Optimal trading algorithms and selfsimilar processes: a p-variation approach

Mauricio Labadie [1] Charles-Albert Lehalle [2]

December 10, 2010

Abstract

Almgren and Chriss (*Optimal execution of portfolio transactions.* Journal of Risk, Vol. 3, No. 2, 2010, pp. 5-39) and Lehalle (*Rigorous strategic trading: balanced portfolio and mean reversion.* Journal of Trading, Summer 2009, pp. 40-46.) developed optimal trading algorithms for assets and portfolios driven by a brownian motion. More recently, Gatheral and Schied (*Optimal trade execution under geometric brownian motion in the Almgren and Chriss framework.* Working paper SSRN, August 2010) addressed the same problem for the geometric brownian motion. In this article we extend these ideas for assets and portfolios driven by a discrete version of a selfsimilar process of exponent $H \in (0, 1)$, which can be either a fractional brownian motion of Hurst exponent H or a truncated Lévy distribution of index $1/H$.

The cost functional we use is not the classical expectation-variance one: instead of the variance, we use the p-variation, i.e. the l_p equivalent of the variance. We find explicitly the trading algorithm for any $p > 1$ and compare the resulting trading curve (that we call p-curve) with the classical expectation-variance curve (the 2-curve). If $p < 2$ we show that the p-curve is below the 2-curve at the beginning of the execution and above at the end of the execution. Therefore, we have a trading pattern that minimizes the market risk (i.e. the risk that the prices will drift away from its current level). On the other hand, if $p > 2$ then the p-curve is above the 2-curve at the beginning of the execution and below at the end. Therefore, this pattern minimizes the market impact.

We also show that the value of p in the p-variation is related to the exponent H of selfsimilarity via $p = 1/H$. In consequence, one can find the right value of p to put into the trading algorithm by calibrating the exponent H via real time series. We believe this result is interesting applications for high-frecuency trading.

[1] CAMS, EHESS, CNRS and University Pierre et Marie Curie (Paris VI). 54 Boulevard Raspail 75006 Paris, France. Email address: mauricio.labadie@gmail.com, labadie@ehess.fr

[2] Head of Quantitative Research, Crédit-Agricole Cheuvreux. 9 Quai du Président Paul Doumer 92400 Courbevoie (Paris La Défense), France. Email address: clehalle@cheuvreux.com

Contents

1 Introduction

1.1 Brownian motion and expectation-variance: a review

This section recalls the framework, notation and results in Almgren and Chriss [1] and Lehalle [8]. Suppose we want to trade an asset S throughout a time horizon $T > 0$. We will also suppose that we have already set the trading schedule, i.e. we will do N trades at times

$$0 = t_0 < t_1 < t_2 < \cdots < t_N = T.$$

Define

$$\tau_n = t_n - t_{n-1}$$

and assume that the price dynamics follows a brownian motion, i.e.

$$S_{n+1} = S_n + \sigma \tau_{n+1}^{1/2} \varepsilon_{n+1}, \tag{1}$$

where $\sigma > 0$ and $(\varepsilon_n)_{1 \leq n \leq N}$ are i.i.d. normal random variables of mean zero and variance 1. Following Almgren and Chriss [1] and Lehalle [8], we will model the temporary market impact as a function h depending on the average trading on each interval. More precisely, if

$$(\nu_1, \ldots, \nu_N), \qquad \sum_{n=1}^{N} \nu_n = 1$$

is the number of units we trade at each time n then

$$h(\nu_n) = \eta \frac{\sigma \nu_n}{V_n}, \qquad \eta > 0, \tag{2}$$

with V_n the historical volume of the asset at time n. This implies that the temporary market impact is directly proportional to the number of traded units ν_n and the volatility, but inversely proportional to the available volume. Under this framework, the wealth process (i.e. the full trading revenue upon completion of all trades) is

$$\begin{aligned} W &= \sum_{n=1}^{N} q_n \nu_n (S_n + q_n h(\nu_n)) \qquad (3) \\ &= \sum_{n=1}^{N} q_n \nu_n S_n + \sum_{n=1}^{N} S_n \frac{\eta \sigma (\nu_n)^2}{V_n}, \end{aligned}$$

where $q_n = 1$ if we buy at time n and $q_n = -1$ if we sell.

The permanent market impact will no be considered here for two reasons. On the one hand, the permanent impact can be added without structural changes on the model, as long as this impact is linear (Lehalle [8]). On the other hand, we believe that there is no clear definition, nor consensus, on what is the *fundamental* price of an asset (Bouchaud and Potters [4]), and on consequence the definition of permanent market impact is rather ambiguous.

For long-only portfolios (i.e. $q_n = +1$) the wealth process as a function of $(x_1, \ldots, x_n)$ takes the form

$$W(x_1, \ldots, x_n) = S_0 + \sum_{n=1}^{N} \sigma \tau_n^{1/2} \varepsilon_n x_n + \sum_{n=1}^{N} \frac{\eta\sigma}{V_n}(x_n - x_{n+1})^2. \tag{4}$$

The expectation and variance of the wealth process (4) as functions of x_n are

$$\begin{aligned} \mathbb{E}(W) &= S_0 + \sum_{n=1}^{N} \frac{\eta\sigma}{V_n}(x_n - x_{n+1})^2, \\ \mathrm{Var}(W) &= \sum_{n=1}^{N} \sigma^2 \tau_n x_n^2 \,. \end{aligned} \tag{5}$$

Therefore, the corresponding cost functional for a level of risk aversion λ is

$$\begin{aligned} J_\lambda(x_1, \ldots, x_N) &= \mathbb{E}(W) + \lambda \mathrm{Var}(W) \\ &= S_0 + \sum_{n=1}^{N} \frac{\eta\sigma}{V_n}(x_n - x_{n+1})^2 + \lambda \sum_{n=1}^{N} \sigma^2 \tau_n x_n^2 \,. \end{aligned} \tag{6}$$

We compute the partial derivative of J_λ:

$$\frac{\partial J_\lambda}{\partial x_n} = -\frac{2\eta\sigma}{V_{n-1}}(x_{n-1} - x_n) + \frac{2\eta\sigma}{V_n}(x_n - x_{n+1}) + 2\lambda\sigma^2\tau_n x_n. \tag{7}$$

Equating (6) to zero we obtain the optimal trading curve via the recursive algorithm

$$x_{n+1} = \left(1 + \frac{\mathcal{D}_{n-1}}{\mathcal{D}_n} + \lambda \frac{\sigma^2 \tau_n}{\mathcal{D}_n}\right) x_n - \frac{\mathcal{D}_{n-1}}{\mathcal{D}_n} x_{n-1} \,, \quad \mathcal{D}_n = \frac{\eta\sigma}{V_n} \,, \tag{8}$$

under the constraints $x_0 = 1$ and $x_{N+1} = 0$.

1.2 Selfsimilarity and Lévy processes

This section recalls the basic definitions and properties of selfsimilar and Lévy processes (see Embrechts and Maejima [5] for more details).

- An $\mathbb{R}^K$-valued stochastic processes $\{S(t) : t \geq 0\}$ is *selfsimilar* if for any $a > 0$ there exists $b > 0$ such that $\{S(at)\} = \{bS(t)\}$ in distribution.
- $\{S(t) : t \geq 0\}$ is *stochastically continuous* if for any $\varepsilon > 0$,

$$\lim_{h \to 0} \mathbb{P}\{|S(t+h) - S(t)| > \varepsilon\} = 0.$$

  In particular, a continuous process is stochastically continuous.
- If $\{S(t) : t \geq 0\}$ is selfsimilar and stochastically continuous then there exists a unique $H \geq 0$ such that $b = a^H$, i.e. $\{S(at)\} = \{a^H S(t)\}$ in distribution. We call H the *exponent of self-similarity* and $\{S(t) : t \geq 0\}$ a *H-selfsimilar process.*

- If $H = 1/2$ we recover the classical brownian motion.
- An $\mathbb{R}^K$-valued stochastic processes $\{S(t) : t \geq 0\}$ is a *Lévy process* if
  - $S(0) = 0$ a.s.
  - It is stochastic continuous at any $t \geq 0$.
  - It has independent and stationary increments.
  - It sample paths are right-continuous and have left limits a.s.
- A probability measure m on $\mathbb{R}^K$ is *stable* if for any $a > 0$ there exists $b > 0$ such that

$$\hat{m}(y)^a = \hat{m}(by) \quad \forall y \in \mathbb{R}^K.$$

  where $\hat{m}$ is the Fourier transform of m.
- If m is stable then there exists a unique $p \in (0, 2]$ such that $b = a^{1/p}$, i.e.

$$\hat{m}(y)^a = \hat{m}(a^{1/p} y) \quad \forall y \in \mathbb{R}^K.$$

  Such probability measure m is called *p-stable*.
- Let $p \in (0, 2]$ and let Z be an $\mathbb{R}^K$-valued random variable with a p-distribution. Then for any $\gamma \in (0, p)$ we have $\mathbb{E}[|Z|^\gamma] < \infty$, but $\mathbb{E}[|Z|^p] = \infty$.
- Let $\{S(t) : t \geq 0\}$ be a Lévy process. Then $\{S(t) : t \geq 0\}$ is H-selfsimilar if and only if $X(1)$ is μ-stable. Moreover, μ and H satisfy $\mu = 1/H$.
- Stable Lévy processes are the only selfsimilar processes with independent and stationary increments.
- If $\mu = 2$ we recover the classical brownian motion.

1.3 Fractional brownian motion

This section recalls the definition and properties of the fractional brownian motion (see Embrechts and Maejima [5], Bouchaud and Potters [4] and Mandelbrot [9] for more details).

- Let $H \in (0, 1]$. The *fractional brownian motion* $\{B^H(t) : t \geq 0\}$ is a real-valued gaussian process such that

$$\begin{aligned} \mathbb{E}[B(t)] &= 0\,, \\ \mathbb{E}[B(t)B(s)] &= \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right)\,. \end{aligned}$$

- If $H = 1/2$ the returns are not correlated and B^H is the classical brownian motion (also called *white noise*).
- If $H > 1/2$ the returns are positively correlated (and the process is called *red noise*). This implies that the asset has a trend.

- If $H < 1/2$ the returns are negatively correlated (*blue noise*). This implies that the asset has a mean-reverting dynamics.

The idea of using power law distribution instead of gaussian ones is not a new one. One of the reasons is that gaussian distributions fail to reproduce several properties of the actual distributions of returns. This properties, known as *stylized facts*, are common of any asset class: the real distribution of returns is approximately symmetric, it has fat tails (i.e. extreme returns are more likely to happen that the normal theory would forecast), it has a high peak (i.e. *leptokurtic*), there is weak autocorrelation between returns but a high autocorrelation in the absolute value of square returns. The last point implies that volatility varies over time and presents clustering patterns, i.e. there are periods of high volatility and low volatility (see Bouchaud and Potters [4] and Embrechts *et al* [11] for more details).

The stylized facts can be reproduced using different distributions such as (truncated) Lévy, Student or fractional brownian motion. Moreover, there is empirical evidence of the selfsimilar behavior of returns. Indeed, Almgren *et al* [3] analyzed real data and found that market impact is a power law of exponent 3/5 of block size, with specific dependence on trade duration, daily volume and volatility. For all this reasons, a model based on sefsimilar processes is a natural step towards the understanding of the real dynamics of market prices.

2 The p-variation model

2.1 Definition and properties of the p-variation

Let $p > 1$ and $y = (y_1, \ldots, y_n) \in \mathbb{R}^N$ be a random vector of mean zero. We define the *p-variation* of y as

$$\mathbb{V}_p(y) = \sum_{n=1}^{N} \mathbb{E}\,[\,|\,y_n|^p\,]\,.$$

The p-variation $\mathbb{V}_p(y)$ and the l_p-norm in $\mathbb{R}^N$ are related via

$$\|y\|_p = \mathbb{V}_p(y)^{1/p}.$$

Notice that if $y = (y_1, \ldots, y_n)$ is a time series of i.i.d. random variables then the 2-variation reduces to the variance, i.e.

$$\mathbb{V}_2(y) = \mathrm{Var}(y)\,.$$

Moreover, it is easy to show that the p-variation defines a metric on $\mathbb{R}^N$, and since all norms in $\mathbb{R}^N$ are equivalent there exist $0 < \alpha < \beta$ such that

$$\alpha\|y\|_p \le \|y\|_2 \le \beta\|y\|_p.$$

Therefore, the variance (i.e. the 2-variation) and the p-variation are two equivalent metrics on $\mathbb{R}^N$ such that

$$\mathbb{V}_p(y) \sim \mathbb{V}_2(y)^{p/2}\,. \tag{9}$$

For real-life applications we have $\mathbb{V}_2(y) < 1$. Since $p \mapsto \mathbb{V}_2(y)^{p/2}$ is decreasing, we have the following statements:

- If $p < 2$ the p-variation $\mathbb{V}_p(y)$ amplifies the effect of the variance.
- If $p > 2$ the p-variation reduces the effect of the variance.

Now let us define the p-variation for a special family of functions of random variables. Let $y = (y_1, \ldots, y_N)$ a random vector of mean zero. Consider the function $F : \mathbb{R}^N \to \mathbb{R}$ defined as

$$F(y) = \sum_{n=1}^{N} y_n.$$

We define the p-variation of F as

$$\mathbb{V}_p(F) = \sum_{n=1}^{N} \mathbb{E}\left[\, |\, y_n|^p \,\right].$$

Observe that if $y = (y_1, \ldots, y_N)$ is a time series of mean zero then $\mathbb{V}_p(F)$ is the p-th moment of the time series y. Finally, for general functions F such that

$$F - \mathbb{E}(F) = \sum_{n=1}^{N} y_n$$

we define their p-variation as

$$\mathbb{V}_p(F) := \mathbb{V}_p(F - \mathbb{E}(F)) = \mathbb{V}_p(y).$$

2.2 Optimal trading algorithms

Let $H \in (0, 1)$ and assume that the price dynamics is

$$S_{n+1} = S_n + \sigma \tau_{n+1}^H \varepsilon_{n+1}, \tag{10}$$

where $(\varepsilon_n)_{1 \le n \le N}$ are identically-distributed random variables such that $\mathbb{E}\left[\varepsilon_n\right] = 0$. Under these assumptions, the wealth process is

$$W(x_1, \ldots, x_n) = S_0 + \sum_{n=1}^{N} \sigma \tau_n^H \varepsilon_n x_n + \sum_{n=1}^{N} \frac{\eta\sigma}{V_n} (x_n - x_{n+1})^2 . \tag{11}$$

Its corresponding expectation and p-variation are

$$\begin{aligned} \mathbb{E}(W) &= S_0 + \sum_{n=1}^{N} \frac{\eta\sigma}{V_n} (x_n - x_{n+1})^2, \qquad (12) \\ \mathbb{V}_p(W) &= \mathbb{V}_p(W - \mathbb{E}(W)) \\ &= \sum_{n=1}^{N} \sigma^p \tau_n^{pH} |x_n|^p . \end{aligned}$$

The corresponding cost functional is

$$\begin{aligned} J_p(x_1, \ldots, x_N) &= \mathbb{E}(W) + \lambda \mathbb{V}_p(W) \qquad (13) \\ &= S_0 + \sum_{n=1}^{N} \frac{\eta\sigma}{V_n} (x_n - x_{n+1})^2 + \lambda \sum_{n=1}^{N} \sigma^p \tau_n^{pH} |x_n|^p . \end{aligned}$$

In order to have a well-defined cost functional J_p in terms of units, we need to have the same time units in both $\mathbb{E}(W)$ and $\mathbb{V}_p(W)$. But notice that $\mathbb{E}(W)$ has linear time units, whereas $\mathbb{V}_p(W)$ has t^{pH} units. In consequence, the only choice for p that renders the p-variance linear in time is $p = 1/H$. Therefore, assuming $H = 1/p$ and

$$E[|\varepsilon_n|^p] = 1 \quad \forall n = 1, \dots, N \tag{14}$$

we obtain

$$\begin{aligned} J_p(x_1, \dots, x_N) &= \mathbb{E}(W) + \lambda \mathbb{V}_p(W) \\ &= S_0 + \sum_{n=1}^{N} \frac{\eta\sigma}{V_n}(x_n - x_{n+1})^2 + \lambda \sum_{n=1}^{N} \sigma^p \tau_n |x_n|^p \,. \end{aligned} \tag{15}$$

The partial derivative of J_p is

$$\frac{\partial J_\lambda}{\partial x_n} = -\frac{2\eta\sigma}{V_{n-1}}(x_{n-1} - x_n) + \frac{2\eta\sigma}{V_n}(x_n - x_{n+1}) + p\lambda\sigma^p\tau_n x_n |x_n|^{p-2}. \tag{16}$$

Equating (16) to zero yields the optimal trading curve via the recursive algorithm

$$x_{n+1} = \left(1 + \frac{\mathcal{D}_{n-1}}{\mathcal{D}_n}\right) x_n - \frac{\mathcal{D}_{n-1}}{\mathcal{D}_n} x_{n-1} + \frac{p\lambda\sigma^p\tau_n}{2\mathcal{D}_n} x_n |x_n|^{p-2}\,, \quad \mathcal{D}_n = \frac{\eta\sigma}{V_n}\,, \tag{17}$$

with the constraints $x_0 = 1$ and $x_{N+1} = 0$. Notice that (17) is a well-defined algorithm for $p > 1$.

We would like to remark that the idea of a risk measure that is linear in time was already introduced by Gatheral and Schied [6]. The risk measure they chose was the expectation of the time-average, whereas our risk measure is a l^p version of the variance. In both cases, the risk measure has the same time units as the expectation of the process. However, in our case there is no a priori on the dynamics of the price process: the right choice of p is a consequence of finding empirically the exponent H of selfsimilarity and using the relation $p = 1/H$.

2.3 Comparison between 2-variation and p-variation algorithms

Let us study the difference between the p-variation algorithm (17) and the 2-variation algorithm (8)

$$\delta_{n+1}(p) = x_{n+1}(p) - x_{n+1}(2) = \frac{\lambda x_n \tau_n}{\mathcal{D}_n}\left(\frac{p}{2}\sigma^p |x_n|^{p-2} - \sigma^2\right). \tag{18}$$

Proposition 1 *Let $\mathbf{x}(p) = (x_0(p), \dots, x_{N+1}(p))$ be the optimal trading curve defined by the p-variation algorithm (17). Suppose that $x_0(p) = 1$ and $x_1(p) = \alpha$ is constant for all p. If $\sigma \le e^{-1/2} \approx 0.6065$ then the mapping*

$$p \mapsto \delta_{n+1}(p)$$

is decreasing for all $n = 2, \dots, N$.

Proof: Suppose $p > 2$. The function

$$f(x, \sigma, p) := \frac{p}{2}\sigma^p |x|^{p-2} - \sigma^2$$

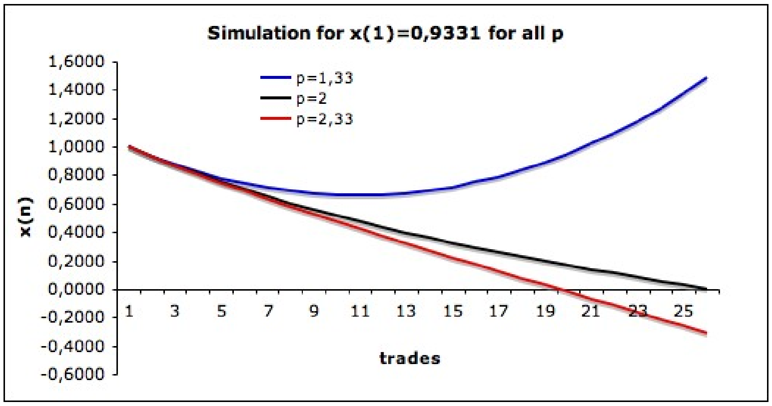

Figure 1: Comparing the p-curves for the same value of $x_1(p)$. As it was shown in Proposition 1, the mapping $p \mapsto \mathbf{x}(p)$ with $x_1(p) = \alpha$ constant for all p is decreasing.

is increasing in $|x|$. Therefore,

$$f(x, \sigma, p) \leq f(1, \sigma, p) \quad \text{for all } x \in (0, 1],\ \sigma > 0,\ p > 2.$$

But notice that

$$f(1, \sigma, p) \leq 0 \iff \sigma \leq \sigma^* := \left(\frac{2}{p}\right)^{\frac{1}{p-2}}.$$

After some straightforward computations from elementary Calculus, it can be shown the function

$$p \mapsto \left(\frac{2}{p}\right)^{\frac{1}{p-2}}$$

is increasing in $p > 2$ and that its minimum is $e^{-1/2} \approx 0.6065$. The case $p < 2$ can be treated similarly. Indeed,

$$f(x, \sigma, p) := \frac{p}{2}\sigma^p |x|^{p-2} - \sigma^2$$

is decreasing in $|x|$, which implies that

$$f(x, \sigma, p) \geq f(1, \sigma, p) \quad \text{for all } x \in (0, 1],\ \sigma > 0,\ p > 2.$$

Therefore

$$f(1, \sigma, p) \geq 0 \iff \sigma \leq \sigma^* := \left(\frac{p}{2}\right)^{\frac{1}{2-p}}.$$

As before, it can be shown that

$$p \mapsto \left(\frac{p}{2}\right)^{\frac{1}{2-p}}$$

is decreasing in $p < 2$ and that its minimum is $e^{-1/2} \approx 0.6065$. $\square$

Proposition 2 *Let $1 < q < p$, $\sigma \le e^{-1/q}$ and suppose that $x_0(p) = x_0(q) = 1$ and $x_1(p) = x_1(q) = \alpha$. Then $\mathbf{x}(q) \ge \mathbf{x}(p)$, i.e. the p-curves are monotone decreasing (see Fig. 1). In particular, the result is independent of the choice of p and q provided $\sigma \le e^{-1} \approx 0.3678$.*

Proof: The difference between the p and q-curves is

$$f(x,\sigma,p,q) = \frac{p}{2}\sigma^p|x|^{p-2} - \frac{q}{2}\sigma^q|x|^{q-2}\,.$$

Therefore,

$$f(x,\sigma,p,q) \le 0 \iff \frac{p}{q}\sigma^{p-q}|x|^{p-q} \le 1 \quad \text{for all } x \in (0,1].$$

Fix $q > 1$. It can be shown with some Calculus that the mapping

$$q \mapsto \left(\frac{q}{p}\right)^{\frac{1}{p-q}}$$

is increasing and that its minimum is $e^{-1/q}$. In conclusion,

$$f(x,\sigma,p,q) \le 0 \quad \text{for all } x \in (0,1],\ \sigma \le e^{-1/q} \text{ and } p > q > 1. \quad \square$$

3 Numerical results

3.1 The p-curves are increasing in p

Let $\mathbf{x}(p) = (x_0(p),, \ldots, x_{N+1}(p))$ be the optimal trading curve defined by the p-variation algorithm (17). Suppose that $x_1(p)$ is chosen in such a way that $x_{N+1}(p) = 0$. In that framework, our numerical simulations show that the p-curves $\mathbf{x}(p)$ is increasing in p (see Fig. 2). This confirms the intuition given in Section 2.1 about the p-variation, i.e. that p is a "tuning parameter" in the sense that it amplifies (resp. reduces) the effect of the variance when $p < 2$ (resp. when $p > 2$).

3.2 The p-variation algorithm is robust with respect to H

Proposition 3 *Let $\pi(e)$ as the error in the estimate p when the estimate of H has an error of e. Then $\pi(e)$ is independent of p and H and*

$$\pi(e) \le \frac{e}{1-e}\,.$$

Proof: Let $\hat{H} = H + \varepsilon H$ be an estimate of H, where ε is the signed relative error. Then the error $\pi(\varepsilon)$ is independent of H and p, i.e.

$$\pi(\varepsilon) = \left|\frac{\hat{p}-p}{p}\right| = \left|\frac{\varepsilon}{1-\varepsilon}\right|\,.$$

In consequence, if we define $e := |\varepsilon|$ then

$$\pi(e) \le \frac{e}{1-e}\,. \quad \square$$

Now let $R(e,p)$ be the maximum error in the p-curves when the estimate of H has an error of e. In our numerical experiments (see Table 1) we found that the error ratio $R(e,p)/e$ is smaller than one. This

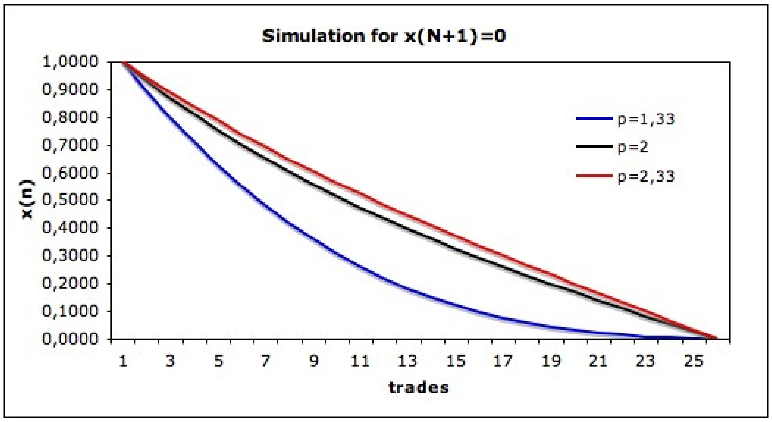

Figure 2: Comparing the p-curves. According to our numerical experiments, the mapping $p \mapsto \mathbf{x}(p)$ is increasing when $x_1(p)$ is chosen in such a way that $x_{N+1} = 0$.

implies that the p-variation algorithm has somewhat a "regularising" effect in the sense that the error in the p-curve is smaller that the error in the estimate of H. Moreover, the mapping

$$p \mapsto \frac{R(e,p)}{e}$$

is monotone decreasing. This was expected since the case $p = 1$ is singular, and as p increases we are further from this (numerical) instability. In conclusion, the p-variation algorithm (17) is robust with respect to H, and this robustness improves as p increases.

p	Error in H e	Error in p-curve $R(e,p)$	Error ratio $R(e,p)/e$
2.33	1%	0.29%	0.29
	5%	1.6%	0.32
	10%	3.74%	0.374
1.33	1%	0.68%	0.68
	5%	3.54%	0.708
	10%	7.58%	0.758

Table 1: Numerical results for the errors. As we can see, the errors in the calibration of H are smoothed out by the p-variation algorithm, and this smoothing or regularising effect is more important as p increases.

4 Relation between p-variation algorithms and selfsimilar processes

4.1 Interpretation in terms of Lévy processes

Observe that the discrete model (10) is selfsimilar of exponent $H = 1/p$. One would like to show that (10) can be the discretization of a familiar continuous stochastic process.

If (10) were the discrete version of a continuous, selfsimilar random process with i.i.d. increments then the process would be necessarily a stable Lévy process. However, any H-selfimilar, Lévy process is p-stable Lévy, where $p = 1/H$, and one of their properties is that the p-th moment is infinite, whereas in the p-variance we implicitly assume that the p-th moment is finite. In consequence, (10) cannot be the discretization of a Lévy process.

In order to overcome the integrability issues of the stable Lévy distribution, we can consider that the prices have a *truncated* Lévy distribution, i.e. a p-stable Lévy distribution for finite values, but either zero (see Mantegna and Stanley [10]) or exponential distribution (see Koponen [7]) for large values. Under this conditions, it follows that the model (10) is "locally" (i.e. inside a compact set in time and space) the discretization of a selfsimilar, Lévy process of index p and the hypothesis of p-th moments is satisfied.

4.2 Interpretation in terms of fractional brownian motion

Unlike stable Lévy processes, the fractional brownian motion has finite moments of all orders and its increments are autocorrelated. However, if the increments in the model (10) are not independent then the p-variation does not coincide with the p-th moment of a fractional brownian motion, since the latter does not take into account the autocorrelations. Nevertheless, we could think that the the p-variance is a first-order approximation of the p-th moment, in which we are negecting the effect of autocorrelations. The good news are that this approximation reproduces the trading patterns that one could expect from a fractional brownian motion:

- If $p < 2$ (i.e. $H > 1/2$) we have positive autocorrelation, i.e. prices have a trend. Therefore, our trading will move prices in the wrong sense for us: if we sell prices will go down, whereas if we buy prices will go up. In consequence, since there is a bigger market risk than for the classical brownian motion, we would like to execute the order faster in order to minimize this risk.
- If $p > 2$ (i.e. $H < 1/2$) we have negative autocorrelation, i.e. prices follow a mean-reverting dynamics. In consequence, since the market risk is less important than for the classical brownian motion, the algorithm focusses on minimizing the market impact, which implies that we trade slower.

4.3 Real trading applications

The p-variation model (10) can be regarded as a generalization of the classical expectation-variance approach, for which there is a parameter p that has a qualitative impact on the trading patterns. Observe that when choosing a value of p and writing the cost functional in terms of the p-variation, we are implicitly assuming that we the price process follows a selfsimilar process of exponent $H = 1/p$ (the classical brownian motion is the particular case $p = 2$). Regardless of the continuous dynamics of (10), i.e. whether it is a truncated Lévy processes or a fractional brownian motion, the right value of p to put into the algorithm can be calibrated empirically using time series: we choose the exponent H as the best fit for the logarithm of the price process, i.e.

$$\log(S(t_n) - S(t_{n-1})) = \alpha + H \times \log(\tau_n) + \varepsilon_n, \quad \tau_n = t_n - t_{n-1}.$$

5 Extensions of the model to portfolios

5.1 General portfolios

This section is a generalization of the multi-asset case in Almgren and Chriss [1] and Lehalle [8] to a selfsimilar process. Suppose we have a portfolio of K assets $\mathbf{S} = (S^1, \ldots, S^K)$. For any n we define the asset vector $\mathbf{S}_n = (S_n^1, \ldots, S_n^K)$, whose dynamics is supposed to be a fractional brownian motion, i.e.

$$\mathbf{S}_{n+1} = \mathbf{S}_n + \tau_n^{1/p} \mathcal{E}_{n+1}, \tag{19}$$

where $(\mathcal{E}_n)_{1 \le n \le N}$ are i.i.d. K-dimensional Gaussian vectors with mean zero and covariance matrix

$$[\mathfrak{S}_n]_{ij} = \mathbb{E}\left[\mathcal{E}_n^i \mathcal{E}_n^j\right] .$$

The wealth function is

$$W = \sum_{k=1}^K S_0^k + \sum_{k=1}^K \sum_{n=1}^N \tau_n^{1/p} \mathcal{E}_n^k x_n^k + \sum_{k=1}^K \sum_{n=1}^N \eta^k \frac{\sigma^k}{V_n^k} (x_n^k - x_{n+1}^k)^2. \tag{20}$$

Its corresponding expectation and p-variation are

$$\begin{aligned} \mathbb{E}(W) &= \sum_{k=1}^K S_0^k + \sum_{k=1}^K \sum_{n=1}^N \eta^k \frac{\sigma^k}{V_n^k} (x_n - x_{n+1})^2, \\ \mathbb{V}_p(W) &= \sum_{n=1}^N \tau_n \, \mathbf{x}_n(p)' \mathfrak{S}_n(p) \mathbf{x}_n(p), \end{aligned} \tag{21}$$

where

$$\mathbf{x}_n(p) = (|x_n^1|^{p/2}, \ldots, |x_n^K|^{p/2})$$

and

$$[\mathfrak{S}_n(p)]_{ij} = \mathbb{E}\left[sign(\mathcal{E}_n^i \mathcal{E}_n^j) \, |\mathcal{E}_n^i|^{p/2} |\mathcal{E}_n^j|^{p/2} \right]$$

(notice that for $p = 2$ we recover the classical variance-covariance matrix). The cost functional is

$$\begin{aligned} J_p(\mathbf{x}_1, \ldots, \mathbf{x}_N) &= \mathbb{E}(W) + \lambda \mathbb{V}_p(W) \\ &= \sum_{k=1}^K S_0^k + \sum_{n=1}^N \left(\lambda \tau_n \, \mathbf{x}_n(p)' \mathfrak{S}_n(p) \mathbf{x}_n(p) + \sum_{k=1}^K \eta^k \frac{\sigma^k}{V_n^k} (x_n - x_{n+1})^2 \right). \end{aligned} \tag{22}$$

The corresponding recursive algorithm is still of the form

$$\mathbf{x}_{n+1} = F(\mathbf{x}_n, \mathbf{x}_{n-1}),$$

where $\mathbf{x}_0$ is given and $\mathbf{x}_{N+1} = (0, \ldots, 0)$. However, the explicit formula is complicated because it involves the partial derivatives of $\mathbf{x}_n(p)' \mathfrak{S}_n(p) \mathbf{x}_n(p)$, which do not have a linear expression for $p \neq 2$.

5.2 Balanced portfolios for brownian motion: a review

Here we recall a particular case of a multi-asset gaussian case, the so-called *balanced portfolio*, that appears in Lehalle [8]. Let us suppose that the price dynamics is a classical brownian motion, i.e.

$$\mathbf{S}_{n+1} = \mathbf{S}_n + \tau_n^{1/2} \mathcal{E}_{n+1},$$

where $(\mathcal{E}_n)_{1 \le n \le N}$ are i.i.d. K-dimensional Gaussian vectors with mean zero and covariance matrix

$$[\mathfrak{S}_n]_{ij} = \mathbb{E}\left[\mathcal{E}_n^i \mathcal{E}_n^j\right] ,$$

and that there is a unique trading strategy

$$(\tilde{\nu}_1, \dots, \tilde{\nu}_N), \qquad \sum_{n=1}^{N} \tilde{\nu}_n = 1 \tag{23}$$

such that the trading schedule of any asset k is a constant multiple of it. More precisely, we will assume that for any $k = 1, \dots, K$ there exists $\pi^k > 0$ such that

$$\nu_n^k = \pi^k \tilde{\nu}_n \qquad \text{for all } n = 1, \dots, N.$$

The number π^k is the total number of shares of asset k to be traded. Under this framework we also have $x_n^k = \pi^k \tilde{x}_n$. Indeed,

$$\begin{aligned} x_n^k &= \sum_{i=n}^{N} \nu_i^k = \sum_{i=n}^{N} \pi^k \tilde{\nu}_i \\ &= \pi^k \sum_{i=n}^{N} \tilde{\nu}_i = \pi^k \tilde{x}_n. \end{aligned} \tag{24}$$

Define $\pi = (\pi^1, \dots, \pi^K)$. From (23) and (24) it follows that the wealth process (20) for a balanced portfolio takes the form

$$W = \sum_{k=1}^{K} S_0^k + \sum_{n=1}^{N} \tau_n^{1/2} \left(\sum_{k=1}^{k} \mathcal{E}_n^k \pi^k \right) \tilde{x}_n + \sum_{n=1}^{N} (\tilde{x}_n - \tilde{x}_{n+1})^2 \left(\sum_{k=1}^{K} \eta^k (\pi^k)^2 \frac{\sigma^k}{V_n^k} \right) ,$$

Therefore, the expectation and variance are

$$\begin{aligned} \mathbb{E}(W) &= \sum_{k=1}^{K} S_0^k + \sum_{n=1}^{N} \mathcal{D}_n^\pi (\tilde{x}_n - \tilde{x}_{n+1})^2, \qquad (25) \\ Var(W) &= \sum_{n=1}^{N} \tau_n \, \pi' \mathfrak{S}_n \pi \, (\tilde{x}_n)^2 , \end{aligned}$$

where as before

$$\mathcal{D}_n^\pi := \sum_{k=1}^{K} \eta^k (\pi^k)^2 \frac{\sigma^k}{V_n^k} .$$

In consequence, the cost functional is

$$J_\lambda(\tilde{x}_1,\ldots,\tilde{x}_N) = \sum_{k=1}^{K} S_0^k + \lambda \sum_{n=1}^{N} \tau_n\, \pi' \mathfrak{S}_n \pi (\tilde{x}_n)^2 + \sum_{n=1}^{N} \mathcal{D}_n^\pi (\tilde{x}_n - \tilde{x}_{n+1})^2.$$

Therefore, the optimal trading curve is the solution of the recursive algorithm

$$\tilde{x}_{n+1} = \left(1 + \frac{\mathcal{D}_{n-1}^\pi}{\mathcal{D}_n^\pi} + \lambda \frac{\tau_n\, \pi' \mathfrak{S}_n \pi}{\mathcal{D}_n^\pi}\right) \tilde{x}_n - \frac{\mathcal{D}_{n-1}^\pi}{\mathcal{D}_n^\pi} \tilde{x}_{n-1}\,, \tag{26}$$

under the constraints $\tilde{x}_1 = 1$ and $\tilde{x}_{N+1} = 0$.

5.3 Balanced portfolios for selfsimilar processes: the p-variation model

We will extend the balanced portfolios for the p-variation model. Suppose that the portfolio follows the dynamics in (19) and define

$$\pi(p) := (|\pi^1|^{p/2}, \ldots, |\pi^K|^{p/2})\,.$$

Using (23) and (24) it can be shown that the wealth process (20) for a balanced portfolio takes the form

$$W = \sum_{n=1}^{N} S_0^n + \sum_{n=1}^{N} \tau_n^{1/p} \left(\sum_{k=1}^{k} \mathcal{E}_n^k \pi^k \right) \tilde{x}_n + \sum_{n=1}^{N} \mathcal{D}_n^\pi (\tilde{x}_n - \tilde{x}_{n+1})^2\,.$$

From (21) it follows that the expectation and p-variation are

$$\begin{aligned} \mathbb{E}(W) &= \sum_{k=1}^{K} S_0^k + \sum_{n=1}^{N} \mathcal{D}_n^\pi (x_n - x_{n+1})^2, \\ \mathbb{V}_p(W) &= \sum_{n=1}^{N} \tau_n\, \pi(p)' \mathfrak{S}_n(p) \pi(p)\, |\tilde{x}_n|^p\,, \end{aligned} \tag{27}$$

where (as before)

$$[\mathfrak{S}_n(p)]_{ij} = \mathbb{E}\left[sign(\mathcal{E}_n^i \mathcal{E}_n^j)\, |\mathcal{E}_n^i|^{p/2} |\mathcal{E}_n^j|^{p/2} \right]\,.$$

In consequence, the cost functional is

$$J_p(\tilde{x}_1,\ldots,\tilde{x}_N) = \sum_{k=1}^{K} S_0^k + \lambda \sum_{n=1}^{N} \tau_n\, \pi(p)' \mathfrak{S}_n(p) \pi(p)\, |\tilde{x}_n|^p + \sum_{n=1}^{N} \mathcal{D}_n^\pi (\tilde{x}_n - \tilde{x}_{n+1})^2\,.$$

In consequence, the corresponding recursive algorithm is

$$\tilde{x}_{n+1} = \left(1 + \frac{\mathcal{D}_{n-1}^\pi}{\mathcal{D}_n^\pi}\right) \tilde{x}_n - \frac{\mathcal{D}_{n-1}^\pi}{\mathcal{D}_n^\pi} \tilde{x}_{n-1} + \frac{p\lambda\tau_n\, \pi(p)' \mathfrak{S}_n(p) \pi(p)}{2\mathcal{D}_n^\pi}\, \tilde{x}_n |\tilde{x}_n|^{p-2}\,, \tag{28}$$

with the constraints $\tilde{x}_1 = 1$ and $\tilde{x}_{N+1} = 0$. Notice that (28) is well-defined for $p > 1$.

6 Conclusions

- Equation (1) can be considered as the discretization of a selfsimilar process. If $(\varepsilon_n)_{1\leq n\leq N}$ are i.i.d. then its continuous version $S(t)$ can be a stable Lévy process of exponent $p = 1/H$. However, for a p-stable Lévy process we have $\mathbb{E}[|S(t)|^p] = \infty$. Nevertheless, if we consider truncated Lévy processes we can ensure that the p-moment is finite.

- If $(\varepsilon_n)_{1\leq n\leq N}$ are gaussian and correlated then $S(t)$ can be a fractional brownian motion of Hurst exponent $H = 1/p$. However, the p-variation does not take into account the autocorrelations. Nevertheless, we can consider that the p-variation is a first order approximation (i.e. neglecting autocorrelations) of the true p-moment.

- We have presented a general algorithm for optimal trading curves when the underlying is a selfsimilar process (i.e. a fractal). We showed numerically that the optimal trading curves are increasing in p. Therefore, the effect of the volatility is decreasing in p.

- When $p = 2$ we recover the original Almgren-Chriss model [1].

- We extended our model to portfolios. For general portfolios the algorithm is fully implicit, but *balanced portfolios*) we have the same results as in the single-asset case.

- In other works (e.g. Gatheral and Schied [6]) the underlying follows a geometrical brownian motion. In our model, however, we are not imposing a gaussian model: it is the empirical data who determine the choice of p. Indeed, we set $p = 1/H$ after calibrating H via the real time series of the process. This can offer interesting applications for practitioners.

- The p-variation algorithm (17) has a regularising effect in the sense that the error in the estimate of H is bigger than the error between the corresponding p-curves. Moreover, this effect increases as p increases.

- We plan to extend our model in an upcoming work, where we will include autocorrelations and nonlinear temporary market impact.

References

[1] Robert Almgren, Neil Chriss, *Optimal execution of portfolio transactions*. Journal of Risk, Vol. 3, No. 2, 2000, pp. 5-39.

[2] Robert Almgren, *Optimal execution with nonlinear impact functions and trading-enhanced risk*. Applied Mathematical Finance 10, pp. 1-18.

[3] Robert Almgren, Chee Thum, Emmanuel Hauptman, Hong Li, *Equity market impact*. Risk, July 2005, pp. 57-62.

[4] Jean-Philippe Bouchaud, Marc Potters, *Theory of financial risk and derivative pricing,* 2nd. edition, Cambridge, 2003.

[5] Paul Embrechts, Makoto Maejima, *Self-similar processes*. Princeton Series in Applied Mathematics, USA, 2002.

[6] Jim Gatheral, Alexander Schied, *Optimal trade execution under geometric brownian motion in the Almgren and Chriss framework.* Working paper SSRN, August 2010.

[7] Ismo Koponen, *Analytic approach to the problem of convergence of truncated Lévy flights towards the gaussian stochastic process.*. Phys. Rev. E 52, pp. 1197-1199, 1995.

[8] Charles-Albert Lehalle, *Rigorous strategic trading: balanced portfolio and mean reversion.* Journal of Trading, Summer 2009, pp. 40-46.

[9] Benoît Mandelbrot, *The (mis)behaviour of markets.* Profile Books, 2008.

[10] Rosario Mategna, Eugene Stanley, *Stochastic process with ultraslow convergence to a gaussian: the truncated Lévy flight.* Phys. Rev. Lett. 73, pp. 2946-2949, 1994.

[11] Alexander J. McNeil, Rüdiger Frey, Paul Embrechts, *Quantitative Risk Management: Concepts, Techniques, and Tools.* Princeton University Press, 2005.