{ l_N(\theta) : \theta \in \mathcal{C} \right\} \ a.s.$$

- Consistency of $\hat{\Theta}_N$

  Suppose that $X \in L^p(\mathcal{F}_T, P)$ with $p > 2$. Then $\hat{\Theta}_N$, $N \geq 1$, is consistent, i.e. we have $\lim_{N\to\infty} \hat{\Theta}_N = \theta_0$ a.s.

- Asymptotic Normality of $\hat{\Theta}_N$

  Suppose that $X \in L^p(\mathcal{F}_T, P)$ with $p > 2$. Then

$$\sqrt{N}\left(\hat{\Theta}_N - \theta_0\right) \xrightarrow{\mathcal{D}} N\left(0, J(\theta_0)^{-1} \mathcal{I}(\theta_0) J(\theta_0)^{-1}\right),$$

  with $J_{ij}(\theta) = D^2 f(\theta) \cdot (e_i, e_j)$ for $1 \leq i, j \leq nd$.

Numerical Results

- Black-Scholes model

$$dS_i(t) = S_i(t)\left(r\,dt + \left(\sigma_i, dW(t)\right)\right), \ \left(S_1(0), \cdots, S_m(0)\right) = S_0$$

Numerical Results

- Black-Scholes model

$$dS_i(t) = S_i(t)\left(r\,dt + (\sigma_i, dW(t))\right), \ \left(S_1(0), \cdots, S_m(0)\right) = S_0$$

- $\mathcal{T} = \{0 = T_0 < T_1 \cdots < T_n = T\}$

Numerical Results

- Black-Scholes model

$$dS_i\,(t) = S_i\,(t)\,(r\,dt + (\sigma_i\,, dW\,(t)))\,,\ \ (S_1\,(0)\,, \cdots, S_m\,(0)) = S_0$$

- $\mathcal{T} = \{0 = T_0 < T_1 \cdots < T_n = T\}$
- $\Phi = \{\phi_1, \cdots, \phi_n\}$ with $\phi_i = \chi_{[T_{i-1},T_i]}/\sqrt{T_i - T_{i-1}}$

Numerical Results

- Black-Scholes model

$$dS_i(t) = S_i(t)\left(r\,dt + (\sigma_i, dW(t))\right),\ (S_1(0), \cdots, S_m(0)) = S_0$$

- $\mathcal{T} = \{0 = T_0 < T_1 \cdots < T_n = T\}$
- $\Phi = \{\phi_1, \cdots, \phi_n\}$ with $\phi_i = \chi_{[T_{i-1}, T_i]}/\sqrt{T_i - T_{i-1}}$
- $S_i(T_j) =$
$S_{0,i} \exp\left(rT_j - \frac{1}{2}\|\sigma_i\|^2 T_j + \sum_{l=1}^{j}(\sigma_i, W(\phi_l))\sqrt{T_l - T_{l-1}}\right)$

Numerical Results

- Black-Scholes model

$$dS_i(t) = S_i(t)\left(r\,dt + (\sigma_i, dW(t))\right),\ (S_1(0), \cdots, S_m(0)) = S_0$$

- $\mathcal{T} = \{0 = T_0 < T_1 \cdots < T_n = T\}$
- $\Phi = \{\phi_1, \cdots, \phi_n\}$ with $\phi_i = \chi_{[T_{i-1},T_i]}/\sqrt{T_i - T_{i-1}}$
- $S_i(T_j) =$ $S_{0,i} \exp\left(rT_j - \frac{1}{2}\|\sigma_i\|^2 T_j + \sum_{l=1}^{j} (\sigma_i, W(\phi_l))\sqrt{T_l - T_{l-1}}\right)$
- Payoff-Function: $X = F(W(\Phi))$, with

$$W(\Phi) = (W(\phi_1), \cdots, W(\phi_n))$$

Numerical Results

- The price

$$\pi(X) = e^{-rT} E_P(F(W(\Phi)))$$

is also given by

$$\pi(X) = e^{-rT} E_P\left(F\left(W^{\Phi(\theta)}(\Phi)\right)\mathcal{E}(W(\Phi(\theta)))\right)$$

Numerical Results

- The price

$$\pi(X) = e^{-rT} E_P(F(W(\Phi)))$$

is also given by

$$\pi(X) = e^{-rT} E_P\left(F\left(W^{\Phi(\theta)}(\Phi)\right)\mathcal{E}(W(\Phi(\theta)))\right)$$

- with

$$F(\eta,\theta) = F(\eta-\theta)\exp\left(-\frac{1}{2}\|\theta\|^2 + (\theta\,,\eta)\right),\ \ \eta,\theta \in \mathcal{C}$$

we get $\pi(X) = e^{-rT} E_P(F(W(\Phi)\,,\theta))$.

Numerical Results

- Monte Carlo estimator

$$\pi_N(X,\theta) = \frac{1}{N}\sum_{i=1}^{N} F(Y_i,\theta)$$

Numerical Results

- Monte Carlo estimator

$$\pi_N(X,\theta) = \frac{1}{N}\sum_{i=1}^{N} F(Y_i,\theta)$$

- Measure of variance reduction

$$VR_N(X,\theta) = \frac{\sum_{i=1}^{N}\left(F(Y_i,\theta) - \pi_N(X,\theta)\right)^2}{\sum_{i=1}^{N}\left(F(Y_i,0) - \pi_N(X,0)\right)^2}.$$

European option

Parameters		Black-Scholes	MM-Algorithm	
σ	K	Price	Price$_N$	VR$_{N_1}^{M_1}$
30%	30	0.13	0.13	16.8
	40	1.28	1.28	10.4
	50	4.68	4.69	6.3
	60	10.53	10.53	4.9
10%	40	0.0042	0.0038	6.4
	50	0.96	0.96	9.5
	60	7.31	7.30	5.9

$N = 40000$ Monte Carlo samples for estimating prices, $N_1 = 500, M_1 = 100$ for estimating optimal drift $\hat{\theta}_0$ with the MM-algorithm. We used: $S_0 = 50, r = 5\%, T = 1$.

European option

Parameters		MM-Algorithm		Robbins-Monro	
σ	K	Price$_N$	VR$_{N_2}^{M_2}$	Price	VR
30%	30	0.13	25.0	0.13	38.4
	40	1.28	11.3	1.28	10.9
	50	4.69	6.3	4.68	6.3
	60	10.53	4.9	10.54	4.8
10%	40	0.004	22.4	0.0042	349.7
	50	0.97	9.7	0.97	9.6
	60	7.30	5.9	7.31	6.3

$N = 40000$ Monte Carlo samples for estimating prices and $N_2 = 2000, M_2 = 500$ for estimating optimal drift $\hat{\theta}_0$ with the MM-algorithm. We used: $S_0 = 50, r = 5\%, T = 1$.

Asian option (n=16)

Parameters		MM-Algorithm				GHS	
σ	K	Price$_N$	VR$_{N_1}^{M_1}$	Price$_N$	VR$_{N_2}^{M_2}$	Price	VR
10%	45	6.06	10.2	6.06	10.9	6.05	11
	50	1.92	7.3	1.92	7.5	1.92	7
	55	0.20	16.7	0.20	21.2	0.20	21
30%	45	7.15	8.5	7.15	8.9	7.15	8.3
	50	4.17	9.3	4.17	9.7	4.17	9.2
	55	2.21	11.6	2.21	12.5	2.21	12

$N = 1.000.000$ Monte Carlo samples for estimating prices, $N_1 = 1000, M_1 = 100$, $N_2 = 2000, M_2 = 50$ for estimating optimal drift $\hat{\theta}_0$ with the MM-algorithm. We used: $S_0 = 50$, $r = 5\%$, $T = 1$.

Asian option (n=64)

Parameters		MM-Algorithm				GHS	
σ	K	Price$_N$	VR$_{N_1}^{M_1}$	Price$_N$	VR$_{N_2}^{M_2}$	Price	VR
10%	45	5.99	5.7	6.00	7.9	6.00	11
	50	1.85	6	1.85	6.9	1.85	7.3
	55	0.17	11	0.17	16.2	0.17	23
30%	45	7.02	6.8	7.02	7.9	7.02	8.3
	50	4.02	7.5	4.02	8.8	4.02	9.2
	55	2.08	9.2	2.08	11.3	2.08	12