Optimal Execution using Statistical Learning

Sophie LARUELLE

LAMA-UPEC, CMAP-Polytechnique

July 6th, 2016

International Conference on Monte-Carlo Techniques

LABORATOIRE D'ANALYSE
ET DE MATHÉMATIQUES APPLIQUÉES

Sous la co-tutelle de :
CNRS
UPEC • UNIVERSITÉ PARIS-EST CRÉTEIL
UPEM • UNIVERSITÉ PARIS-EST MARNE-LA-VALLÉE

Outline

Outline

1. Why Using Statistical Learning for Optimal Execution ?

2. Introduction to Stochastic Approximation Algorithms

3. Applications to Optimal Execution
   - Intraday volume curve for optimal trade scheduling
   - Optimal split of orders across liquidity pools
   - Optimal posting price of limit orders: learning by trading
   - Optimal split and posting price of limit orders across lit pools

Why Using Statistical Learning for Optimal Execution ?

- **High frequency market data** : Database with only the trades or all the events that may happen in the LOB $\Longrightarrow$ massive database.
- **Evolutions of technology** : The speed of processors and their computational throughput have both increased. The traders use more powerful computers or dedicated hardware like FPGAs to accelerate their decision taking process.
- **Parameter Estimation** : Non stationary condition of parameters, global estimation to take into account the correlations $\Longrightarrow$ statistical learning algorithm to estimate market parameters (like some use in pattern recognition for computer vision).
- **Interaction with the LOB** : Interactions between market participants imply to estimate online the parameters and not from an historical database (misspecification risk) $\Longrightarrow$ update parameter values in trading algorithm.

Outline

Motivations

In many fields, we often are faced with optimization or zero search problems.

Examples: In Finance,

- extraction of implicit parameter (volatility),
- calibration, optimization of an exogenous parameter for variance reduction (regression, importance sampling, etc).

Common point: all the concerned functions have a representation as an expectation, *i.e.*

$$h(\theta) = \mathbb{E}\left[H(\theta, Y)\right], \quad \text{where } Y \text{ is a } q\text{-dimensional random vector.}$$

The stochastic approximation is a tool based on simulation to solve optimization or zero search problems.

Measurement : Robbins-Monro procedure

A dose θ of a chemical product creates a **random effect** measured by $F(\theta, Y)$, Y being a random variable with distribution μ and $F : \mathbb{R}^2 \mapsto \mathbb{R}$.

We assume that the **mean effect**

$$f(\theta) = \mathbb{E}\left[F(\theta, Y)\right] \quad \text{is non-decreasing.}$$

We want to determine the dose θ^* which creates a mean effect of a given threshold α, *i.e.* to solve

$$f(\theta^*) = \alpha.$$

- **Naive idea**: $f(\theta_n) \approx \frac{1}{N_n} \sum_{k=1}^{N_n} F(\theta_n, Y_k)$, Y_k i.i.d. with distribution μ, $N_n \underset{n \to \infty}{\longrightarrow} \infty$.
- **Other idea**: use $F(\theta_n, Y_{n+1})$ instead of $f(\theta_n)$ with $\gamma_n \searrow 0$ ("local" randomization).

The **Robbins-Monro procedure** is the following

- choose arbitrarily θ_0 and administer it to a subject which reacts with the effect $F(\theta_0, Y_1)$.
- *Recurrence*: at instant n, choose a dose θ_n administered to a subject (independent of the previous ones), the effect is $F(\theta_n, Y_{n+1})$.

As $(Y_n)_{n\geq 1}$ is a sequence of i.i.d. random variables with distribution μ, then

$$f(\theta_n) = \mathbb{E}\left[F(\theta_n, Y_{n+1}) \mid F(\theta_0, Y_1), \ldots, F(\theta_{n-1}, Y_n)\right].$$

The **Robbins-Monro algorithm** for the choice of θ_n then reads

$$\theta_{n+1} = \theta_n - \gamma_n \left(F(\theta_n, Y_{n+1}) - \alpha\right), \quad (\gamma_n) \text{ non-negative tending to 0.}$$

By setting $H(\theta, y) := F(\theta, y) - \alpha$, this procedure is then a zero search of the function

$$h(\theta) := f(\theta) - \alpha = \mathbb{E}\left[F(\theta, Y) - \alpha\right] = \mathbb{E}\left[H(\theta, Y)\right].$$

Convergence: Martingale approach or ODE method

- **Martingale approach**: Existence of a Lyapunov function and control of both martingale increment and remainder term.
- **ODE method**: Study of the SA as a discretization of

$$ODE_h \equiv \dot{\theta} = -h(\theta).$$

Assume that $(\gamma_n)_{n\geq 1}$ is a non-negative sequence such that

$$\sum_{n\geq 1} \gamma_n = +\infty \quad \text{and} \quad \sum_{n\geq 1} \gamma_n^2 < +\infty.$$

Then,

$$\theta_n \xrightarrow[n\to\infty]{a.s.} \theta^*.$$

Rate of convergence: CLT (SDE method)

Let θ^* be an equilibrium point of $\{h = 0\}$. Assume that the function h is differentiable at θ^* and that all the eigenvalues of $Dh(\theta^*)$ have a non-negative real part. Specify the step sequence as follows

$$\forall n \geq 1, \quad \gamma_n = \frac{\alpha}{n}, \quad \alpha > \frac{1}{2\mathcal{R}e(\lambda_{min})} \tag{1}$$

where λ_{min} denotes the eigenvalue of $Dh(\theta^*)$ with the lowest real part.

Then, the previous convergence $a.s.$ is ruled on $\{\theta_n \xrightarrow{a.s.} \theta^*\}$ by the following CLT

$$\sqrt{n}\,(\theta_n - \theta^*) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0, \alpha\Sigma)$$

with

$$\Sigma := \int_0^{+\infty} \left(e^{-\left(Dh(\theta^*)-\frac{I_d}{2\alpha}\right)u}\right)^t \Gamma e^{-\left(Dh(\theta^*)-\frac{I_d}{2\alpha}\right)u} du.$$

Outline

Outline

Intraday volume curve for optimal trade scheduling

Let split the trading day into N equal slices. Assume a trader has estimated from historical data a reference volume curve

$$(V_1^0, \dots, V_N^0)$$

where V_n^0 is the total (estimated) traded volume of the slice n. Then we deduce the estimated weights for the slice n

$$w_n^0 := \frac{V_n^0}{\sum_{k=1}^N V_k^0}, \quad n = 1, \dots, N.$$

At the end of the n^{th} slice, the real value of the traded volume V_n is known. Thus the weighted volume curve can be updated as follows

$$w_0 = w_0^0, \quad w_{n+1} = \frac{V_{n+1}^0 + (V_n - V_n^0)}{\sum_{k=1}^n V_k + \sum_{k=n+1}^N V_k^0} \quad n = 0, \dots, N-1,$$

and used in trading algorithm (for VWAP or IS for example) to update the trading curve.

Outline

Static modelling

The principle of a *Dark pool* is the following:

- It proposes a bid price with no guarantee of executed quantity at the occasion of an OTC transaction.
- Usually this price is lower than the bid price offered on the regular market.

So one can model the impact of the existence of N dark pools ($N \geq 2$) on a given transaction as follows:

- Let $V > 0$ be the random volume to be executed,
- Let $\theta_i \in (0, 1)$ be the *discount factor* proposed by the dark pool i.
- Let r_i denote the percentage of V sent to the dark pool i for execution.
- Let $D_i \geq 0$ be the quantity of securities that can be delivered (or mase available) by the dark pool i at price $\theta_i S$.

Cost of the executed order

The reminder of the order is to be executed on the regular market, at price S.

Then the **cost** C of the whole executed order is given by

$$\begin{aligned} C &= S\sum_{i=1}^{N}\theta_i \min\left(r_i V, D_i\right) + S\left(V - \sum_{i=1}^{N}\min\left(r_i V, D_i\right)\right) \\ &= S\left(V - \sum_{i=1}^{N}\rho_i \min\left(r_i V, D_i\right)\right) \end{aligned}$$

where

$$\rho_i = 1 - \theta_i \in (0,1), i = 1, \ldots, N.$$

Set $\mathcal{P}_N := \left\{ r = (r_i)_{1\le i\le N} \in \mathbb{R}_+^N \mid \sum_{i=1}^{N} r_i = 1 \right\}$.

Design of the stochastic algorithm

We us a Lagrangian approach to solve this maximization problem under constraints. We obtain that

$$r^* \in \arg\min_{\mathcal{P}_N} C$$
$$\Updownarrow$$
$$\forall i \in I_N, \quad \mathbb{E}\left[V\left(\rho_i \mathbb{1}_{\{r_i^* V < D_i\}} - \tfrac{1}{N}\textstyle\sum_{j=1}^N \rho_j \mathbb{1}_{\{r_j^* V < D_j\}}\right)\right] = 0.$$

Consequently, this leads to the following **recursive zero search procedure**

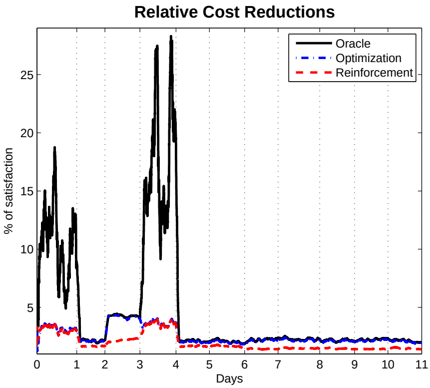
$$r_i^{n+1} = r_i^n + \gamma_{n+1} H_i(r^n, Y^{n+1}), \; r^0 \in \mathcal{P}_N, \; 1 \leq i \leq N, \qquad (2)$$

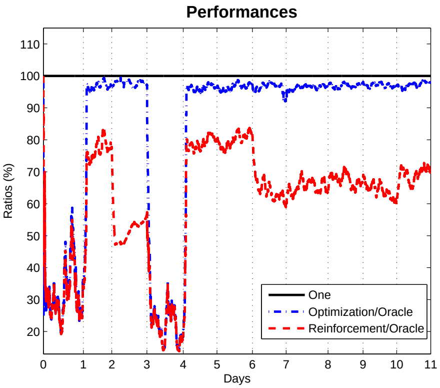
where $Y^n := (V^n, D_1^n, \ldots, D_N^n)$ and

$$H_i(r, Y) = V\left(\rho_i \mathbb{1}_{\{r_i V < D_i\}} - \frac{1}{N}\sum_{j=1}^N \rho_j \mathbb{1}_{\{r_j V < D_j\}}\right)$$

with $(Y^n)_{n \geq 1}$ is a sequence of random vectors with non negative components such that, for every $n \geq 1$ and $1 \leq i \leq N$, $(V^n, D_i^n) \stackrel{d}{=} (V, D_i)$.

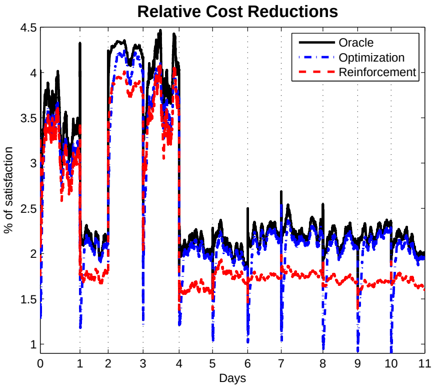
Long-term optimization

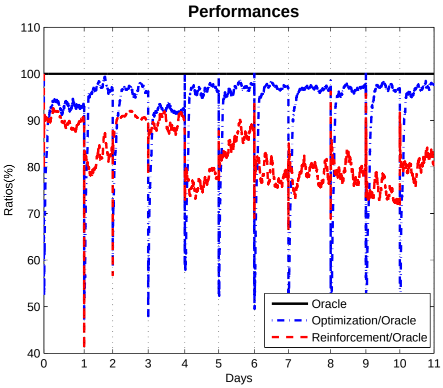




Daily resetting of the procedure

We reset the step γ_n at the beginning of each day and the satisfaction parameters and we keep the allocation coefficients of the preceding day.





Outline

1. Why Using Statistical Learning for Optimal Execution ?

2. Introduction to Stochastic Approximation Algorithms

3. Applications to Optimal Execution
   - Intraday volume curve for optimal trade scheduling
   - Optimal split of orders across liquidity pools
   - Optimal posting price of limit orders: learning by trading
   - Optimal split and posting price of limit orders across lit pools

Modeling and design of the algorithm

We consider on a short period T a **Poisson process of "execution"** of buy orders

$$\left(N_t^{(\delta)}\right)_{0\le t\le T} \quad \text{with intensity} \quad \Lambda_T(\delta, S) := \int_0^T \lambda(S_t - (S_0 - \delta))dt \tag{3}$$

where

- $0 \le \delta \le \delta_{\max}$ with $\delta_{\max} \in (0, S_0)$ denotes the *depth of the limit order book*,
- $(S_t)_{t\ge 0}$ is a stochastic process modeling the dynamic of the *fair price of a security stock*,
- the function λ is defined on the whole real line as a finite non increasing convex function.

Optimization Problem

Then we introduce a **market impact penalization function** $\Phi : \mathbb{R} \mapsto \mathbb{R}_+$, nondecreasing and convex, with $\Phi(0) = 0$ to model the additional cost of the execution of the remaining quantity.

Then the *resulting cost of execution* on a period $[0, T]$ reads

$$C(\delta) := \mathbb{E}\left[(S_0 - \delta)\left(Q_T \wedge N_T^{(\delta)}\right) + \kappa S_T \Phi\left(\left(Q_T - N_T^{(\delta)}\right)_+\right)\right] \quad (4)$$

where $\kappa > 0$.

Our aim is then to **minimize this cost**, namely to solve the following optimization problem

$$\min_{0 \leq \delta \leq \delta_{\max}} C(\delta). \quad (5)$$

To solve this optimization problem, we will devise a **stochastic algorithm constrained** to stay in $[0, \delta_{\max}]$.

Design of the algorithm

Once the two points are checked, we can devise the algorithm following the **standard stochastic approximation with projection**, namely

$$\delta_{n+1} = \text{Proj}_{[0,\delta_{\max}]}\left(\delta_n - \gamma_{n+1} H\left(\delta_n, \left(\bar{S}_{t_i}^{(n+1)}\right)_{0\le i\le m}\right)\right), \ \delta_0 \in [0, \delta_{\max}], \tag{6}$$

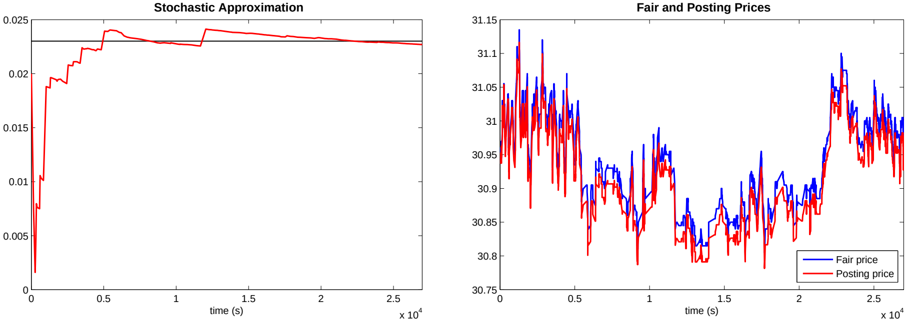
where

- $\text{Proj}_{[0,\delta_{\max}]}$ denotes the projection on $[0, \delta_{\max}]$,
- the positive step sequence $(\gamma_n)_{n\ge 1}$ satisfies at least the minimal *decreasing step* assumption

$$\sum_{n\ge 1} \gamma_n = +\infty \quad \text{and} \quad \gamma_n \to 0, \tag{7}$$

- $\left\{\left(\bar{S}_{t_i}^{(n)}\right)_{0\le i\le m}, n \ge 0\right\}$ is either a sequence of i.i.d. copies of the true underlying dynamics of $(S_{t_i})_{0\le i\le m}$ or at least of its Euler scheme or a sequence sharing some averaging properties (*e.g.* stationary α-mixing).

δ and posting price obtained by SA



Outline

Framework and Modelling

- Assume that a trader wants to buy a volume V of an asset across N lit pools with limit orders.
- She has to determinate the proportions $r = (r^i)_{1 \le i \le N}$ to sent to each lit pools and the posting prices $(S - \delta) = (S - \delta^i)_{1 \le i \le N} \in [0, \delta_{\max}]^N$.
- The execution flow at the distance δ^i of the reference price S is modelled by a random variable $Q^i(\delta^i) = \bar{Q}^i e^{-k^i \delta^i}$ where $\bar{Q}^i$ is a positive random variable modeling the executed quantity at the first limit an $k^i > 0$.
- If she is not fully executed, she sends a market order of the remaining quantity.

Mean Execution Cost

The **mean resulting cost of execution** is the **sum of each mean execution costs** on the lit pools, namely it reads

$$C(r,\delta) := \sum_{i=1}^{N} \mathbb{E}\left[(S-\delta^i)(r^i V \wedge Q^i(\delta^i)) + \kappa S(r^i V - Q^i(\delta^i))_+\right], \quad (8)$$

where $\kappa > 0$ is a free tuning parameter.

Our aim is then to **minimize this cost** by choosing the proportions and the distances to post at, namely to solve the following optimization problem

$$\min_{r \in \mathcal{P}_N, \delta \in [0,\delta_{\max}]^N} C(r,\delta). \quad (9)$$

We take advantage of the representation of C and its first two derivatives as expectations to devise a recursive stochastic algorithm.

Design of the stochastic algorithm

Based on a **Lagrangian approach** for the optimal proportions and on the **representations as expectations** for C' and C'', we can formally devise a recursive stochastic gradient descent

$$r_{n+1} = \mathrm{Proj}_{\mathcal{P}_N}\left(r_n - \gamma_{n+1} H(r_n, \delta_n, \bar{Q}_n e^{-k\delta_n})\right), \quad n \geq 0,$$

$$\delta_{n+1} = \mathrm{Proj}_{[0,\delta_{\max}]^N}\left(\delta_n - \gamma_{n+1} G(r_n, \delta_n, \bar{Q}_n e^{-k\delta_n})\right), \quad n \geq 0,$$

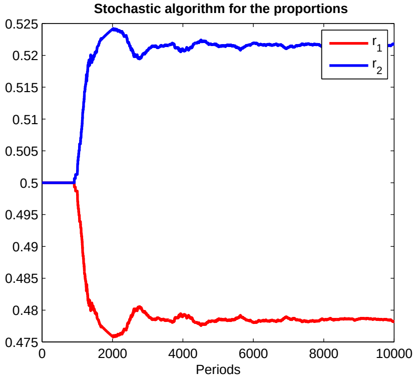
where, for every $i \in \{1, \dots, N\}$,

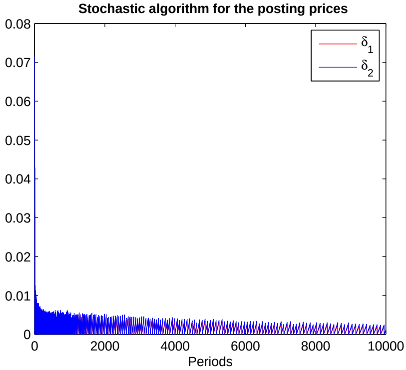
$$H(r_n^i, \delta_n^i, \bar{Q}_n^i e^{-k^i\delta_n^i}) = V\left((S - \delta_n^i)\mathbb{1}_{\{r_n^i V \leq \bar{Q}_n^i e^{-k^i\delta_n^i}\}} + \kappa S \mathbb{1}_{\{r_n^i V \geq \bar{Q}_n^i e^{-k^i\delta_n^i}\}}\right)$$
$$-\frac{1}{N}\sum_{j=1}^N V\left((S - \delta_n^j)\mathbb{1}_{\{r_n^j V \leq \bar{Q}_n^j e^{-k^j\delta_n^j}\}} + \kappa S \mathbb{1}_{\{r_n^j V \geq \bar{Q}_n^j e^{-k^j\delta_n^j}\}}\right),$$

and

$$G(r_n^i, \delta_n^i, \bar{Q}_n^i e^{-k^i\delta_n^i}) = -r_n^i V \wedge \bar{Q}_n^i e^{-k^i\delta_n^i} + k^i(\kappa S - (S - \delta_n^i))\mathbb{1}_{\{r_n^i V \leq \bar{Q}_n^i e^{-k^i\delta_n^i}\}}.$$

Convergence of the SA procedure





Some References

- M. Duflo, *Algorihtmes Stochastiques*, volume 23 of Math. and Applications (Berlin), Springer-Verlag, Berlin, 1996.
- H. J. Kushner, D. S. Clark, *Stochastic approximation methods for constrained and unconstrained systems*, volume 26 of *Applied Mathematical Sciences*. Springer-Verlag, New York, second edition, 1978.
- S. Laruelle, C.-A. Lehalle, G. Pagès, Optimal split of orders across liquidity pools: a stochastic algorithm approach, *SIAM Journal on Financial Mathematics*, 2(1):1042–1076, 2011.
- S. Laruelle, G. Pagès, Stochastic Approximation with Averaging Innovation Applied to Finance, *Monte Carlo Methods and Applications*, 18 (1):1–51, 2012.
- S. Laruelle, C.-A. Lehalle, G. Pagès, Optimal posting price of limit orders: learning by trading, *Mathematics and Financial Economics*, 7(3):359–403, 2013.
- S. Laruelle, Faisabilité de l'apprentissage des paramètres d'un algorithme de trading sur des données réelles, *Cahier de la Chaire Finance and Sustainable Development, Hors-Série Microstructure des marchés*, juin 2013.
- S. Laruelle, C.-A. Lehalle, *Market Microstructure in Practice*, World Scientific, 2013.
- S. Laruelle, Optimal split and posting price of limit orders across lit pools using stochastic approximation, *Initiative de Recherche "Microstructure des marchés" de Kepler-Cheuvreux*, sous la tutelle de l'Institut Europlace de Finance, avril 2014.

Thank you for your attention