King's Research Portal

DOI:
10.3233/AF-200311

Document Version
Peer reviewed version

Link to publication record in King's Research Portal

Citation for published version (APA):

Citing this paper
Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights
Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
You may not further distribute the material or use it for any profit-making activity or commercial gain.
You may freely distribute the URL identifying the publication in the Research Portal.

Take down policy
If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 01. Nov. 2023
Portfolio Optimization for Cointelated Pairs: SDEs vs Machine Learning

Babak Mahdavi-Damghani\textsuperscript{1}, Konul Mustafayeva\textsuperscript{2}, Cristin Buescu\textsuperscript{2}, and Stephen Roberts\textsuperscript{1}

\textsuperscript{1}Oxford-Man Institute of Quantitative Finance, Oxford, UK
\textsuperscript{2}Department of Mathematics, King’s College London, London, UK

Abstract

With the recent rise of Machine Learning (ML) as a candidate to partially replace classic Financial Mathematics (FM) methodologies, we investigate the performances of both in solving the problem of dynamic portfolio optimization in continuous-time, finite-horizon setting for a portfolio of two assets that are intertwined.

In the Financial Mathematics approach we model the asset prices not via the common approaches used in pairs trading such as a high correlation or cointegration, but with the cointelation model in Mahdavi-Damghani (2013) that aims to reconcile both short-term risk and long-term equilibrium. We maximize the overall P&L with Financial Mathematics approach that dynamically switches between a mean-variance optimal strategy and a power utility maximizing strategy. We use a stochastic control formulation of the problem of power utility maximization and solve numerically the resulting HJB equation with the Deep Galerkin method introduced in Sirignano and Spiliopoulos (2018).

We turn to Machine Learning for the same P&L maximization problem and use clustering analysis to devise bands, combined with in-band optimization. Although this approach is model agnostic, results obtained with data simulated from the same cointelation model gives a slight competitive advantage to the ML over the FM methodology.

\textbf{Keywords:} Pairs Trading, Cointelation, Portfolio Optimization, Stochastic Control, Band-wise Gaussian Mixture, Deep Learning.

\textsuperscript{1}Link to the repository containing all relevant codes can be found here: https://github.com/babakmahdavi
1 Introduction

The concept of co-movement and correlation have been shown to be usually misunderstood by practitioners yielding the proposed Cointelation model. We solve the portfolio optimization problem employing a more general set of admissible strategies than long/short strategies used in pairs trading.

A pairs trading strategy involves matching a long position with a short position in two assets with a high correlation. Pairs trading was pioneered in the mid 1980s by a group of quantitative researchers from Morgan Stanley. For an introduction to pairs trading see Vidyamurthy (2004). The securities in a pairs trade must have a high positive correlation, which is the primary driver behind the strategy’s profits.

Pairs trading is based on the high historical correlation of two assets and a trader’s view that the two securities will maintain a specified correlation. A pairs trading strategy is applied when a trader identifies a correlation discrepancy. More specifically, the trader monitors performance of two historically correlated securities. When the correlation between the two securities temporarily weakens, i.e. the spread widens, the trader applies a trading strategy which shorts the high asset and buys the low asset. As the spread narrows again to some equilibrium value, a profit results.

However, many authors argue that correlation is an inappropriate measure of dependency in financial markets, since returns often exhibit a nonlinear co-dependence (e.g. Alexander (2001), Wilmott (2007)). Damghani showed that the measured correlation taken from the returns of a mean-reverting processes is misleading: indeed a strong positive correlation does not necessarily imply that two stochastic processes move in the same direction and vice versa. Also Correlation measures the short term risk but Cointegration, on the other hand, tests the long-term equilibrium relationships between assets and has been extensively used in low frequency pairs trading ( see Vidyamurthy (2004)). Cointegration tests do not measure how well two variables move together, but rather whether the difference between their means remains constant. Sometimes series with high correlation will also be cointegrated, and vice versa, but this is not always the case.

The cointelation model was introduced in (Mahdavi-Damghani (2013)) as a hybrid model which reconciles correlation and cointegration by capturing both short-term risk and long-term equilibrium. The rationale for the long term risk is that during the time of rare market crashes all assets prices fall. However, in the more bullish periods, the short term risk increases, the long term risk becomes less pronounced and the “macro” driver less visible. These influences are accompanied with mean reversion forces from one asset to the other.

In this setting we consider a continuous-time, finite horizon portfolio optimization problems for pairs of assets whose prices follow the cointelation model of Mahdavi-Damghani (2013). Generally, the optimization problem is to find the optimal control

\[
\tilde{\omega}^* = \arg\max_{\tilde{\omega} \in \mathcal{A}} U(X_t^{\tilde{\omega}}, Y_t^{\tilde{\omega}})
\]

where \( U(x) \) is a utility function, \( \tilde{\omega} = (\tilde{\omega}_1, \tilde{\omega}_2) \) is a vector of proportions of wealth invested
in each asset, \( A \) is a set of admissible strategies: either \( \tilde{w}_1 = -\tilde{w}_2 \) (long/short) or \( \tilde{w}_1, \tilde{w}_2 > 0 \) with \( \tilde{w}_1 + \tilde{w}_2 = 1 \) (long only).

We solve the portfolio optimization problem in (1) with Financial Mathematics and Machine Learning methodologies and compare their performance. In the Financial Mathematics approach we use SDE evolution of asset prices, whereas the Machine Learning approach does not assume an underlying model and applies generally to any pair of assets.

In Section 2 we review the cointelation model. In Section 3 we use the classical Financial Mathematics criteria: mean-variance optimization and power utility maximization. In Section 4 we use clustering analysis from Machine Learning to solve the P&L maximization problem. We present the results of each approach in Section 5 and discuss them comparatively.

2 Review of cointelation model for pairs of asset

We first present the usual way correlation is calculated in the financial industry (see e.g. p.274 Wilmott (2007), Alexander (2001)). Assume we have two assets with prices modeled by stochastic processes \((X_t)_{t\geq 0}\) and \((Y_t)_{t\geq 0}\) on a probability space \((\Omega, \mathcal{F}, \mathbb{P})\). We have \(N\) observations of \(X\) and \(Y\) at intervals \(\Delta t\), i.e. \(X(t_i)\) and \(Y(t_i)\) for all \(i = 1, ..., N\) and \(\Delta t = t_i - t_{i-1}\). Here \(\Delta t \in \{1, 5, 22, 252\}\) corresponds to daily, weekly, monthly and yearly data. The \(\Delta t\)-returns on \(i\)-th data point of assets \(X\) and \(Y\) is

\[
R_X(t_i, \Delta t) = \frac{X(t_i + \Delta t) - X(t_i)}{X(t_i)}
\]

\[
R_Y(t_i, \Delta t) = \frac{Y(t_i + \Delta t) - Y(t_i)}{Y(t_i)}.
\]

The sample volatilities of time series of asset prices \(X\) and \(Y\) are then

\[
\sigma_X(\Delta t) = \sqrt{\frac{1}{\Delta t(N-1)} \sum_{i=1}^{N} (R_X(t_i, \Delta t) - \bar{R}_X)^2}
\]

\[
\sigma_Y(\Delta t) = \sqrt{\frac{1}{\Delta t(N-1)} \sum_{i=1}^{N} (R_Y(t_i, \Delta t) - \bar{R}_Y)^2},
\]

where \(\bar{R}_X, \bar{R}_Y\) are the sample average of all the returns in the series of \(X\) and \(Y\), respectively.

The sample covariance between the returns of assets \(X\) and \(Y\) is given by

\[
\sigma_{XY}(\Delta t) = \frac{1}{\Delta t(N-1)} \sum_{i=1}^{N} (R_X(t_i, \Delta t) - \bar{R}_X)(R_Y(t_i, \Delta t) - \bar{R}_Y).
\]

In this paper we consider the measured correlation, which is the sample cross-correlation given by

\[
\rho_{XY}(\Delta t) = \frac{\sigma_{XY}(\Delta t)}{\sigma_X(\Delta t)\sigma_Y(\Delta t)}.
\]
For correlation to be an appropriate choice of measure of co-dependence the assumption of linear dependency between series needs to be satisfied (see Chapter 1.4, Alexander (2001)). Often in financial markets with a non-linear dependence between returns, the correlation is a misleading measure of co-dependency and is misleading, especially when used to capture long-term relationship between assets (see Mahdavi-Damghani et al. (2012) and Mahdavi-Damghani (2013)).

An alternative statistical measure to correlation is cointegration. If two time series $X_t$ and $Y_t$ are integrated of order $d$ and there exists $\beta$ such that a linear combination $X_t + \beta Y_t$ is integrated of order less that $d$, then $X_t$ and $Y_t$ are cointegrated (see Engle and Granger (1991)). Since the spread of cointegrated asset prices is mean reverting, they have a common stochastic trend, i.e. the asset prices are ‘tied together’ in the long term, although they might drift apart in the short-term (see Alexander (1999)). Because the cointegration requires sophisticated statistical analysis, it has not been used as widely as correlation in the financial industry.

Although correlation and cointegration are related, they are different concepts. High correlation does not necessarily imply high cointegration, and neither does high cointegration imply high correlation (e.g. see Figure 4 in Mahdavi-Damghani et al. (2012)). Two assets may be perfectly correlated over short timescales yet diverge in the long run, with one growing and the other decaying. Conversely, two assets may follow each other, with a certain finite spread, but with any correlation, positive, negative or varying.

Mahdavi-Damghani (2013) proposed cointegration as a hybrid model that aims to mediate between correlation and cointegration. It captures both short-term and long-terms relationships between the assets.

**Definition 1.** Consider a filtered probability space by $(\Omega, F, (F_t)_{t \geq 0}, \mathbb{P})$, with the historical probability measure, $\mathbb{P}$. The cointegration model for a pairs of assets with prices $X_t$ and $Y_t$ defined in Mahdavi-Damghani (2013) as

\[
\begin{align*}
dX_t &= \mu X_t dt + \sigma X_t dW_t, \\
dY_t &= \kappa (X_t - Y_t) dt + \eta Y_t d\tilde{W}_t, \\
d\langle W, \tilde{W} \rangle_t &= \rho dt,
\end{align*}
\]

where $\mu \in \mathbb{R}$, $\sigma > 0$, $X(t_0) = x_0$ are the drift, diffusion coefficients and initial value of asset price $X$; $0 < \kappa \leq 1$, $\eta > 0$, $Y(t_0) = y_0 > 0$ are the rate of mean reversion, volatility and initial value of the asset price $Y$; $(\tilde{W}(t))_{t \geq 0}$ and $(W(t))_{t \geq 0}$ are two correlated Brownian motions with constant correlation coefficient $-1 \leq \rho \leq 1$ that generate the filtration $(F_t)_{t \geq 0}$.

The processes $(X)_{t \geq 0}$ and $(Y)_{t \geq 0}$ are called the leading process and the lagging process, respectively. This is due to the fact that the lagging process reverts around the leading process.

We present here the concepts of inferred correlation function and number of crosses

---

2A time series $X_t$ is integrated of order $d$ if $(1 - L)^d X_t$ is a stationary process. Here $L$ is a lag operator.
formula introduced in [Mahdavi-Damghani (2013)] in order to devise a test whether two pairs are cointelated.

Let \( \rho^*_{XY}(\Delta t) \) be the inferred correlation function between two times series of cointelated asset prices defined as follows

\[
\rho^*_{XY}(\Delta t) = \sup_{0<\Delta \tau \leq \Delta t} \rho_{XY}(\Delta \tau).
\] (9)

Sometimes there may not be enough data to calculate \( \Delta t \)-inferred (measured) correlation of cointelated assets. In [Mahdavi-Damghani (2013)] the following formula for approximation of inferred correlation (9) was proposed via examining various data sets:

\[
\rho^*_{XY}(\Delta t) \approx \rho + (1 - \rho) \left[ 1 - \exp \left( -\lambda \kappa (\Delta t - 1) \right) \right],
\] (10)

where \( \kappa \in [0, 1] \), \( \lambda > 0 \), \( \rho \in [-1, 1] \). The parameter \( \lambda \approx 1.75 \) for "regular financial data", although it is itself a function in general. Thus, if one does not have enough empirical data to calculate, for example, the yearly (252 days) inferred correlation, the formula in equation (10) allows to approximate it using only \( \kappa \) and \( \rho \) parameters of cointelation model in (8) and setting \( \Delta t = 252 \), \( \lambda \approx 1.75 \).

The motivation for inferred correlation approximation formula (10), is that in the discrete version of the processes in equation (8) the measured correlation increases as the time increment, \( \Delta t \), increases (e.g. correlations calculated using daily, weekly, monthly returns). Moreover, the measured correlation of cointelated pairs will converge to 1 faster as the speed of mean reversion parameter \( \kappa \) increases. If we set \( \rho = -1 \) in (8), the inferred correlation of cointelated asset prices may cover the whole correlation spectrum \([-1, 1]\) (see Figure 1).

Another way for testing if two times series are cointelated is to study how many times the normalized series cross paths. If one discretizes equation (8), then one can approximate the expectation of the number of times, \( \Gamma_{x,y} \), the two stochastic process, \( x = X_i \in [1, 2, \ldots, N] \) and \( y = Y_i \in [1, 2, \ldots, N] \), cross paths as follows

\[
E[\Gamma_{x,y}(\kappa, N)] \approx N \left[ \gamma (1 - \kappa) + \frac{1}{2} \sqrt{\kappa} \right]
\] (11)

with \( N \) is the length of the data, \( \gamma \) is a positive constant and \( \kappa \) is the speed of mean reversion in equation (8).

Compared to the number of times purely correlated SDEs (eg: without the mean reversion component, i.e. when \( \kappa = 0 \)) the number of times the discrete version of the cointelated SDEs cross paths is larger than if they were random, and the bigger the \( \kappa \) the more often the paths of discretized SDEs cross each other per unit of time.

Then two stochastic processes are cointelated (see [Mahdavi-Damghani (2013)]) if

- Inferred correlation formula in equation (10) is verified;
- The number of crosses formula in equation (11) is verified;
• the underlying assets have a reasonable physical connection that would suggest their spread should mean revert (e.g. oil and BP share prices).

The parameters in cointelation model (8) can be estimated using the inferred correlation formula (10) and the number of crosses formula (11) (see Mahdavi-Damghani (2013)). Similarly to the variance reduction methodology described in Mahdavi-Damghani et al. (2012), Mahdavi-Damghani (2013), we can define

\[
B_+ = \frac{\left| \max(X_t - Y_t, t \in [0, T]) \right|}{2},
\]

\[
B_- = \frac{\left| \min(X_t - Y_t, t \in [0, T]) \right|}{2}.
\]

We note that the estimation of \( \kappa \) has a higher variance when

\[
Z_{\rho} = B_+ > |X_t - Y_t| > B_-,
\]

where \( \rho \), on the other hand has quality samples. The reverse is true when

\[
Z_{\kappa} = |X_t - Y_t| > B_+ \bigcup |X_t - Y_t| < B_-.
\]

We can therefore sample \( \kappa \) in \( Z_{\kappa} \) and \( \rho \) in \( Z_{\rho} \). Figure 2 illustrates this.

3 Financial Mathematics approach for portfolio optimization problem

We consider the portfolio of two assets and model the their prices with the cointelation in (8). We approach the optimization problem of this portfolio with classic Financial Mathematics criteria: mean-variance and power utility maximization. Since the cointeled assets are characterized by both correlation and mean-reversion components, we formulate the mean-variance optimization problem for long only strategies and we calculate the optimal strategies to make profit on correlation. To make profit on mean-reversion property of the cointeled assets we use stochastic control formulation of the power utility maximization problem for long/short strategies and calculate the optimal weights. We then maximize portfolio P&L by dynamically switching between these two optimal strategies.

3.1 Mean-variance optimization

We first review fundamental notions and concepts for mean-variance optimization.

Returns: A portfolio considers a combination of \( n \) potential assets, with an initial capital \( V(0) \) and weights \( w_1, w_2, ..., w_n \), such that \( \sum_i^n w_i = 1 \). \( w_i V(0) \) is the amount invested in security \( i \) for \( i = 1, 2, ..., n \) at time \( t = 0 \). The number of shares to invest in security \( i \) at
time $t = 0$ is
\[ n_i = \frac{w_i V(0)}{S_i(0)}. \] (15)

The value of portfolio at time $t$ is
\[ V(t) = \sum_{i=1}^{N} n_i S_i(t). \] (16)

Given the number of shares $n_i$ with $i = 1, ..., n$, the percentage of the portfolio invested in asset $i$ at time $t$ is
\[ w_i(t) = \frac{n_i S_i(t)}{\sum_{i=1}^{N} n_i S_i(t)}, \] (17)
with $\sum_{i=1}^{N} w_i(t) = 1$. The rate of return of asset $i$ at time $t$ (i.e. over $[t - \Delta t, t]$) is given by
\[ R_i(t) = \frac{S_i(t) - S_i(t - \Delta t)}{S_i(t - \Delta t)} = \frac{S_i(t)}{S_i(t - \Delta t)} - 1. \] (18)

The rate of return of portfolio, $R_p(t)$, is then
\[ R_p(t) = \frac{V(t) - V(t - \Delta t)}{V(t - \Delta t)}. \] (19)

We can show that the return of portfolio is a linear combination of the returns of individual assets as follows
\[
R_p(t) = -1 + \frac{V(t)}{V(t - \Delta t)} = -1 + \sum_{i=1}^{N} \frac{n_i S_i(t)}{\sum_{i=1}^{N} n_i S_i(t - \Delta t)} \\
= -1 + \sum_{i=1}^{N} \frac{n_i S_i(t - \Delta t) S_i(t)}{\sum_{i=1}^{N} n_i S_i(t - \Delta t) S_i(t - \Delta t)} = -1 + \sum_{i=1}^{N} w_i(t) (R_i(t) + 1) \\
= \sum_{i=1}^{N} w_i(t) R_i(t). \] (20)

Sometimes it is more convenient to use log returns, which are defined for asset $i$ by
\[ r_i(t) = \ln \left( \frac{S_i(t)}{S_i(t - \Delta t)} \right). \] (21)

It should be pointed out that for short period of time the log return is approximately equal to the rate of return
\[ r_i(t) = \ln \left( \frac{S_i(t)}{S_i(t - \Delta t)} \right) = \ln(R_i(t) + 1) \approx R_i(t). \] (22)

Therefore we do not distinguish between these two returns, as long as the time increment, $\Delta t$, is short compared to the rate of return. Going forward we will use daily logarithmic
returns. Thus, the return of portfolio, $r_p$, at time at time $t$ in this case becomes

$$r_p = \sum_{i=1}^{N} w_i r_i. \quad (23)$$

**Expectation and variance of returns:** By the linearity property of expected value operator, the expected return of portfolio, $E(r_p)$, is

$$E(r_p) = E \left( \sum_{i=1}^{N} w_i r_i \right) = \sum_{i=1}^{N} w_i E(r_i) = \sum_{i=1}^{N} w_i \mu_i = w^\top \mu, \quad (24)$$

where $\mu_i$ denotes the expected return of asset $i$ and $w^\top = [w_1, w_2, ..., w_n]$, $\mu = [\mu_1, \mu_2, ..., \mu_n]^\top$.

The variance of the return of portfolio, $\text{Var}(r_p)$, is given by

$$\text{Var}(r_p) = E \left[ \left( \sum_{i=1}^{N} w_i r_i - E(r_p) \right)^2 \right] = E \left[ \left( \sum_{i=1}^{N} w_i (r_i - E(r_i)) \right)^2 \right]$$

$$= \sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j E \left[ (r_i - E(r_i))(r_j - E(r_j)) \right]$$

$$= \sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j \sigma(r_i, r_j) = w^\top \Sigma w, \quad (25)$$

where $\Sigma$ denotes the covariance matrix of the asset returns, composed of all covariances between the returns of assets $i$ and $j$ defined as $\sigma(r_i, r_j)$. The variance of asset $i$’s return, which constitute the diagonal of the covariance matrix, is $\sigma(r_i, r_i)$.

**Optimal investment strategy using mean-variance criterion**

We consider a portfolio consisting of two assets. The uncertainty is modelled by a probability space $(\Omega, \mathcal{F}, P)$ with a filtration $(\mathcal{F}_t)_{t \geq 0}$ generated by two-dimensional Brownian motion: $(W, \tilde{W})$. Denote by $X(t)$ and $Y(t)$ the prices of two assets at time $t$, with dynamics following cointelation model in [8]. The investment behavior is modelled by an investment strategy $h = (h_1, h_2)$. Here, $h_i \in [0, 1]$, $i = 1, 2$, denotes the percentage of total wealth invested in $i$-th asset (see equation [17]). Let $h_1(t)$ and $h_2(t)$ denote respectively the portfolio weights for assets $X$ and $Y$ at time $t$. The holdings are allowed to be adjusted continuously up to a fixed horizon $T$.

Denoting by $V^h_t$ the value of portfolio at time $t$ associated to a strategy $h$ we have

$$V^h(t) = \frac{h_1(t)V^h(t)}{X(t)} X(t) + \frac{h_2(t)V^h(t)}{Y(t)} Y(t), \quad (26)$$
with initial wealth $V^h(t_0) = v_0$. We restrict our considerations to self-financing strategies, where the value of the portfolio changes only because the asset prices change, i.e. there is no inflow or withdrawal of money (see [Harrison and Kreps (1979)]). In this case the dynamic of the wealth process is

$$dV^h(t) = V^h(t) \left[ h_1(t) \frac{dX(t)}{X(t)} + h_2(t) \frac{dY(t)}{Y(t)} \right].$$  \hspace{1cm} (27)$$

Let $\mathcal{A}^1$ denote the set of all admissible strategies, $h = (h_1, h_2)$, satisfying:

(i) Given $v_0 > 0$ the wealth process $V^{v_0,h}(\cdot)$ corresponding to $w_0, h$ satisfies

$$V^{v_0,h}(t) \geq 0, \quad 0 \leq t \leq T,$$

(ii) $h_i(t) \geq 0$ for all $i = 1, 2$,

(iii) $\sum_{i=1}^2 h_i(t) = 1$.

An investment strategy, $h \in \mathcal{A}^1$, is called optimal if there exists no other strategy $\tilde{h} \in \mathcal{A}^1$ such that $E(r_p(h)) \geq E(r_p(\tilde{h}))$ and $Var(r(h)) \leq Var(r(\tilde{h}))$ with at least one inequality being strict (see [Li and Ng (2000)]).

We define a utility function, $U(t,h)$, as in [Bodie et al. (1999)]:

$$U(t,h) = 2 \tau E[r_p(t)] - \sigma^2[r_p(t)],$$

where $\tau \geq 0$ is the risk tolerance coefficient. Then according to [Garcia et al. (2017)] we have the following proposition.

**Proposition 1** (Mean-Variance Criterion). Finding an optimal strategy for mean-variance criteria is equivalent to the utility maximization problem:

$$\max_{h(t)} U(t,h)$$

with constraints

- $\sum_{i=1}^N h_i = 1,$
- $h_i \geq 0 \ \forall i.$

and $U(t,h)$ given in (29).

Thus we have optimization problem in equation (30). From equation (20) we have that the rate of return of our portfolio, $R_p$, over $[t - \Delta t, t]$ is

$$R_p(t) = \frac{V^h(t) - V^h(t - \Delta t)}{V^h(t - \Delta t)} = \sum_{i=1}^2 h_i(t) R_i(t),$$

$$\hspace{1cm} (31)$$
where $R_i$ is the rate of return of individual assets. The log return of our portfolio, $r_p$ is given by

$$r_p(t) = h_1 r_1(t) + h_2 r_2(t),$$

where $r_i(t) \approx R_i(t)$, as we showed in equation (22).

**Lemma 1.** Denote by $V^h(t)$ the value of the portfolio corresponding to the admissible strategy $h \in \mathcal{A}$. Then:

(i) The expectation of portfolio return over $[t - \Delta t, t]$ is

$$E(r_p(t)) = h_1 E[r_X(t)] + h_2 E[r_Y(t)],$$

(ii) The variance of portfolio return over $[t - \Delta t, t]$ is

$$\text{Var}(r_p(t)) = h_1^2 \text{Var}[r_X(t)] + h_2^2 \text{Var}[r_Y(t)] + 2h_1h_2 \text{Cov}[r_Xr_Y(t)],$$

where $r(X_t) = \ln \left( \frac{X_t}{X_{t-\Delta t}} \right)$ and $r(Y_t) = \ln \left( \frac{Y_t}{Y_{t-\Delta t}} \right)$ the daily log returns of assets $X$ and $Y$ and

- $E(r_X(t)) = (\mu - \frac{\sigma^2}{2})\Delta t$ is the expected return of the asset price $X$ over the horizon $[t - \Delta t, t]$;
- $E(r_Y(t)) = \left[ \ln \left( a e^{\kappa \Delta t} + (Y_0 - a) e^{-\kappa \Delta t} \right) - \frac{a (2(\mu + \kappa) \Delta t) + de(\mu + \sigma \rho) \Delta t}{2(ae^{\kappa \Delta t} + (Y_0 - a) e^{-\kappa \Delta t})^2} - \ln(Y_{t-\Delta t}) \right] - \frac{(Y_0^2 - c - d) e^{2(\kappa - \kappa) \Delta t}}{2(ae^{\kappa \Delta t} + (Y_0 - a) e^{-\kappa \Delta t})^2} + \frac{1}{2}$ is the expected return of the asset price $Y$ over the horizon $[t - \Delta t, t]$;
- $\text{Var}(r_X(t)) = \sigma^2 \Delta t$ is the variance of return of asset price $X$ over the horizon $[t - \Delta t, t]$;
- $\text{Var}(r_Y(t)) = \frac{a e^{2\kappa \Delta t} (Y_0 - a)^2}{(ae^{\kappa \Delta t} + (Y_0 - a) e^{-\kappa \Delta t})^2} + \frac{d e^{2(\kappa - \kappa) \Delta t}}{2(ae^{\kappa \Delta t} + (Y_0 - a) e^{-\kappa \Delta t})^2} + \frac{(Y_0^2 - c - d) e^{2(\kappa - \kappa) \Delta t}}{(ae^{\kappa \Delta t} + (Y_0 - a) e^{-\kappa \Delta t})^2} - 1$ is the variance of return of asset price $Y$ over the horizon $[t - \Delta t, t]$;
- $\text{Cov}(r_X(t), r_Y(t)) = \ln \left( \frac{b e^{(\mu + \kappa) \Delta t} + (X_0 - b)e^{(\kappa - \kappa) \Delta t}}{a \kappa (Y_0 - a) e^{\kappa \Delta t} + (X_0 - a X_0) e^{(\kappa - \kappa) \Delta t}} \right)$ is the covariance of returns of two asset prices $X$ and $Y$ over the horizon $[t - \Delta t, t]$.

**Proof.** See Appendix [A].

The optimal weights for mean-variance criterion were derived in Soeryana et al. (2017). We state the following proposition from Soeryana et al. (2017) applied to the cointelation model (8).

**Proposition 2.** The optimal solution for the problem in (30) for cointelation model (8) is:

$$h^*(t) = \frac{1}{e^\epsilon} \Sigma^{-1}(t)e \tau \left[ \Sigma^{-1}(t)M(t) - \frac{e^\epsilon \Sigma^{-1}(t)M(t)}{e^\epsilon \Sigma^{-1}(t)e} \right],$$

where $\Sigma^{-1}(t) = \sigma^2 \Delta t$, $\epsilon = \frac{\sigma^2}{2}$, $\tau = \frac{1}{\sigma^2}$, $\Sigma^{-1}(t)$ is the inverse of the covariance matrix, and $\Sigma(t) = \text{Cov}(X(t), Y(t))$. The formula above represents the optimal weights for the mean-variance criterion over the horizon $[t - \Delta t, t]$. 

10
Replacing these formulas for expectation, variance and covariance of the returns of asset prices in equation (35), we get optimal strategies for mean-variance optimization problem. We will present numerical examples in Section 5.

3.2 Stochastic control for pairs trading

Power utility maximization problem

We now use a stochastic control approach to the power utility maximization problem. Here we mainly follow Mudchanatongsuk et al. (2008), but with modified dynamics for asset prices. More specifically, they assume the price dynamics of one of the assets is a geometric Brownian motion and model the log-spread as an Ornstein-Uhlenbeck process. We, however, assume the dynamics of asset prices are governed by the cointegration model in equation (8), where one of the assets follow the geometric Brownian motion and the second asset mean reverts around the first one.

Let \((\Omega, \mathcal{F}, P)\) be a complete probability space with a filtration \((\mathcal{F}_t)_{t \geq 0}\) generated by two-dimensional Brownian motion: \((W, \tilde{W})\). We consider the same market as in Subsection 3.1: two assets which follow the cointegration model (8).

We assume an initial wealth \(v_0 > 0\) at time \(t = 0\). Initial wealth is held in a margin account. For simplicity we assume that the interest rate for margin account is 0, \(r = 0\). Margining constraints can be quite punitive financially. The holdings are allowed to be adjusted continuously up to a fixed horizon \(T\). The investment behavior is modelled by an investment strategy \(\pi = (\pi_1, \pi_2)\). Here, \(\pi_i(t), i = 1, 2\), denotes the percentage of total wealth invested in \(i\)-th asset at time \(t\) (see equation (17)). Let \(\pi_1(t), \pi_2(t)\) be respectively the portfolio weights for assets \(X\) and \(Y\) at time \(t\). We only allow pairs trading: short one of the asset and long the other in equal dollar amount, i.e. \(\pi_1(t) = -\pi_2(t)\). In addition, we restrict our considerations to self-financing strategies.

We define admissible control and controlled process as in Korn and Kraft (2002).

Definition 2 (Control). Given a subset \(U\) of \(\mathbb{R}^2\), we denote by \(\mathcal{U}_0\) the set of all progressively measurable processes \(\pi = \{\pi_t, t \geq 0\}\) valued in \(U\). The elements of \(\mathcal{U}_0\) are called control processes.

Denote by \(V^\pi(t)\) the value of portfolio corresponding to strategy \(\pi\) at time \(t\), which is given by

\[
V^\pi(t) = \frac{\pi_1(t)V^\pi(t)}{X(t)}X(t) + \frac{\pi_2(t)V^\pi(t)}{Y(t)}Y(t). \tag{36}
\]

The dynamics of the portfolio value \(V^\pi\) associated with strategy \(\pi = (\pi_1, \pi_2)\) is given by

\[
dV^\pi(t) = V^\pi(t) \left[ \pi_1(t) \frac{dX(t)}{X(t)} + \pi_2(t) \frac{dY(t)}{Y(t)} \right] \tag{37}
\]

Replacing the dynamics for \(X(t)\) and \(Y(t)\) into (37) we get:

\[
dV^\pi(t) = V^\pi(t) \left[ \pi_1(\mu dt + \sigma dW(t)) - \pi_1 \left( \kappa \left( \frac{X(t)}{Y(t)} - 1 \right) dt + \eta d\tilde{W}(t) \right) \right]. \tag{38}
\]
**Lemma 2.** Denote $Z(t) := \frac{X(t)}{Y(t)}$. For the cointelation model (8) we obtain that $Z(t)$ has the dynamics

$$dZ(t) = [\mu + \eta^2 - \sigma \rho - \kappa(Z(t) - 1)]Z(t)dt + Z(t)(\sigma dW(t) + \eta d\tilde{W}(t)).$$  

(39)

**Proof.** By Ito’s quotient rule:

$$d \left( \frac{X(t)}{Y(t)} \right) = \frac{dX(t)}{X(t)} Y(t) - \frac{dY(t)}{Y(t)} X(t) + \frac{d\langle X, Y \rangle_t}{Y(t)} X(t) - \frac{d\langle X, Y \rangle_t}{X(t)} Y(t).$$

(40)

Writing this in terms of $Z(t)$ gives

$$dZ(t) = Z(t)(\mu dt + \sigma dW_t - \kappa(Z(t) - 1)dt - \eta d\tilde{W}(t)) + \frac{\eta^2Y^2(t)}{Y^2(t)} dt - \sigma \rho dt)$$

$$= [\mu + \eta^2 - \sigma \rho - \kappa(Z(t) - 1)]Z(t)dt + (\sigma dW(t) - \eta d\tilde{W}(t))Z(t),$$

(41)

which proves the lemma.

For each control process $\pi \in \mathcal{U}_0$ we rewrite the dynamics of two-dimensional state process, $P = (V, Z)$, as follows

$$dP(t) = a(t, P(t), \pi(t))dt + b(t, P(t), \pi(t))dB(t).$$

(42)

with initial value of $P(t_0) = p_0$ and $B = (W, \tilde{W})$ being the two-dimensional Brownian motion. The process $P$ is called the controlled process. Let $[t_0, T]$ with $0 \leq t_0 < T < \infty$ be the relevant time interval and define $Q := [t_0, T) \times \mathbb{R}^2$. The coefficient functions

$$a : Q \times U \to \mathbb{R}^2,$$

$$b : Q \times U \to \mathbb{R}^{2 \times 2},$$

are all continuous. Further, for all $\pi \in U$ let $a(\cdot, \cdot, \pi)$ and $b(\cdot, \cdot, \pi)$ be in $C^1(Q)$. We then define

**Definition 3 (Admissible control).** Denoting $\mathcal{A}^2$ the set of all admissible controls, we say a control $\{\pi(t)\}_{t \in [t_0, T]}$ will be called admissible if the following conditions hold

(i) $\forall k \in \mathbb{N}$ the integrability condition

$$E \left( \int_{t_0}^T |\pi(s)|^k ds \right) < \infty$$

(44)

is satisfied,

(ii) the corresponding state process $P^\pi$ satisfies

$$E^{t_0, p_0} \left( \sup_{t \in [t_0, T]} |P^\pi(t)|^k \right) < \infty,$$
(iii) only pairs trading is allowed: short one of the asset and long the other

$$\pi_1 = -\pi_2. \quad (45)$$

Since we consider self-financing portfolio, then by equation \((45)\) the dynamics of the state process, \(P = (V^\pi, Z)\), becomes

\[
dV^\pi(t) = V^\pi(t) \left[ (\pi_1[\mu - \kappa(Z(t) - 1)]) dt + \pi_1[\sigma dW(t)] + \eta d\tilde{W}(t) \right], \quad V^\pi(0) = v_0, \\
dZ(t) = [\mu + \eta^2 - \sigma \eta \rho - \kappa(Z(t) - 1)] Z(t) dt + [\sigma dW(t) - \eta d\tilde{W}(t)] Z(t), \quad Z(0) = z_0.
\]

**Optimal investment strategy**

We assume that an investor’s preference is represented by the power utility function

$$U(x) = \frac{1}{\gamma} x^\gamma, \quad (46)$$

with \(x \geq 0\) and risk aversion parameter \(\gamma < 1\). Our aim is to maximize the objective functional \(J\) over all admissible controls, i.e. determine an admissible control \(\pi(\cdot)\) such that for each initial value \((t_0, v_0)\) the utility functional below is maximized:

$$J(t_0, v_0, z_0; \pi) := \mathbb{E} [U(V^\pi(T)) | V_{t_0} = v_0, Z_{t_0} = z_0]. \quad (47)$$

The optimization problem is to find \(\tilde{v}(t, v, z)\) and \(\pi \in \mathcal{A}^2\) such that

$$\tilde{v}(t, v, z) := \sup_{\pi(\cdot) \in \mathcal{A}^2} J(t, v, z, \pi) = J(t, v, z, \pi^*). \quad (48)$$

Consider the function \(G(t, v, z)\) such that \(G \in C^{1,2}(Q)\). The Hamilton-Jacobi-Bellman (HJB) equation corresponding to the stochastic control problem \((48)\) is

$$\frac{\partial G}{\partial t}(t, v, z) + \sup_{\pi(\cdot) \in \mathcal{A}^2} \mathcal{L}^\pi G(t, v, z) = 0, \quad (49)$$

subject to terminal condition

$$G(T, v, z) = v^\gamma. \quad (50)$$

The infinitesimal generator, \(\mathcal{L}^\pi G(t, v, z)\) in \((49)\) associated with the two dimensional state process \(P = (V, Z)\) is given by

$$\mathcal{L}^\pi G(t, v, z) = \frac{1}{2} [\pi_1^2 (\sigma^2 - 2 \sigma \eta \rho + \eta^2) v^2 G_{vv} + 2 \pi_1 (\sigma^2 - 2 \sigma \eta \rho + \eta^2) v_z G_{vz} + (\sigma^2 - 2 \sigma \eta \rho + \eta^2) z^2 G_{zz}] + [\pi_1 (\mu - \kappa(z - 1))] v G_v + [\mu + \eta^2 - \sigma \eta \rho - \kappa(z - 1)] z G_z. \quad (51)$$
Theorem 1. If there exists an optimal control $\pi^*(\cdot)$ then $G$ coincides with the value function:

$$G(t, v, s) = \tilde{v}(t, v, z) = J(t, v; \pi^*).$$

Using separation ansatz we reduce a 3-dimensional HJB equation in (49) to the following 2-dimensional PDE:

$$\tilde{\sigma}(\gamma - 1)ff_t - \frac{1}{2} \tilde{\sigma}^2 \gamma^2 f_z^2 - \frac{1}{2} \gamma [\mu - \kappa(z - 1)]^2 f + \frac{1}{2} \tilde{\sigma}(\gamma - 1)z^2 ff_z z - \tilde{\sigma} \gamma [\mu - \kappa(z - 1)]zf z + \tilde{\sigma}(\gamma - 1)[\mu + \eta^2 - \sigma \eta \rho - \kappa(z - 1)]zf z,$$

with $f(T, z) = 1$, $(t, z) \in [0, T] \times \mathbb{R}$, $\forall z \in \mathbb{R}$, (52)

where $\tilde{\sigma} = \sigma^2 - 2\sigma \eta \rho + \eta^2$.

The issue at this stage is that this PDE does not have a closed for solution. This is a non standard PDE, which is not high dimensional but is nonlinear which makes using finite difference methods or any standard numerical methods inadequate. For this reason we propose to use the "Deep Galerkin Method" to solve the PDE in (52). Once the solution is found, we can write the optimal strategy as

$$\pi^*_1 = \frac{\tilde{\sigma} z G_{vz} + [\mu - \kappa(z - 1)]G_v}{\tilde{\sigma} v G_{vv}} = -\frac{\tilde{\sigma} z (f_z v^{\gamma - 1}) + [\mu - \kappa(z - 1)](fv^{\gamma - 1})}{\tilde{\sigma} v (fv^{2\gamma - 2\gamma} (\gamma - 1))}$$

$$= -\frac{\tilde{\sigma} z f_z + [\mu - \kappa(z - 1)]f}{\tilde{\sigma} f(\gamma - 1)} = -\frac{zf_z}{(\gamma - 1)f} - \frac{[\mu - \kappa(z - 1)]}{\tilde{\sigma}(\gamma - 1)}. \quad (53)$$

See Appendix B for the details.

3.3 Deep learning for solving PDE in stochastic control

Without an analytical solution to the non-standard 2-dimensional PDE in (52), we approximate the solution with an algorithm "Deep Galerkin Method" (DGM) proposed in [Sirignano and Spiliopoulos (2018)]. DGM is a merger of the Galerkin method and deep neural network machine learning algorithm. The Galerkin method is a popular numerical method which seeks a reduced-form solution to a PDE as a linear combination of basis functions. The deep learning algorithm, or DGM, uses a deep neural network instead of a linear combination of basis functions. The algorithm is trained on batches of randomly sampled time and space points, therefore it is mesh free.

**Brief review of DGM**

In general case, consider a PDE with $d$ spatial dimensions:

$$\frac{\partial u}{\partial t}(t, x; \theta) + Lu(t, x) = 0, \quad (t, x) \in [0, T] \times \Omega,$n$$

$$u(t, x) = g(t, x), \quad x \in \partial \Omega,$n$$

$$u(t = 0, x) = u_0(x), \quad x \in \Omega \quad (54)$$
where \( x \in \Omega \subset \mathbb{R}^d \) and \( \mathcal{L} \) is an operator of all the other partial derivatives. The goal is to approximate the \( U(t,x) \) with deep neural network \( f(t,x;\theta) \). Here \( \theta \in \mathbb{R}^K \) are the neural network parameters. We want to minimize the objective function associated to the problem which consists of three parts:

1. A measure of how well the approximation satisfies the PDE:
   \[
   \left\| \frac{\partial f}{\partial t}(t,x;\theta) - \mathcal{L}f(t,x;\theta) \right\|_{[0,T] \times \Omega,\nu_1}^2. \tag{55}
   \]

2. A measure of how well the approximation satisfies the boundary condition:
   \[
   \left\| \frac{\partial f}{\partial t}(t,x;\theta) - g(t,x) \right\|_{[0,T] \times \partial\Omega,\nu_2}^2. \tag{56}
   \]

3. A measure of how well the approximation satisfies the initial condition:
   \[
   \left\| \frac{\partial f}{\partial t}(0,x;\theta) - u(0,x) \right\|_{\Omega,\nu_3}^2. \tag{57}
   \]

Here all three errors are measured in terms of \( L^2 \)-norm, i.e. \( \|f(y)\|_{Y,\nu}^2 = \int_{Y} |f(y)|^2 \nu(y) dy \) with \( \nu(y) \) being a density on region \( Y \).

The sum of all three terms above gives us the objective function associated with the training of the neural network:

\[
J(f) = \left\| \frac{\partial f}{\partial t}(t,x;\theta) - \mathcal{L}f(t,x;\theta) \right\|_{[0,T] \times \Omega,\nu_1}^2 + \left\| \frac{\partial f}{\partial t}(t,x;\theta) - g(t,x) \right\|_{[0,T] \times \partial\Omega,\nu_2}^2 + \left\| \frac{\partial f}{\partial t}(0,x;\theta) - u(0,x) \right\|_{\Omega,\nu_3}^2. \tag{58}
\]

Thus, the goal is to find a set of parameters \( \theta \) such that the function \( f(t,x;\theta) \) minimizes the error \( J(f) \). When the dimension \( d \) is large, estimating \( \theta \) by directly minimizing \( J(f) \) is infeasible. Therefore, one can minimize the error \( J(f) \) using a machine learning approach: stochastic gradient descent, where we use a sequence of time and space points drawn randomly. The algorithm for DGM method is described in Algorithm 1 below.

**Remark 1.** The learning rate, \( \alpha_n \), is a configurable hyperparameter\(^3\) used in the training of neural networks that controls how much to change the model in response to the estimated error. Each time the model weights are updated. Learning rate has a small positive value, often in the range between 0.0 and 1.0. Similar to Al-Aradi et al. (2018), we set \( \alpha_0 = 0.001 \). Note that our learning rate \( \alpha_n \) must decrease with \( n \) (see Sirignano and Spiliopoulos (2018)) and a simple enough way to do that is by using an exponential weighted method where \( \alpha_n \leftarrow \alpha_{n-1} \cdot \lambda \) with \( \lambda \in ]0,1[ \).

\(^3\)In machine learning, a hyperparameter is a parameter whose value is set before the learning process begins whereas, the values of other parameters are derived via training.
Algorithm 1 Deep Galerkin Method

Require: $L(f), u(), g()$
Ensure: $L_1^L + L_2^L + L_3^L$ is minimized

Generate random points:
1: $(t_n, x_n) \sim U \sim [0, 1]^2$
2: $(\tau_n, z_n) \sim U \sim [0, 1]^2$
3: $w_n \sim U \sim [0, 1]$
4: $s_n \sim (t_n, x_n; \tau_n, z_n; w_n)$

Calculate the squared error:
5: $L_1^n \leftarrow \left( \frac{\partial f}{\partial t}(t_n, x_n; \theta_n) - L_f(t_n, x_n; \theta_n) \right)^2$
6: $L_2^n \leftarrow \left( \frac{\partial f}{\partial \tau}(\tau_n, z_n; \theta_n) - g(\tau_n, z_n) \right)^2$
7: $L_3^n \leftarrow \left( \frac{\partial f}{\partial x}(0, x_n; \theta_n) - u(0, w_n) \right)^2$
8: $G(\theta_n, s_n) \leftarrow L_1^n + L_2^n + L_3^n$

Take a descent step at the random points:
9: $-\arg\max_{\theta_n} G(\theta_n, s_n)$
10: $\alpha_n \leftarrow \alpha_{n-1} \star \lambda$
11: $\theta_{n+1} \leftarrow \theta_n - \alpha_n \nabla_\theta G(\theta_n, s_n)$
Repeat until tolerance level $10^{-8}$ for convergence criterion is achieved

The neural network (NN) architecture used in DGM is like a long short-term memory (LSTM) network though with small differences (see Sirignano and Spiliopoulos (2018)). We describe below the architecture of this NN:

$$S^l = \sigma(w^1 \cdot x + b^1)$$
$$Z^l = \sigma(u^{z,l} \cdot x + w^{z,l} \cdot S^l + b^{z,l}) \quad l = 1, \ldots, L$$
$$G^l = \sigma(u^{g,l} \cdot x + w^{g,l} \cdot S^l + b^{g,l}) \quad l = 1, \ldots, L$$
$$R^l = \sigma(u^{r,l} \cdot x + w^{r,l} \cdot S^l + b^{r,l}) \quad l = 1, \ldots, L$$
$$H^l = \sigma(u^{h,l} \cdot x + w^{h,l} \cdot (S^l \odot R^l) + b^{h,l}) \quad l = 1, \ldots, L$$
$$S^{l+1} = (1 - G^l) \odot H^l + Z^l \odot S^l \quad l = 1, \ldots, L$$
$$f(t, x, \theta) = w \cdot S^{L+1} + b$$

with $\odot$ denoting Hadamard multiplication, L number of layers and $\sigma$ the activation function.

The rest of the superscripts refer to the neurones for our NN architecture of Figures 3 and 4.

Remark 2. We can see the Bird Eye view of the DGM Al-Aradi et al. (2018), Sirignano and Spiliopoulos (2018) method in Figure 3 and its details in Figure 4. The rationale is explained in Al-Aradi et al. (2018), Sirignano and Spiliopoulos (2018).
Testing DGM on Merton problem

Because the DGM method was relatively new, we wanted to test the algorithm ourselves. More specifically the method was tested with several nonlinear, high-dimensional PDEs independently in Al-Aradi et al. (2018) and Sirignano and Spiliopoulos (2018), including the nonlinear HJB equations. We have tested the DGM algorithm on the HJB equation for the Merton problem ourselves. More specifically, Figures 5 and 6 show the plots of the analytical and approximated surface with DGM solution. We found the performance of the algorithm quite impressive. Indeed, Figure 7 shows the difference between the analytical and the approximate solution. Most of the time, the error is between 0% and 1%. The only criticism, though negligible that we can make is that the approximate solution does not do as well around $t = 0$ (the maximum error of 4% is around $t = 0$). This corroborates with the findings in Al-Aradi et al. (2018). The authors do not give an explanation of why this is the case but we did not think that this small imperfection was big enough for us to abandon the methodology.

Solution to our PDE problem using DGM

Recall the PDE we want to solve is given in equation (52). In the absence of a closed form solution to this PDE we approximate the solution with the DGM algorithm described above. Figure 8 shows the approximate solution to the PDE in (52) for different parameter values. Recall, once we have the numerical solution $f$ for the PDE above, we obtain the optimal weights as following:

\[
\pi^*_1 = -\frac{\tilde{\sigma} z G_{vz} + [\mu - \kappa(z - 1)] G_v}{\tilde{\sigma} v G_{vv}} = -\frac{\tilde{\sigma} z (f z v^{\gamma - 1} \gamma) - [\mu - \kappa(z - 1)](f v^{\gamma - 1} \gamma)}{\tilde{\sigma} v (f v^{\gamma - 2} \gamma (\gamma - 1))} = -\frac{z f z}{(\gamma - 1)f} - \frac{[\mu - \kappa(z - 1)]}{\tilde{\sigma}(\gamma - 1)},
\]

with $\pi^*_1 = -\pi^*_2$.

3.4 Dynamic Switching: optimal strategies of mean-variance and power utility

Although in the previous two cases we assume that an investor has a certain risk preferences as modelled by a utility function (MVC and power utility), it is interesting to consider a limiting case where the investor can be always persuaded to go for more money (identical utility function $U(x) = x$, which is essentially the power utility function with risk aversion parameter $\gamma = 1$) when deciding between MVC or power utility. Assuming that an investors’ preference is modelled either as in equation (29) or as in equation (46), in order to improve further the portfolio returns we employ dynamic switching between the two optimal strategies

\[
\psi^*(t) = \begin{cases} 
\pi^*(t), & \text{if } V^{\pi^*}(t) \geq V^{h^*}(t), \\
h^*(t), & \text{otherwise},
\end{cases}
\]
where $\pi^*(t)$ and $h^*(t)$ are given in equations (60) and (35) and $V^{\pi^*}(t)$ and $V^{h^*}(t)$ are given in equations (36) and (26). The motivation behind the dynamic switching is that the investor wants to benefit from both the mean-reversion and the correlation elements of the cointegration model (8). More specifically, as the spread between two assets increases the investor implements pairs trading and makes profit, otherwise the MVC approach is used. The portfolio return over investment horizon $[0,T]$ with $T = 1000$ days is

$$R(r_p) = \frac{V(0) - V(T)}{V(0)}.$$  

(62)

We perform 500 simulations with the same model and present in Table 1 the average results. The average return at terminal time $T$ obtained by using dynamic switching optimal strategies is higher than the average returns calculated by employing MVC or power utility maximizing optimal strategies.

4 Machine Learning formulation of the optimization problem

4.1 The portfolio optimization problem

We assume an initial wealth $\tilde{w}_0 > 0$ at time $t = 0$. The investment behavior is modeled by an investment strategy $w = (w_1, w_2)$. Here $w_1(t), w_2(t)$ denote the percentages of wealth invested in asset $X$ and $Y$ respectively at time $t$. Let $V(t)$ denote the portfolio value at time $t$ and $V^{PnL}(t) := V(t) - V(0)$ denote the profit and loss (P&L) over $[0,t]$. At each time $t$ we allow either pairs trading: $w_1(t) = -w_2(t)$ or long only strategies without leverage: $w_1(t) + w_2(t) = 1$ with $w_1(t), w_2(t) > 0$.

The general optimization problem is to find an optimal strategy, $w(t)$, such that the terminal P&L is maximized:

$$w^*(t) = \operatorname*{argmax}_{w(t) \in A} V^{PnL}(w,T),$$  

(63)

where $V^{PnL}(w,T)$ is profit and loss corresponding to the strategy $w$ at time terminal time $T$. We use clustering analysis to devise the bands and in each band we solve the following optimization problem

$$w_i^*(t) = \operatorname*{argmax}_{w_i(t) \in A} V^{PnL}(w_i,t),$$  

(64)

where $i = 1, \ldots, n$ is the number of bands, $V^{PnL}(w_i,t)$ is profit and loss corresponding to the strategy $w_i$ at time $t$. Then the overall solution $w^*$ is obtained via a linear interpolation of optimal weights per band $w_i^*$. The advantage of the proposed method is that we do not impose a certain model on the asset prices, inline with Data Curation methodologies discussed by Bailey Jr (2018). Only data observations are required to calculate the optimal weight, meaning that the complex SDE calibration is avoided.
4.2 Review of Band-Wise Gaussian Mixture model

We review band-wise Gaussian mixture model because it inspires our method of selecting the bands. Consider a probability space \((\Omega, \mathcal{F}, P)\) and let \((P_t)_{t \geq 0}\) denote the asset price. Mahdavi-Damghani and Roberts (2017) has recently introduced a generalised bumping SDE for the price dynamics of asset \(P_t\). The SDE contains some secondary parameters whose purpose is empirical manual fitting. The generalized SDE is given by

\[
dP_t = \theta_{t,\tau}(\mu_{t,\tau} - P_t)dt + \sigma P_t^\alpha (1 - P_t^2)^\beta dW_t,
\]

Here \(\theta_t\) is the speed of mean reversion, \(\mu_t\) is the long term mean, \(\alpha\) is the positivity flag enforcer, \(\beta\) is the \([-1, +1]\) boundary flag enforcer and \(\{\bigcup_{i=t-\tau} W_i\}\) is the set of historical deviations of the assumed model’s distribution (e.g.: all the historical absolute returns in the context of a normal diffusion).

This generalised SDE gives as a special case the cointelation model: take \(\theta = -\mu, \mu = 0, \alpha = 1\) and \(\beta = 0\) for the dynamic of \(X\) in (8); take \(\theta_{t,\tau} = \kappa, \mu_{t,\tau} = X_t, \alpha = 1, \beta = 0\) for the dynamics of \(Y\) in (8). The SDE in (65) can also model:

- Proportional returns (log-normal diffusion) when \(\theta = 0, \alpha = 1, \beta = 0\).
- Absolute returns (normal diffusion) when \(\theta = 0, \alpha = 0, \beta = 0\),
- Mean reverting returns where we enforce positivity of returns (e.g. CIR Cox et al. (1985) diffusion when \(\alpha = 1/2\) and \(\beta = 0\)),
- Mean reverting returns where we do not enforce positivity of the returns (e.g OU Uhlenbeck and Ornstein (1930) diffusion when \(\alpha = 0\) and \(\beta = 0\)).

In general calibrating parameters of the SDE in (65) to a real data is complex. Using data simulated with (65) their empirical distribution is approximated for the purpose of prediction by a band-wise Gaussian mixture model. This is done for a sequence of bands which are created using Machine Learning clustering method (see Mahdavi-Damghani and Roberts (2017)).

Let \(P = \{p_1, \ldots, p_n\}\) be a set of empirical random variables sampled using equation (65) with cumulative distribution function \(F(p)\) and density \(f(p)\). Denote \(O = \{p_1, \ldots, p_n\}\) the ordered set of \(P\) such that \(p_1 < p_2 < \ldots < p_n\) and

\[
O_h^j = \{p_{[(n(i-1)+1)/h]}, \ldots, p_{[(n(i)/h)]}\}.
\]

Then the band-wise Gaussian mixture model for the empirical distribution function of the data simulated using the SDE in equation (65) is given as follows:

\[
\hat{F}_n(p_i|\mathcal{F}_t) = \frac{1}{n} \sum_{j=1}^{h} \sum_{i=\eta}^{\zeta} 1_{p_i \in O_h^j}
\]

(66)
with \( \eta = \lceil n((i - 1) + 1)/h \rceil \) and \( \zeta = \lfloor n(i)/h \rfloor \). For example in the case bands \( h = 3 \), using a Gaussian Mixture such that

\[
\hat{F}_n(p_t|F_t) = \mathcal{N}(-3, 1)1_{p_t \in O^+_{31}} + \mathcal{N}(0, 1)1_{p_t \in O^+_{32}} + \mathcal{N}(3, 1)1_{p_t \in O^+_{33}},
\]

we obtain the approximate stratification in Figure 9. The stratification is made so that the cardinality in each \( O^j \) region remains approximately the same, as opposed to being the result of a geometrical separation function of \( p_{(1)} \) and \( p_{(n)} \). Theorem 1 in [Mahdavi-Damghani and Roberts (2017)] ensures a good approximation of the generalised SDE (65) by the Gaussian mixture model (66). The calibration for the band-wise Gaussian mixture is given in Algorithm 2. For our optimization problem we take a similar approach of dividing the range of observations into bands via the clustering algorithm, and then perform an optimization in each band via perturbation of weights.

Algorithm 2 BAND-WISE GAUSSIAN MIXTURE \((P, h)\)

**Require:** array \( P_{1:n} \) and number of bands \( h \)

**Ensure:** \( \Omega^{(1:h)}, [B^+_{(1:h)}, B^-_{(1:h)}] \) are returned.

**Sorting state:**
1. \( X_{(1:h)} \leftarrow \text{QuickSort}(X_{1:n}) \)
2. \( [B^+_{(1:h)}, B^-_{(1:h)}] \leftarrow \text{FindPercentileBands}(X_{(1:n)}, h) \)
3. \( \Omega^{(1: \lfloor n/h \rfloor)} \leftarrow \emptyset \)

**Allocation state:**
4. for \( j = 1 \) to \( h \) do
5.  for \( i = 1 \) to \( n \) do
6.     if \( B^-_{(1:h)} \leq P_{(i)} < B^+_{(1:h)} \) then
7.       Amend(\( \Omega^{(j)}, P_{(i)} \))
8.     end if
9.  end for
10. end for

**Checking Approximation state:**
11. \( \hat{\mu}_{1:h} \leftarrow \text{mean}(\Omega^{(1:h)}) \)
12. \( \hat{\sigma}_{1:h} \leftarrow \text{stdev}(\Omega^{(1:h)}) \)
13. Print(\( \bigcup_{i=1}^{h} \mathcal{N}(\hat{\mu}_i, \hat{\sigma}_i) \))

**Return state:**
14. \( \Omega^{(1:h)}, [B^+_{(1:h)}, B^-_{(1:h)}] \)

4.3 Optimal Machine Learning strategy

Based on the idea of band-wise Gaussian mixture model, we use clustering analysis to create bands, however not for the observed asset price data, but for the spread between two asset prices in \( \mathcal{S} \), i.e. \( X_t - Y_t \). Inside of each band instead of specifying the distribution as in band-wise Gaussian mixture, we test a set of strategies that maximizes the corresponding
P&L. We record the optimal strategies within each band, and in live trading, whenever the spread of asset prices falls in a certain band we employ the optimal strategy for this specific band. We now present the trading signal that translates to investment strategy in machine learning approach.

The Bayesian set-up: We set from equation (8) \( B_t = X_t - Y_t \) and have

\[
B_t = \{ B_{n,t}^+, B_{n-1,t}^+, \ldots, B_{1,t}^+, B_{1,t}^-, \ldots, B_{n-1,t}^-, B_{n,t}^- \},
\]

such that \( B_{n,t}^+ > B_{n-1,t}^+ > \ldots > B_{1,t}^+ > 0 > B_{1,t}^- > \ldots > B_{n-1,t}^- > B_{n,t}^- \). We know that depending on the spread, the resulting approximated distribution of the samples differ (see Mahdavi-Damghani and Roberts (2017)). The calibration algorithm will then consist of creating different zones and test the performance of the possible strategies (“Long/Long”, “Long/Short”, “Short/Long”, “Short/Short”) within these bands. We take a direct approach (see Remark 3) consisting of 3 strategies and their cumulative P&Ls. Fixing the bands \([a_i, b_i]\), with \( i = 1, 2, ..., n \) we consider the following strategies:

- Strategy \( S^{++} \) in which we are long both \( X \) and \( Y \) at time \( t \) in between bands \([a_i, b_i]\) and with P&L \( V_{[a_i,b_i],t}^{++} \).
- Strategy \( S^{+-} \) in which we are long \( X \) and short \( Y \) at time \( t \) in between bands \([a_i, b_i]\) and with P&L \( V_{[a_i,b_i],t}^{+-} \).
- Strategy \( S^{-+} \) in which we are short \( X \) and long \( Y \) at time \( t \) in between bands \([a_i, b_i]\) and with P&L \( V_{[a_i,b_i],t}^{-+} \).

The P&Ls corresponding to these strategies are defined as following:

\[
V_{[a_i,b_i],T}^{++} = \sum_{t=0}^{T} \left[ w_{[a_i,b_i],t}^{++} \Delta X_t + (1 - w_{[a_i,b_i],t}^{++}) \Delta Y_t \right] 1_{a_i < \Delta t \leq b_i},
\]

\[
V_{[a_i,b_i],T}^{+-} = \sum_{t=0}^{T} \left[ w_{[a_i,b_i],t}^{+-} \Delta X_t - (1 - w_{[a_i,b_i],t}^{+-}) \Delta Y_t \right] 1_{a_i < \Delta t \leq b_i},
\]

\[
V_{[a_i,b_i],T}^{-+} = \sum_{t=0}^{T} \left[ -w_{[a_i,b_i],t}^{-+} \Delta X_t + (1 - w_{[a_i,b_i],t}^{-+}) \Delta Y_t \right] 1_{a_i < \Delta t \leq b_i}.
\]

Remark 3. We call this approach direct, since ideally the number of strategies should consist of a more granular weight distribution. However for the sake of comparing with Financial Mathematics approach we consider the same set of strategies: long only, long/short.

We denote the maximum P&L achieved by each of these strategies by \( V_{[a_i,b_i],T}^{+-**,} \), as given by equation (68) and define \( S_{[a_i,b_i],T}^{**} \) of P&L \( V_{[a_i,b_i],T}^{**} \) (equation (69)), the optimal strategy

\[\text{E.g: sharpe ratio maximization}\]
using Gaussian Learning in band \([a_i, b_i]\).

\[
V_{[a_i, b_i], T}^\pm, \ast = \arg\max_{w_{[a_i, b_i], t} \in [0, 1]} \ V_{[a_i, b_i], T}^\pm, w_{[a_i, b_i], t} \in [0, 1]
\]  \hspace{1cm} (68)

\[
V_{[a_i, b_i], T}^{\pm \ast} = \max(V_{[a_i, b_i], T}^{\pm +, \ast}, V_{[a_i, b_i], T}^{\pm -, \ast}, V_{[a_i, b_i], T}^{\pm, \ast}).
\]  \hspace{1cm} (69)

In live trading we recombine the optimal weights per bands into an overall optimal solution via a linear interpolation:

\[
w^*(t) = \sum_{i=1}^{n} w_i^* \mathbf{1}_{(X_t - Y_t) \in [a_i, b_i]}.
\]  \hspace{1cm} (70)

Although we do not have a proof that the resulting interpolated strategy in (70) is optimal, we use it as a benchmark that still improves over the results with Financial Mathematics approach. Our goal is to apply Machine Learning approach to a pair of assets that exhibit some dependence, but this approach can be used for any model, i.e. it is model agnostic.

We further provide Algorithm 3 as the pseudo-code for the calibration process. Note that in both Algorithms 2 and 3 we have used a QuickSort which can be substituted by other sorting algorithms. Note that the use of self explanatory functions such as \texttt{returnCorrespondingStrat}(x,y) in line 20 of Algorithm 3 which given the set of strategies and the P&L returns, as its name indicates, outputs the corresponding strategy that maximizes P&L. The function \texttt{forecast}(x,y,z) in line 22 of Algorithm 3 takes as input the set of trained strategies and the current level of \(X_t\) and \(Y_t\) and returns a prediction of where the signals for the latter two should be. Finally the use of the \texttt{argmax} function in lines 13-16 can be replaced by a simple for loop but in the interest of not making the pseudocode too crowded we have kept it this way.

\textbf{Remark 4.} Mahdavi-Damghani and Roberts (2017) show that a reasonable risk manager or trader can assume the generalized SDE (65) with \(\beta = 0\) and an \(\alpha = 1\), in order to enforce positivity for the simulated scenarios of our risk factor. This very reasonable assumption would have crashed the whole risk engine if it is no longer satisfied in the real markets. The approach we advocate would have, however, been able to continue its dynamical learning scenario without any problem since it is model agnostic.

\section{5 Numerical results}

\subsection{5.1 Simulation}

Figure 10 illustrates the ML and the DS approaches on one single simulated path. Note that when implementing the ML approach with a horizon of 1000 days, we double this data for training, i.e. we use 2000 historical daily prices. We have performed two sets of 500 simulations and we have gathered their results in the following two examples.

\textbf{Example 1.} We have simulated 500 paths of \(X\) and \(Y\) based on cointelation model (8) with parameters \(\mu = 0.05, \sigma = 0.17, \eta = 0.16, \kappa = 0.1, \rho = -0.6\). Figure 11 illustrates that
Algorithm 3 Band-Wise ML for Cointelation \((P, h)\)

Require: array \(P_{1:n}\) and number of bands \(h\)
Ensure: \(\Omega^{(1:h)}, [B^+_{(1:h)}, B^-_{(1:h)}]\) are returned

Sorting state:
1. \(P_{(1:h)} \leftarrow \text{QuickSort}(P_{1:h})\)
2. \([B^+_{(1:h/2)}, B^-_{(1:h/2)}]\) \(\leftarrow \text{FindPercentileBands}(P_{1:n}, h)\)
3. \(B_{(1:h)} \leftarrow [B^+_{(1:h/2)}, B^-_{(1:h/2)}]\)
4. \(\Omega^{(1:\lceil n/h \rceil)} \leftarrow []\)

Allocation state:
5. \(\text{for } j = 1 \text{ to } h \text{ do}\)
6. \(\text{for } i = 1 \text{ to } n \text{ do}\)
7. \(\text{if } P_{(i)} \in B^i \text{ then}\)
8. \(\text{Amend}(\Omega^{(j)}, P_{(i)})\)
9. \(\text{end if}\)
10. \(\text{end for}\)
11. \(\text{end for}\)

Optimize the 3 types of P&L for each band:
12. \(\text{for } i = 1 \text{ to } h \text{ do}\)
13. \(V^{++,*}_{B_i,T} \leftarrow \arg \max_{w_{B_i,t} \in [0,T]} V^{++}_{B_i,T}\)
14. \(V^{+-,*}_{B_i,T} \leftarrow \arg \max_{w_{B_i,t} \in [0,T]} V^{+-}_{B_i,T}\)
15. \(V^{-+,*}_{B_i,T} \leftarrow \arg \max_{w_{B_i,t} \in [0,T]} V^{-+}_{B_i,T}\)
16. \(\text{end for}\)

Rank and return best strategy for each band:
17. \(\text{for } i = 1 \text{ to } h \text{ do}\)
18. \(V^{**,*}_{B_i,T} \leftarrow \max(V^{++,*}_{B_i,T}, V^{+-,*}_{B_i,T}, V^{-+,*}_{B_i,T})\)
19. \(S^{*}_{T} \leftarrow (S^{++,*}_{B_i,T}, S^{+-,*}_{B_i,T}, S^{-+,*}_{B_i,T})\)
20. \(S^{**}_{B_i,T} \leftarrow \text{returnCorrespondingStrat}(V^{**,*}_{B_i,T}, S^{*}_{T})\)
21. \(\text{end for}\)

Forecasting:
22. \(\text{signal}^S, \text{signal}^S_l \leftarrow \text{forecast}(S^{**}_{B_i,T}, S_t, S_{l,t})\)

Return buy/sell signals:
23. \(\text{signal}^S, \text{signal}^S_l\)

the Machine Learning approach with long/short strategies \((ML_{LS})\), on average performs slightly better in terms of P&L than the Stochastic Control approach \((SC)\). However, based on histogram none of the approaches perform significantly better or significantly worse than the other at any time.

Example 2. We have simulated 500 paths of \(X\) and \(Y\) based on cointelation model \((8)\) with
parameters $\mu = 0.05, \sigma = 0.17, \eta = 0.16, \kappa = 0.1, \rho = -0.6$. Figure 12 illustrates how the ML approach seems to perform slightly better in terms of P&L than the FM approach about 55\% of the time, while being outperformed the other 45\% of the time. However, based on histogram we have noticed that sometimes the ML approach is being outperformed significantly more than it outperforms FM approach.

From our performance histogram in Figures 11 and 12 we have concluded that for parameters $\mu = 0.05, \sigma = 0.17, \eta = 0.16, \kappa = 0.1, \rho = -0.6$ of cointelation model (8) we have the following rankings for the approaches:

$$SC < ML_{LS} < FM < ML.$$  

The reason for ML with full set of strategies (long only and long/short) outperforming the DS most of the time might be the fact that in long only optimal strategies of ML approach we have more variety in weights, whereas the closed form formula (35) in FM gives us almost constant weights (with small fluctuations).

5.2 Market Data: Bitcoin vs NVDA

5.2.1 Rational

During the peer reviewing process we were as to use the model with real market data (on top of the simulations presented previously). The recommendation was to use two-asset model with uncorrelated memoryless jumps in both the leading asset and the spread to produce more realistic model. A pure cryptocurrency model was not recommended so we thought a hybrid model would be best. We decided to propose NVDA and Bitcoin to be cointelated pairs for the following reason. When cryptocurrencies prices go up, then the incentive to mine cryptocurrencies increases which increases, in turn, the value of companies involved in mining. The most well known cryptocurrency is Bitcoin, and the most well known mining equity remains NVDA. We have decided to study this pair, assuming they are cointelated. The data set was downloaded using a free data provider (yahoo finance) for the last 5 years. In figure 15 we have displayed the price of Bitcoin, NVDA as well as their normalised spread in our timescale. The volatility of bitcoin being significantly higher than the one of NVDA, we decided to proxy their current volatility as their rolling 6 months standard deviation. We then constructed their spread using an inverse weighting function of their current volatility. Please refer to our github code in order to gain full transparency for the normalisation process. We roughly separated the “in sample” and “out of sample” data in half (2 years and 3 month each: the first 6 months being “burnt” in order to calculate current volatility). As per the described methodology the bands were chosen to have the same size (0 to 33 percentile, 33 to 67 percentile and finally 67 to 100 percentile). We also incorporated a simple cost function (adding conservatively a couple of bps per position change: 4 in total assuming a long position would become short and vice versa). We recorded our in and out of
sample Sharpe ratios in Table 2. We can observe few points. First, over-fitting is noticeable since the SR of our overall strategy drops by 1.2. However, this is mitigated by the fact that the out-of-sample SR remain at 1.7, a very reasonable figure at the low frequencies. Added to this point, all bands are doing reasonably well in both the in and the out of sample. We take this as being evidence of a reasonable over-fitting. Figure 16 plots the spread of the cointegrated pair on top as well as its 67 and 33 percentile bands.

6 Conclusion

6.1 Possible directions for future work

One direction for future work is to consider portfolio optimization problem for n-dimensional cointegration model. For instance, when \( n = 3 \) we can have something of the following form:

\[
\begin{align*}
    dS^a_t &= \sigma S^a_t dW^a_t \\
    dS^b_t &= \theta (S^a_t - S^b_t) dt + \sigma S^a_t dW^b_t \\
    dS^c_t &= \theta (S^a_t - S^c_t) dt + \sigma S^c_t dW^c_t
\end{align*}
\]  

One natural question would first be about how to model this triplet? For instance would equation (72) with \( S^b \) and \( S^c \) reversion around \( S^a \) be more in-line with the pair from equation (8) or would \( S^b \) reversion around \( S^a \) and \( S^c \) reversion around \( S^b \) be better? Are they equivalent or is one more useful? What happens as \( n \) increases? We plan to examine these questions in the future.

6.2 Conclusion

We have studied the portfolio optimization problem of two assets that follow the cointegration model using two approaches: Financial Mathematics and Machine Learning. We first implemented the FM approach, where we use classic financial mathematics criteria: mean-variance and power utility maximization. Without an analytical solution to the PDE (52), we resort to the DGM method, a deep learning algorithm, to solve it numerically. The second approach we implemented is ML using clustering. The latter approach is easier to implement, it is model agnostic, therefore avoids the complex SDE calibration. In our case the Machine Learning approach slightly outperforms the Financial Mathematics approach.

\footnote{Note that this is almost always true in an low signal to noise ratio in the lower frequencies.}
Appendices

A Proof of Lemma [1]

Since $X_t$ is a geometric Brownian motion, we have

$$E[r(X_t)] = (\mu - \frac{\sigma^2}{2}) \Delta t \tag{72}$$

where $X_{t-\Delta t}$ is a known constant at time $t - \Delta t$. The expectation of log return of asset $Y$ is

$$E[r(Y_t)] = E[\ln(Y_t)] - \ln(Y_{t-\Delta t}), \tag{73}$$

where $Y_{t-\Delta t}$ is a known constant at time $t - \Delta t$. We use Taylor expansion to approximate expected value and variance of $\ln(Y_t)$ and covariance of $\ln(Y_t)$ and $\ln(X_t)$ (see Benaroya et al. (2005), p.165-167):

$$E[\ln(Y_t)] \approx \ln (E[Y_t]) - \frac{\sigma^2}{2} E[Y_t^2], \tag{74}$$

$$\sigma^2[\ln(Y_t)] \approx \frac{\sigma^2}{E[Y_t^2]}, \tag{75}$$

$$\sigma[\ln(Y_t) \ln(X_t)] \approx \ln \left(1 + \frac{\sigma Y_t X_t}{E[X_t] E[Y_t]}\right). \tag{76}$$

First, we need to derive $E[Y_t]$. From equation (8) we have

$$Y_t = Y_{t-\Delta t} + \kappa \int_{t-\Delta t}^t (X_s - Y_s)ds + \eta \int_{t-\Delta t}^t Y_s dZ_s. \tag{77}$$

Taking expectation on both sides we have

$$E[Y_t] = Y_{t-\Delta t} + \kappa \int_{t-\Delta t}^t E[X_s - Y_s] ds. \tag{78}$$

Differentiating on both sides we get

$$\frac{dE[Y_t]}{dt} = \kappa E[X_t] - \kappa E[Y_t] = \kappa X_{t-\Delta t} e^{\mu \Delta t} - \kappa E[Y_t]. \tag{79}$$

Denoting $E[Y_t] = y(t)$ we obtain an ordinary differential equation (ODE):

$$y' = -\kappa y + \kappa X_{t-\Delta t} e^{\mu \Delta t}. \tag{80}$$

The solution is given by

$$y(t) = E[Y_t] = ae^{\mu \Delta t} + (Y_{t-\Delta t} - a)e^{-\kappa \Delta t}, \tag{81}$$

26
where
\[ a = \frac{\kappa X_{t-\Delta t}}{\mu + \kappa}. \]

In order to derive \( E[Y^2_t] \) we first compute \( E[X_t Y_t] \). Applying integration by parts (IBP) to (8) we get
\[
d(X_t Y_t) = X_t dY_t + Y_t dX_t + dX_t dY_t
= \kappa X_t^2 dt - \kappa X_t Y_t dt + \eta X_t Y_t dW_t + \mu X_t Y_t dt + \sigma X_t Y_t dZ_t + \sigma \eta \rho X_t Y_t dt.
\]
(82)

Thus
\[
X_t Y_t = X_{t-\Delta t} Y_{t-\Delta t} + \kappa \int_{t-\Delta t}^{t} X_s^2 ds + \eta \int_{t-\Delta t}^{t} X_s Y_s dW_s +
(\mu - \kappa + \sigma \eta \rho) \int_{t-\Delta t}^{t} X_s Y_s ds + \sigma \int_{t-\Delta t}^{t} X_s Y_s dZ_s.
\]

Taking expectation and differentiating on both sides
\[
\frac{dE[X_t Y_t]}{dt} = \kappa E[X_t^2] + (\mu - \kappa + \sigma \eta \rho) E[X_t Y_t].
\]
(83)

Denoting \( E[X_t Y_t] = x(t) \) we obtain ODE
\[
x'(t) = \kappa E[X_t^2] + (\mu + \sigma \eta \rho - \kappa) y.
\]
(84)

Since \( X_t \) is GBM, its second moment is given by
\[
E[X_t^2] = E[X_{t-\Delta t}^2 e^{(2\mu - \sigma^2)\Delta t + 2\sigma W_t}] = X_{t-\Delta t}^2 e^{(2\mu + \sigma^2)\Delta t}.
\]
(85)

Thus (84) becomes
\[
x'(t) = \kappa X_{t-\Delta t} e^{(2\mu + \sigma^2)\Delta t} + (\mu - \kappa + \sigma \eta \rho) y.
\]
(86)

Using variation of parameters method we get the solution
\[
x(t) = E[X_t Y_t] = b e^{(2\mu + \sigma^2)\Delta t} + (X_{t-\Delta t} Y_{t-\Delta t} - b) e^{(\mu - \kappa + \sigma \eta \rho)\Delta t},
\]
(87)

where
\[
b = \frac{\kappa X_{t-\Delta t}}{\mu + \sigma^2 + \kappa - \sigma \eta \rho}.
\]

Now we are ready to compute \( E[Y^2_t] \). By Itô’s lemma the dynamics of \( Y^2_t \) is
\[
dY^2_t = 2Y_t dY_t + (dY_t)^2 = (\eta^2 - 2\kappa)Y^2_t dt + 2\kappa X_t Y_t dt + 2\eta Y^2_t dZ_t.
\]
(88)

Integrating on both sides
\[
Y^2_t = Y^2_0 + 2(\eta^2 - \kappa) \int_0^t Y^2_s ds + 2\kappa \int_0^t X_s Y_s ds + 2\eta \int_0^t Y^2_s dZ_s.
\]
(89)
Taking expectation on both sides and and differentiating

\[
\frac{dE[Y_t^2]}{dt} = 2(\eta^2 - \kappa)E[Y_t^2] + 2\kappa E[X_tY_t].
\]

Defining \( E[Y_t^2] = z(t) \) and replacing the value for \( E[X_tY_t] \) form equation (87) we obtain an ODE

\[
z' = (\eta^2 - \kappa)z + 2\kappa be^{(2\mu+\sigma^2)\Delta t} + 2\kappa(X_t-\Delta t) e^{(\mu-\kappa+\sigma \eta)p}\Delta t.
\]

Using again variation of parameters we obtain the following solution

\[
z(t) = E[Y_t^2] = ce^{(2\mu+\sigma^2)\Delta t} + de^{(\mu-\kappa+\sigma \eta)p}\Delta t + (Y_t^2 - c - d) e^{2(\eta^2 - \kappa)\Delta t},
\]

with \( c = \frac{2\kappa b}{2\mu+\sigma^2-\eta^2+2\kappa} \) and \( d = \frac{2\kappa(X_t-\Delta t) - b}{\mu-2\mu^2+\kappa+\sigma \eta p} \).

Now we are ready to approximate \( E[\ln(Y_t)] \). From (74) we have

\[
E[\ln(Y_t)] \approx \ln[E[Y_t]] - \frac{E[Y_t^2]}{2E[Y_t]^2} + \frac{1}{2} = \ln \left( ae^{\mu \Delta t} + (Y_t-\Delta t - a) e^{-\kappa \Delta t} \right) + \frac{1}{2} \frac{ce^{(2\mu+\sigma^2)\Delta t} + de^{(\mu-\kappa+\sigma \eta)p}\Delta t}{2(ae^{\mu \Delta t} + (Y_t-\Delta t - a) e^{-\kappa \Delta t})^2} - \frac{(Y_t^2 - c - d) e^{2(\eta^2 - \kappa)\Delta t}}{2(ae^{\mu \Delta t} + (Y_t-\Delta t - a) e^{-\kappa \Delta t})^2} - \ln(Y_t-\Delta t)
\]

and

\[
E[r(Y_t)] = E[\ln(Y_t)] - \ln(Y_t-\Delta t) \approx \ln \left( ae^{\mu \Delta t} + (Y_t-\Delta t - a) e^{-\kappa \Delta t} \right) + \frac{1}{2} - \frac{ce^{(2\mu+\sigma^2)\Delta t} + de^{(\mu-\kappa+\sigma \eta)p}\Delta t}{2(ae^{\mu \Delta t} + (Y_t-\Delta t - a) e^{-\kappa \Delta t})^2} - \frac{(Y_t^2 - c - d) e^{2(\eta^2 - \kappa)\Delta t}}{2(ae^{\mu \Delta t} + (Y_t-\Delta t - a) e^{-\kappa \Delta t})^2} - \ln(Y_t-\Delta t)
\]

From (75) we have

\[
Var[r(Y_t)] = Var[\ln(Y_t)] \approx E[Y_t^2] - 1 = \frac{ce^{(2\mu+\sigma^2)\Delta t} + de^{(\mu-\kappa+\sigma \eta)p}\Delta t}{(ae^{\mu \Delta t} + (Y_t-\Delta t - a) e^{-\kappa \Delta t})^2} + \frac{(Y_t^2 - c - d) e^{2(\eta^2 - \kappa)\Delta t}}{(ae^{\mu \Delta t} + (Y_t-\Delta t - a) e^{-\kappa \Delta t})^2} - 1
\]

and

\[
Var[r(X_t)] = Var[\ln(X_t)] = \sigma^2 \Delta t.
\]

From (76) we obtain the covariance:

\[
Cov[r(X_t)r(Y_t)] = Cov[\ln(X_t)\ln(Y_t)] \approx \ln \left( \frac{E[X_t Y_t]}{E[X_t] E[Y_t]} \right) = \ln \left( \frac{be^{(2\mu+\sigma^2)\Delta t}(X_t-\Delta t) e^{(\mu-\kappa+\sigma \eta)p}\Delta t}{a \sigma \kappa e^{(2\mu+\sigma^2)\Delta t}(Y_t-\Delta t - a) e^{(\mu-\kappa+\sigma \eta)p}\Delta t} \right).
\]
B Dimension reduction of 3-dim HJB \((49)\)

For ease of notation let \( \hat{\sigma} = \sigma^2 - 2\sigma\eta + \eta^2 \) and rewrite \((49)\):

\[
G_t + \sup_{\pi_t} \left\{ \frac{1}{2} \left( \pi_t^2 \sigma v^2 G_{vv} + \hat{\sigma} z^2 G_{zz} + 2 \pi_t \hat{\sigma} v z G_{vz} \right) \right. \\
\left. + (\pi_t [\mu - \kappa(z - 1)]) v G_v + (\mu + \eta^2 - \sigma \eta \rho - \kappa(z - 1)) z G_z \right\} = 0. \tag{97}
\]

The first order condition for the maximization is

\[
\pi_t^* \hat{\sigma} v G_{vv} + \hat{\sigma} z G_{vz} + [\mu - \kappa(z - 1)] G_v = 0. \tag{98}
\]

Now assuming \( G_{vv} < 0 \) the first order condition is sufficient, yielding

\[
\pi_t^* = -\frac{\hat{\sigma} z G_{vz} + [\mu - \kappa(z - 1)] G_v}{\hat{\sigma} v G_{vv}}. \tag{99}
\]

Replacing \((99)\) back into \((97)\) yields:

\[
G_t + \frac{1}{2} \left\{ (\hat{\sigma} z G_{vz} + [\mu - \kappa(z - 1)] G_v)^2 \right. \\
\left. - \hat{\sigma} z G_{vz} + [\mu - \kappa(z - 1)] G_v \sigma v z G_{vz} \right\} + [\mu + \eta^2 - \sigma \eta \rho - \kappa(z - 1)] z G_z \\
- \frac{\hat{\sigma} z G_{vz} + [\mu - \kappa(z - 1)] G_v}{\sigma v G_{vv}} [\mu - \kappa(z - 1)] v G_v = 0.
\]

Multiplying both sides of equation by \( \hat{\sigma} G_{vv} \) we get:

\[
\hat{\sigma} G_t G_{vv} + \frac{1}{2} (\hat{\sigma} z G_{vz} + [\mu - \kappa(z - 1)] G_v)^2 - (\hat{\sigma} z G_{vz} + [\mu - \kappa(z - 1)] G_v) \hat{\sigma} z G_{vz} \\
+ \frac{1}{2} \hat{\sigma} z^2 G_{zz} G_{vv} - (\hat{\sigma} z G_{vz} + [\mu - \kappa(z - 1)] G_v) \mu - \kappa(z - 1)] G_v \\
+ \hat{\sigma} [\mu + \eta^2 - \sigma \eta \rho - \kappa(z - 1)] z G_z G_{vz} = 0.
\]

Expanding gives

\[
\hat{\sigma} G_t G_{vv} + \frac{1}{2} \hat{\sigma} z^2 G_{vz}^2 + \frac{1}{2} [\mu - \kappa(z - 1)]^2 G_v^2 + \hat{\sigma} z [\mu - \kappa(z - 1)] G_v G_{vz} \\
- \hat{\sigma}^2 z^2 G_{vz}^2 - \hat{\sigma} z [\mu - \kappa(z - 1)] G_v G_{vz} + \frac{1}{2} \hat{\sigma} z^2 G_{zz} G_{vv} - \hat{\sigma} z [\mu - \kappa(z - 1)] G_v G_{vz} \\
- [\mu - \kappa(z - 1)]^2 G_v^2 + \hat{\sigma} [\mu + \eta^2 - \sigma \eta \rho - \kappa(z - 1)] z G_z G_{vz} = 0. \tag{100}
\]

Which further simplifies to

\[
\hat{\sigma} G_t G_{vv} - \frac{1}{2} [\mu - \kappa(z - 1)]^2 G_v^2 + \frac{1}{2} \hat{\sigma} z^2 G_{zz} G_{vv} + \hat{\sigma} z [\mu - \kappa(z - 1)] G_v G_{vz} \\
- \frac{1}{2} \hat{\sigma} z^2 G_{vz}^2 + \hat{\sigma} [\mu + \eta^2 - \sigma \eta \rho - \kappa(z - 1)] z G_z G_{vz} = 0. \tag{101}
\]
At this stage we were able to turn our four variable PDE into three, but we can get eliminate one more. For this we consider the following separation ansatz:

\[ G(t, v, z) = f(t, z)v^{\gamma}, \]  

(102)

with the terminal condition

\[ f(T, z) = 1 \quad \forall z. \]  

(103)

We compute the derivatives of (102):

\[
\begin{align*}
G_t &= f_t v^{\gamma}, \\
G_v &= f v^{\gamma} - 1, \\
G_z &= f_z v^{\gamma}, \\
G_{vv} &= f v^{\gamma} - 2 \gamma, \\
G_{vz} &= f_z v^{\gamma} - 1, \\
G_{zz} &= v^{\gamma} f_{zz}, \\
\end{align*}
\]

Replace derivative back into (101) and divide by \( v^{2(\gamma-1)} \) to get

\[
\begin{align*}
\tilde{\sigma} (\gamma - 1) f f_t - \frac{1}{2} \tilde{\sigma}^2 \gamma z^2 f_{zz} - \frac{1}{2} \gamma [\mu - \kappa(z - 1)] f^2 f + \frac{1}{2} \tilde{\sigma} (\gamma - 1) z^2 f f_{zz} - \\
\tilde{\sigma} \gamma [\mu - \kappa(z - 1)] z f f_z \tilde{\sigma} (\gamma - 1) [\mu + \eta^2 - \sigma \rho - \kappa(z - 1)] f f_z = 0. \tag{104}
\end{align*}
\]

We now have a PDE with only two variables instead of four.

References


Wilmott, P. (2007), Paul Wilmott Introduces Quantitative Finance, John Wiley Sons Ltd., Chichester, West Sussex.
Criterion | Average portfolio return $\hat{R}(r_p)$
--- | ---
MVC | 35% 
SC | 61% 
DS | 83%

Table 1: Average over 500 simulations of portfolio returns at terminal time $T$ (day 1000) with dynamic switching (DS) is higher than average portfolio return with only stochastic control (SC) or only mean-variance-criterion (MVC).
<table>
<thead>
<tr>
<th>Band</th>
<th>In Sample SR</th>
<th>Out of Sample SR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>2.9</td>
<td>1.7</td>
</tr>
<tr>
<td>Band Low</td>
<td>1.3</td>
<td>.5</td>
</tr>
<tr>
<td>Band Middle</td>
<td>1.4</td>
<td>1.4</td>
</tr>
<tr>
<td>Band high</td>
<td>2.2</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Table 2: In and out of sample sharpe ratios by band.
Figure 1: (Up) Simulated path of cointelation model \( \text{[8]} \) with \( \rho = -1, \theta = 0.1, \sigma = 0.01 \); (Down) Corresponding measured correlation \( \text{[7]} \) as a function of the time increment increases from \(-1\) to \(1\).

Figure 2: Hypothetical spread split in three different zones for risk management or/and trading purposes.

Figure 3: Bird’s-eye perspective of overall DGM architecture \([\text{Al-Aradi et al. (2018)}] \).

Figure 4: Operations within a single DGM layer \([\text{Al-Aradi et al. (2018)}] \).

Figure 5: Analytical solution of the Merton Problem.

Figure 6: Approximate solution of Merton problem using DGM.

Figure 7: Error between analytical and approximate solution of Merton problem.

Figure 8: Approximate solutions to PDE \((\text{104})\) with DGM for four different scenarios of \( \rho \) and \( \mu \) and fixed \( \sigma = 0.2, \eta = 0.19, \gamma = 0.5 \).

(a) Approximate solution with low \( \mu = 0.01 \) and low \( \rho = -0.5 \).
(b) Approximate solution with low \( \mu = 0.01 \) and high \( \rho = 0.5 \).
(c) Approximate solution with high \( \mu = 0.4 \) and low \( \rho = -0.5 \).
(d) Approximate solution with high \( \mu = 0.4 \) and high \( \rho = 0.5 \).

Figure 9: Two examples of Gaussian Mixture Simulations with different number of bands.

(a) Empirical distribution of random variable sampled from cointelation model \( \text{[8]} \) in three different zones described in Figure 2.
(b) Empirical distribution of random variable sampled from cointelation model \( \text{[8]} \) in five different zones: two additional zones were added to the initial three zones in Figure 2.

Figure 10: (a) one simulated scenario based on cointelation model \( \text{[8]} \) with parameters: \( \mu = 0.05, \sigma = 0.17, \eta = 0.16, \kappa = 0.1, \rho = -0.6 \) and scaled spread: \( \kappa(X_t - Y_t) \); (b) portfolio return and optimal weight of asset \( X \) with Dynamic Switching approach; (c) portfolio return and optimal weight of asset \( X \) and \( Y \) with Machine Learning approach.

Figure 11: Histogram of excess \((P\&L)\) for \(ML_{LS}\) vs SC at terminal time \(T\).

Figure 12: Histogram of excess \((P\&L)\) for ML vs FM at terminal time \(T\).

Figure 13: In FM approach optimal long/short strategies are more volatile than optimal long only strategies.

Figure 14: In ML approach optimal long/short strategies are slightly more volatile than optimal long only strategies.
Figure 15: Cumulative returns in the last 5 years for:

(top) Bitcoin (BTC),
(middle) NVDA,
(bottom) their volatility normalised log spread.

Figure 16: Strategy decomposition details:

(top) volatility normalised log spread (in blue) with 50 percentile bands (in red),
(middle) trading signal (+1 means long BTC and short NVDA and -1 the contrary),
(bottom) cumulative return of the strategy.
Figure 1
Figure 2
Within a DGM layer, the mini-batch inputs along with the output of the previous layer are transformed through a series of operations that closely resemble those in Highway Networks. Below, we present the architecture in the equations along with a visual representation of a single DGM layer in Figure 5.3:

$$S_1 = \sigma(w_1 \cdot x + b_1)$$

$$Z_\ell = \sigma(u_z,\ell \cdot x + w_z,\ell \cdot S_{\ell} + b_z,\ell) \quad \ell = 1, \ldots, L$$

$$G_\ell = \sigma(u_g,\ell \cdot x + w_g,\ell \cdot S_{\ell} + b_g,\ell) \quad \ell = 1, \ldots, L$$

$$R_\ell = \sigma(u_r,\ell \cdot x + w_r,\ell \cdot S_{\ell} + b_r,\ell) \quad \ell = 1, \ldots, L$$

$$H_\ell = \sigma(u_h,\ell \cdot x + w_h,\ell \cdot (S_{\ell} \odot R_{\ell}) + b_h,\ell) \quad \ell = 1, \ldots, L$$

$$S_{\ell+1} = (1 - G_\ell) \odot H_\ell + Z_\ell \odot S_\ell \quad \ell = 1, \ldots, L$$

where $\odot$ denotes Hadamard (element-wise) multiplication, $L$ is the total number of layers, $\sigma$ is an activation function and the $u$, $w$ and $b$ terms with various superscripts are the model parameters.

Similar to the intuition for LSTMs, each layer produces weights based on the last layer, determining how much of the information gets passed to the next layer. In Sirignano and Spiliopoulos (2018) the authors also argue that including repeated element-wise multiplication of nonlinear functions helps capture "sharp turn" features present in more complicated functions. Note that at every iteration the original input enters into the calculations of every intermediate step, thus decreasing the probability of vanishing gradients of the output function with respect to $x$.

Compared to a Multilayer Perceptron (MLP), the number of parameters in each hidden layer of the DGM network is roughly eight times bigger than the same...
Figure 5.3: Operations within a single DGM layer.

Figure 4
Figure 5
Figure 6
Figure 7
Figure 9
Figure 10
Figure 11

\[ PnL_{ML,ls} - PnL_{SC} = 20.50 \]
\[ PnL_{ML,ls} = 627.01 \]
\[ PnL_{SC} = 606.52 \]
Figure 12

\[
\hat{PnL}_{ML} - \hat{PnL}_{FM} = 13.45
\]
\[
\hat{PnL}_{ML} = 844.96
\]
\[
\hat{PnL}_{FM} = 831.52
\]
Figure 13
Figure 14
Figure 15
Figure 16