Optimal Trade Execution: Viscosity Solutions and HJB Equations

Peter Forsyth¹ S. Tse² H. Windcliff² S. Kennedy²

¹Cheriton School of Computer Science University of Waterloo

> ²Morgan Stanley New York

Global Derivatives 2013, Amsterdam, April 18 Stream C: 11:10-12:30

The Basic Problem

Broker buys/sells large block of shares on behalf of client

- Large orders will incur costs, due to price impact (liquidity) effects
 - ightarrow e.g. rapidly selling a large block of shares will depress the price
- Slow trading minimizes price impact, but leaves exposure to stochastic price changes
- Fast trading will minimize risk due to random stock price movements, but price impact will be large
- What is the optimal strategy?

An Interesting Example of Price Impact

Remember Jérôme Kerviel

- Rogue trader at Société Générale
- The book value of Kerviel's portfolio, January 19, 2008 ¹
 → -2.7 Billion €
- SocGen decided to unwind this portfolio as rapidly as possible
- Over three days, the total cost of unwinding the portfolio was $\rightarrow~-6.3$ Billion \in
- The price impact of rapid liquidation caused the realized loss to more than double the book value loss

¹Report of the Commission Bancaire

Formulation

- P = Trading portfolio
 - = B + AS
- B = Bank account: keeps track of gains/losses
- S = Price of risky asset
- A = Number of units of the risky asset
- T = Trading horizon

For Simplicity: Sell Case Only

Sell
$$t = 0 \rightarrow B = 0, S = S_0, A = A_0$$
$$t = T \rightarrow B = B_L, S = S_T, A = 0$$

- B_L is the cash generated by trading in [0, T)
 → Plus a final sale at t = T to ensure that zero shares owned.
- Success is measured by B_L (proceeds from sale).
- Maximize *E*[*B*_{*L*}], minimize *Var*[*B*_{*L*}]

Price Impact Modelling

In practice, a hierarchy of models is used

- Level 1 Considers all buy/sell orders of a large financial institution, over many assets
 - Simple model of asset price movements, considers correlation between assets
 - Output: "sell 10⁷ shares of RIM today".
- Level 2 Single name sell strategy (trade schedule over the day)
 - Level 2 models attempt to determine optimal strategy for selling a single name, assuming trades occur continuously, at rate \boldsymbol{v}
 - Price impact is a function of trade rate
 - Output: "sell 10⁵ shares of RIM between 10:15-10:45"
- Level 3 Fine grain model
 - Level 3 models assume discrete trades, and try to trade optimally based on an order book model.
 - Output: "place sell order for 1000 shares at 10:22"

We focus on Level 2 models today.

Basic Problem

Trading rate v (A = number of shares)

$$\frac{dA}{dt} = v$$
.

Suppose that S follows geometric Brownian Motion (GBM) under the objective measure

$$dS = (\eta + g(v))S \ dt + \sigma S \ dZ$$

$$\eta \text{ is the drift rate of } S$$

$$g(v) \text{ is the permanent price impact}$$

$$\sigma \text{ is the volatility}$$

$$dZ \text{ is the increment of a Wiener process}$$

٠

Basic Problem II

To avoid round-trip arbitrage (Huberman, Stanzl (2004))

 $g(v) = \kappa_p v$

 κ_p permanent price impact factor (const.)

The bank account B is assumed to follow

$$\frac{dB}{dt} = rB - vS_{exec}$$

r is the risk-free return

 S_{exec} is the execution price

 $= Sf(v)$

 $f(v)$ is the temporary price impact

 $(-vS_{exec})$ represents the rate of cash generated when buying shares at price S_{exec} at rate v (v < 0 if selling).

Temporary Price Impact: $S_{exec} = f(v)S$

Temporary price impact and transaction cost function f(v) is assumed to be

$$f(v) = [1 + \kappa_s \operatorname{sgn}(v)] \exp[\kappa_t \operatorname{sgn}(v)|v|^{\beta}]$$

$$\kappa_s \text{ is the bid-ask spread parameter}$$

$$\kappa_t \text{ is the temporary price impact factor}$$

$$\beta \text{ is the price impact exponent}$$

Optimal Strategy Define:

 $\begin{array}{rcl} X & = & (S(t), A(t), B(t)) & = & \mbox{State} \\ B_L & = & \mbox{Liquidation Value} \\ v(X,t) & = & \mbox{trading rate} \end{array}$

Let

$$\underbrace{E_{t,x}^{v(\cdot)}[\cdot]}_{Reward} = E[\cdot|X(t) = x] \text{ with } v(X(u), u), u \ge t$$

being the strategy along path $X(u), u \ge t$

$$\underbrace{\mathsf{Var}_{t,x}^{v(\cdot)}[\cdot]}_{\textit{Risk}} = \mathsf{Var}[\cdot|X(t) = x] \text{ Variance under strategy } v(\cdot)$$

Mean Variance: Standard Formulation

We construct the efficient frontier by finding the optimal control $v(\cdot)$ which solves (for fixed λ)

$$\sup_{v} \left\{ \underbrace{E^{v}[B_{L}]}_{Reward} - \lambda \underbrace{Var^{v}[B_{L}]}_{Risk} \right\}$$
(1)

- \bullet Varying $\lambda \in [0,\infty)$ traces out the efficient frontier
- $\lambda={\rm 0}; \to$ we seek only maximize cash received, we don't care about risk.
- $\lambda=\infty \rightarrow$ we seek only to minimize risk, we don't care about the expected reward.

Mean Variance: Standard Formulation

The objective is to determine the strategy $v(\cdot)$ which maximizes



Solving (2) for various λ traces out a curve in the expected value, standard deviation plane.

• Let $v_t^*(x, u), u \ge t$ be the optimal policy for (2). Then $v_{t+\Delta t}^*(x, u), u \ge t + \Delta t$ is the optimal policy which maximizes

$$\sup_{v(X(u),u\geq t+\Delta t))} \left\{ \underbrace{E_{t+\Delta t,X(t+\Delta t)}^{v}[B_{L}]}_{Reward as seen at t+\Delta t} -\lambda \underbrace{Var_{t+\Delta t,X(t+\Delta t)}^{v}[B_{L}]}_{Risk as seen at t+\Delta t} \right\}.$$

Pre-commitment Policy

However, in general

$$\underbrace{v_t^*(X(u), u)}_{optimal \ policy \ as \ seen \ at \ t} \neq \underbrace{v_{t+\Delta t}^*(X(u), u)}_{optimal \ policy \ as \ seen \ at \ t+\Delta t} ; \underbrace{u \ge t + \Delta t}_{any \ time > t+\Delta t} ,$$
(3)

 \hookrightarrow Optimal policy is not *time-consistent*. The strategy which solves problem (2) has been called the *pre-commitment* policy (Basak,Chabakauri: 2010; Bjork et al: 2010)

- Much discussion on the economic meaning of such strategies.
- Possible to formulate a time-consistent version of mean-variance.
- Or other strategies: mean quadratic variation
- Different applications may require different strategies.

Ulysses and the Sirens: A pre-commitment strategy



Ulysses had himself tied to the mast of his ship (and put wax in his sailor's ears) so that he could hear the sirens song, but not jump to his death.

Pre-commitment

Problem:

- Since the pre-commitment strategy is not time consistent, there is no natural dynamic programming principle
- We would like to formulate this problem as the solution of an HJB equation.
- How are we going to do this?

Solution:

• Use embedding technique (Zhou and Li (2000), Li and Ng (2000))

Embedding

Equivalent formulation: for fixed λ , if $v^*(\cdot)$ maximizes

$$\sup_{v(X(u),u\geq t),v(\cdot)\in\mathbb{Z}} \left\{ \underbrace{E_{t,x}^{v}[B_{L}]}_{Reward} - \lambda \underbrace{Var_{t,x}^{v}[B_{L}]}_{Risk} \right\} ,$$

$$\mathbb{Z} \text{ is the set of admissible controls}$$
(4)

then there exists a $\gamma = \gamma(t, x, E[B_L])$ such that $v^*(\cdot)$ minimizes

$$\inf_{\nu(\cdot)\in\mathbb{Z}} E_{t,x}^{\nu(\cdot)} \left[\left(B_L - \frac{\gamma}{2} \right)^2 \right] \,. \tag{5}$$

Note we have effectively replaced parameter λ by γ in (5).

Construction of Efficient Frontier

We can alternatively now regard γ as a parameter, and determine the optimal strategy $v^*(\cdot)$ which solves

$$\inf_{v(\cdot)\in\mathbb{Z}} E_{t,x}^{v(\cdot)} \left[(B_L - \frac{\gamma}{2})^2 \right] \,. \tag{6}$$

Once $v^*(\cdot)$ is known, we can easily determine $E_{t,x}^{v^*(\cdot)}[B_L]$, $E_{t,x}^{v^*(\cdot)}[(B_L)^2]$, by solving an additional linear PDE.

For given γ , this gives us $(E_{t,x}^{v^*(\cdot)}[B_L], \operatorname{Std}_{t,x}^{v^*(\cdot)}[B_L])$, a single point on the efficient frontier.

Repeating the above for different γ generates points on the efficient frontier. $^{\rm 2}$

 $^{^2 {\}rm Strictly}$ speaking, since some values of γ may not represent points on the original frontier, we need to construct the upper convex hull of these points (Tse, Forsyth, Li (2012)) .

Hamilton Jacobi Bellman (HJB) Equation

Let

$$V(s, \alpha, b, \tau) = \inf_{v(\cdot) \in \mathbb{Z}} \left\{ E_{t,x}^{v(\cdot)} \left[(B_L - \frac{\gamma}{2})^2 \mid S(t) = s, A(t) = \alpha, B(t) = b \right] \right\}$$
$$x = (s, \alpha, b)$$
$$s = \text{ stock price}$$

$$\alpha =$$
 number of units of stock

$$T =$$
 Trading horizon

$$au = T - t$$

$$\mathbb{Z} = [v_{\min}, 0]$$
 (Only selling permitted)

HJB Equation for Optimal Control $v^*(\cdot)$

We can use dynamic programming³ to solve for

$$\inf_{\nu(\cdot)\in\mathbb{Z}} \mathcal{E}_{t,x}^{\nu(\cdot)} \left[\left(B_L - \frac{\gamma}{2} \right)^2 \right] \,. \tag{7}$$

Then, using usual arguments, $V(s, \alpha, b, \tau)$ is determined by

$$V_{\tau} = \mathcal{L}V + rbV_{b} + \inf_{v \in \mathbb{Z}} \left[-vsf(v)V_{b} + vV_{\alpha} + g(v)sV_{s} \right]$$
$$\mathcal{L}V \equiv \frac{\sigma^{2}s^{2}}{2}V_{ss} + \eta sV_{s}$$
$$\mathbb{Z} = [v_{min}, 0]$$

with the payoff $V(s, lpha, b, au=0) = (b - \gamma/2)^2$. ⁴

³But this is not time-consistent since $\gamma = \gamma(t, x, E[B_L])$ ⁴But note that v is arbitrary if $V_b = V_\alpha = V_s = 0$ But solving the HJB equation requires some work

I will give a brief description of how to do this (later).

- But this is considered too complex by most
- So, the original (Almgren and Chriss) paper made several approximations (e.g. $v(\cdot)$ independent of S(t)).
- In fact, a careful read of this paper, shows that the objective function (after the approximations) is not actually mean-variance, but is mean quadratic-variation

Formally, the quadratic variation risk measure is defined as

$$E\left[\int_{t}^{T} \left(A(t')dS(t')\right)^{2}\right].$$
(8)

Informally (if P = B + AS)

$$\left(A(t')dS(t')\right)^2 = \left(dP(t')\right)^2$$

i.e. the quadratic variation of the portfolio value process.

Originally suggested as an alternate risk measure by Brugièrre (1996).

This measures risk in terms of the variability of the stock holding position, along the entire trading path.

Find optimal strategy $v(\cdot)$ which maximizes (for fixed λ)

$$\sup_{v(\cdot)\in\mathbb{Z}}\left\{\underbrace{E_{t,s,\alpha}^{v(\cdot)}\left[B_{L}\right]}_{Reward}-\lambda\underbrace{E_{t,s,\alpha}^{v(\cdot)}\left[\int_{t}^{T}\left(dP(t')\right)^{2}\right]}_{Risk}\right\}$$

where

$$B_L = \int_t^{T^-} (Cash Flows from selling) dt' + (Final Sale at t = T)$$

One can easily derive the HJB equation for the optimal control $v^*(\cdot)$

$$V_{\tau} = \eta s V_s + \frac{\sigma^2 s^2}{2} V_{ss} - \lambda \sigma^2 \alpha^2 s^2 + \sup_{v \in \mathbb{Z}} \left[e^{r\tau} (-vf(v))s + g(v)sV_s + vV_{\alpha} \right].$$

- The control is time consistent in this case
- If we assume Arithmetic Brownian Motion, then HJB equation has analytic solution (Almgren, Chriss(2001))
 - Control is independent of S(t)

One could argue that mean quadratic variation is a reasonable risk measure

- Risk is measured along the entire trading path
- In contrast, Mean variance only measures risk at end of path
- $\bullet~{\sf Time}{\text{-}{\sf consistency}} \to {\sf smoothly varying controls}$

But

Mean Quadratic Variation \neq Mean Variance

How do We Measure Performance of Trading Algorithms?

Imagine we carry out many hundreds of trades

We then examine post-trade data⁵

- Determine the realized mean return and standard deviation (relative to the pre-trade or *arrival* price)
- Assuming the modeled dynamics very closely match the dynamics in the real world
 - $\rightarrow\,$ Optimal pre-commitment Mean Variance strategy will result in the largest realized mean return, for given standard deviation
- So, if we measure performance in this way
 - We should use Mean Variance optimal control
 - But this is not what's done in industry
 - \rightarrow Effectively, a Mean Quadratic Variation Control is used (Almgren, Chriss (2001))

⁵Apparently, some clients actually do this

HJB Equations

Both mean-variance and mean quadratic variation problems reduce to solving HJB equations

Mean Variance:

$$V_{\tau} = \mathcal{L}V + rbV_{b} + \inf_{v \in \mathbb{Z}} \left[-vsf(v)V_{b} + vV_{\alpha} + g(v)sV_{s} \right]$$
$$\mathcal{L}V \equiv \frac{\sigma^{2}s^{2}}{2}V_{ss} + \eta sV_{s}$$
$$\mathbb{Z} = [v_{min}, 0]$$

Mean Quadratic Variation:

$$V_{\tau} = \mathcal{L}V - \lambda \sigma^2 \alpha^2 s^2 + \sup_{v \in \mathbb{Z}} \left[e^{r\tau} (-vf(v))s + g(v)sV_s + vV_{\alpha} \right].$$

Viscosity Solutions

In general, these types of non-linear PDEs may not have differentiable solutions

- What does it mean to solve a differential equation when the solution is not differentiable?
- Need to relax the definition of solution
- By relaxing our definition of solution, we can now have multiple solutions to our non-linear PDE
 - \rightarrow Need to ensure that our numerical method converges to the financially relevant solution (dynamic program solution)
 - \rightarrow This is the viscosity solution
- For examples of cases where seemingly reasonable discretizations converge to the incorrect solution, see (Pooley et al, 2003, IMA J. Num. Anal.))

Simple Case: 1-d HJB equation

Consider a simple case: a single factor optimal control problem

$$\mathcal{L}^{Q}V \equiv a(x, au, Q)V_{xx} + b(x, au, Q)V_{x} - c(x, au, Q)V$$

 $Q = ext{control}$
 $c(x, au, Q) \ge 0$

$$egin{array}{rcl} V_{ au}&=&\sup_{Q\in \hat{Q}}iggl\{ \mathcal{L}^{Q}V+d(x, au,Q)iggr\} \ \hat{Q}&=& ext{set of admissible controls} \end{array}$$

Both the mean-variance and mean quadratic variation optimal trading problems are multi-d generalizations of this PDE

Viscosity Solution

We can write our general HJB equation in the form

$$g(V_{xx}, V_x, V_\tau, V, x, \tau) = V_\tau - \sup_{Q \in \hat{Q}} \left\{ \mathcal{L}^Q V + d(x, \tau, Q) \right\}$$
$$= 0$$

Suppose we have a C^{∞} test function ϕ such that $\phi \geq V$, and ϕ touches V at a single point (x_0, τ_0) . For simplicity, assume that V is continuous. Otherwise, ϕ should touch the upper semi-continuous envelope of V.

Subsolution

Figure: If, for any point (x_0, τ_0) , for any test function $\phi \ge V$, where ϕ touches V at the single point (x_0, τ_0) , $g(\phi_{xx}, \phi_x, \phi_\tau, \phi, x_0, \tau_0) \le 0$, then V is a viscosity subsolution.



Supersolution

Figure: If, for any point (x_0, τ_0) , for any test function $\phi \leq V$, where ϕ touches V at the single point (x_0, τ_0) , $g(\phi_{xx}, \phi_x, \phi_\tau, \phi, x_0, \tau_0) \geq 0$, then V is a viscosity supersolution.



Viscosity Solution

Any solution which is both a subsolution and a supersolution is a *viscosity solution*

Note that we never evaluate $g(V_{xx}, V_x, V_\tau, ...)$ but only $g(\phi_{xx}, \phi_x, \phi_\tau, ...)$, \rightarrow no problems with non-differentiable V.

Numerical issues:

- We want to ensure that our numerical scheme converges to the viscosity solution
- Sufficient conditions known which ensure that a numerical scheme converges to the viscosity solution (Barles, Souganidis (1991))

Technical Point I

Remark

There may also be some points where a smooth C^{∞} test function cannot touch the solution from either above or below. As a pathological example, consider the function

$$f(x) = \begin{cases} \sqrt{x} & x \ge 0, \\ -\sqrt{-x} & x < 0. \end{cases}$$
(9)

This function cannot be touched at the origin from below (or above) by any smooth function with bounded derivatives. Note that the definition of a viscosity solution only specifies what happens when the test function touches the viscosity solution at a single point (from either above or below). The definition is silent about cases where this cannot happen.

Discretization: General Form

Let V_i^n be the approximate value of the solution, i.e. $V_i^n \simeq V(x_i, \tau^n)$.

Then we can write a general discretization of the HJB equation $g(V_{xx}, V_x, V_\tau, V, x, \tau)$ at node (x_i, τ^{n+1})

$$G_{i}^{n+1}(h, V_{i}^{n+1}, V_{i+1}^{n+1}, V_{i-1}^{n+1}, V_{i}^{n}, V_{i+1}^{n}, V_{i-1}^{n}, ...) = G_{i}^{n+1}(h, V_{i}^{n+1}, \{V_{m}^{p}\}_{\substack{p \neq n+1 \\ \text{or } m \neq i}}) = 0$$

 $\{V_m^p\}_{\substack{p \neq n+1 \\ \text{or } m \neq i}}$ refers to the discrete solution values at nodes neighbouring (in space and time) node (x_i, τ^{n+1}) .

h is the mesh size/timestep size parameter

Example: Linear Heat Equation Discretize

$$egin{array}{rcl} V_t &= V_{xx} \ V(x=0,t) = V(x=1,t) = 1 \end{array}$$

using a mesh $x_i = i\Delta x, i = 0, ..., i_{max}, t^n = n\Delta t, V(x_i, t^n) \simeq V_i^n$.

$$\frac{V_i^{n+1} - V_i^n}{\Delta t} = \frac{V_{i-1}^{n+1} - 2V_i^{n+1} + V_{i+1}^{n+1}}{(\Delta x)^2}$$
$$\Delta t = h \; ; \; \Delta x = h$$

$$\begin{aligned} G_{i=0}^{n+1} &= V_{i=0}^{n+1} - 1 \\ G_{i}^{n+1} &= \frac{V_{i}^{n+1} - V_{i}^{n}}{h} - \left(\frac{V_{i-1}^{n+1} - 2V_{i}^{n+1} + V_{i+1}^{n+1}}{h^{2}}\right); \quad i \neq 0, i_{\max} \\ G_{i=i_{\max}}^{n+1} &= V_{i=i_{\max}}^{n+1} - 1 \end{aligned}$$

Sufficient Conditions: Convergence

We use the results from (Barles, Sougandinis, 1991).

Theorem (Convergence)

Any numerical scheme which is consistent, I_{∞} stable, and monotone, converges to the viscosity solution.

Stability

- Usual requirement (discrete solution bounded in I_∞ as mesh, timestep \rightarrow 0)
- Can prove using maximum analysis
- Usually, only fully implicit timestepping is unconditionally I_{∞} stable (but not always, e.g. jump diffusions).

Monotonicity

Definition (Monotonicity)

The discretization

$$G_i^{n+1}(h, V_i^{n+1}, \left\{V_m^p\right\}_{\substack{p \neq n+1 \\ \text{or } m \neq i}})$$

is monotone if

$$\begin{array}{ll} G_i^{n+1}(h,V_i^{n+1},\left\{X_m^p\right\}_{\substack{p\neq n+1\\ \text{or }m\neq i}}) &\leq & G_i^{n+1}(h,V_i^{n+1},\left\{Y_m^p\right\}_{\substack{p\neq n+1\\ \text{or }m\neq i}}) \\ & \forall & X_m^p \geq Y_m^p, \forall m,p \end{array}$$

• Simple algebraic condition: easy to check

Monotonicity cont'd

- A discretization which is a positive coefficient method is monotone (usually)
- For 1-d problems, multi-dimensional problems with diffusion only in one direction, forward/backward/central differencing generate a positive coefficient scheme
- Often results in low order schemes (first order)
- Has nice financial interpretation: enforces discrete no-arbitrage inequalities. i.e. if payoff from contingent claim A is greater than claim B (on the same underlying) then the value of Claim A must always be greater than Claim B at all earlier times (independent of mesh size h).

Consistency

The precise definition of consistency in the viscosity sense is quite long winded:

- If we assume that the solution and $g(\cdot)$ are continuous everywhere
 - $\rightarrow~$ This boils down to classical definition of consistency
- Can be proven in any given case by using Taylor series applied to infinitely differentiable test functions.
- In general, the solution may not be continuous at the boundary
 - In this case, the definition of consistency in the viscosity sense is more relaxed than usual definition
 - Discretized equation $G(\cdot)$ only has to converge to an operator which is
 - \leq max limit of $g(\cdot)$ as we approach the boundary
 - \geq min limit of $g(\cdot)$ as we approach the boundary

(10)

Why do we ever need this complex definition of consistency?

At points near the boundary, the viscosity definition of consistency is more relaxed than the classical definition of consistency.

- Sometimes, a scheme (i.e. a semi-Lagrangian method) may never be consistent (in the classical sense) near the boundary.
- But the scheme is consistent in the viscosity sense!
- Consistency in the viscosity sense can be very useful in such situations.

Summary: what do we need to know about Viscosity Solutions?

- Convergence to the viscosity solution ensured if discretization is Consistent, I_{∞} stable and monotone.
- Consistency: most of the time, this is just classical consistency, applied to infinitely differentiable test functions
 - $\rightarrow\,$ Consistency in the viscosity sense is very forgiving, when it comes to points near the boundaries
- I_∞ Stability: standard maximum analysis can be used to prove this
- Monotonicity: most interesting condition: preserves discrete arbitrage inequalities
 - ightarrow Positive coefficient discretization guarantees this property
- Seemingly reasonable discretizations can converge to wrong values if these conditions are not satisfied!

Back to our Optimal Trade Execution Problem

Two Objective Functions:

Mean Variance: Find trading rate $v(\cdot)$ which minimizes

$$\left\{\underbrace{E_{(t,x)}^{\nu(\cdot)}[B_L]}_{Reward} - \lambda \underbrace{Var_{(t,x)}^{\nu(\cdot)}[B_L]}_{Risk}\right\}$$
(11)

where

$$x = (s, \alpha, b)$$

$$s = \text{stock price}$$

 $\alpha ~=~$ number of units of stock

$$B_L = \text{gain from selling}$$
 (12)

- This would be the correct objective function if we evaluated trading efficiency by looking at post-trade data
- But it is not time consistent

This is the industry standard (i.e. Almgren and Chriss)

Mean Quadratic Variation Find trading rate $v(\cdot)$ which minimizes

$$\left\{\underbrace{E_{t,s,\alpha}^{\nu(\cdot)}\left[B_{L}\right]}_{Reward} -\lambda \underbrace{E_{t,s,\alpha}^{\nu(\cdot)}\left[\int_{t}^{T} \left(dP(t')\right)^{2}\right]}_{Risk}\right\}$$
(13)

where

$$s = \text{stock price}$$

$$\alpha =$$
 number of units of stock

$$B_L = \text{gain from selling}$$
 (14)

- This would be the correct objective function if we want to control a measure of risk *during* the trade execution, not just at the end
- It is time consistent, but it's not mean variance!

Recall the HJB Equations

Both mean-variance and mean quadratic variation problems reduce to solving HJB equations

Mean Variance:

$$V_{\tau} = \mathcal{L}V + rbV_{b} + \inf_{v \in \mathbb{Z}} \left[-vsf(v)V_{b} + vV_{\alpha} + g(v)sV_{s} \right]$$
$$\mathcal{L}V \equiv \frac{\sigma^{2}s^{2}}{2}V_{ss} + \eta sV_{s}$$
$$\mathbb{Z} = [v_{min}, 0]$$

Mean Quadratic Variation:

$$V_{\tau} = \eta s V_s + \frac{\sigma^2 s^2}{2} V_{ss} - \lambda \sigma^2 \alpha^2 s^2 + \sup_{v \in \mathbb{Z}} \left[e^{r\tau} (-vf(v))s + g(v)sV_s + vV_{\alpha} \right]$$

HJB Equation: Mean Variance

Define the Lagrangian derivative

$$\frac{DV}{D\tau}(v) = V_{\tau} - V_s g(v) s - V_b (rb - vf(v)s) - V_{\alpha}v,$$

which is the rate of change of V along the characteristic curve

$$s = s(\tau)$$
; $b = b(\tau)$; $\alpha = \alpha(\tau)$

defined by the trading velocity v through

$$rac{ds}{d au} = -g(v)s, \quad rac{db}{d au} = -(rb - vf(v)s), \quad rac{dlpha}{d au} = -v.$$

HJB Equation: Lagrangian Form

We can then write the Mean Variance HJB equation as

$$\mathcal{L}V - \sup_{v(\cdot)\in Z} rac{DV}{D au}(v) = 0.$$

 $\mathcal{L}V \equiv rac{\sigma^2 s^2}{2} V_{ss} + \eta s V_s$

Numerical Method:

- Discretize the Lagrangian form directly (semi-Lagrangian method)
- Timestepping algorithm
 - Solve local optimization problem at each grid node
 - Discretized linear PDE solve to advance one timestep
- Provably convergent to the viscosity solution of the HJB PDE
- Similar approach for the Mean Quadratic Variation HJB PDE

Numerical Method: Efficient Frontier

Recall that (Mean Variance)

$$V(s, lpha, b, au = 0) = (b - \gamma/2)^2$$

Numerical Algorithm

- Pick a value for γ
 - Solve HJB equation for optimal control $v = v(s, \alpha, b, \tau)$
 - Store control at all grid points
 - Simulate trading strategy using a Monte Carlo method (use stored optimal controls)
 - Compute mean, standard deviation
 - This gives a single point on the efficient frontier
- Repeat

Similar approach for Mean Quadratic Variation

Numerical Examples

Simple case: GBM, zero drift, zero permanent price impact

$$dS = \sigma S \ dZ$$

Temporary Price Impact:

$$f(v) = exp(\kappa_t v)$$

Т	r	S _{init}	$\alpha_{\textit{init}}$	Action	V _{min}
1/250	0.0	100	1.0	Sell	-1000/T
(One Day)					

Case	σ	κ_t	Percentage of Daily Volume
1	1.0	$2 imes 10^{-6}$	16.7%
2	0.2	$2.4 imes10^{-6}$	20.0%
3	0.2	$6 imes 10^{-7}$	5.0%
4	0.2	$1.2 imes 10^{-7}$	1.0%
5	0.2	$2.4 imes10^{-8}$	0.2%

 $\sigma = 1.0$, 16.7% daily volume, $S_{init} = 100$



 σ = .2, 20% daily volume, S_{init} = 100



 σ = .2, 5% daily volume, S_{init} = 100



 σ = .2, 1% daily volume, S_{init} = 100



 σ = .2, 0.2% daily volume, S_{init} = 100



Optimal trading rate: $t = 0, \alpha = 1, b = 0$



Mean Share Position (α) vs. Time



Standard Deviation of Share Position (α) vs. Time



Conclusions: Mean Variance

Pros:

- If performance is measured by post-trade data (mean and variance)
 - $\rightarrow~$ This is the truly optimal strategy
- Significantly outperforms Mean Quadratic Variation for low levels of required risk (fast trading)

Cons:

- Non-trivial to compute optimal strategy
- Very aggressive in-the-money strategy
- Share position has high standard deviation
- Optimal trading rate is *almost ill posed*: many nearby strategies give almost same efficient frontier in some cases

 $\rightarrow\,$ Simple example: zero standard deviation 6

⁶Recall that v is arbitrary if $V_s = V_b = V_{\alpha} = 0$

Conclusions: Mean Quadratic Variation

Pros:

- Simple analytic solution for Arithmetic Brownian Motion Case
- Trading rate a smooth, predictable function of time
 - ightarrow For GBM case, only weakly sensitive to asset price S
- Almost same results as Mean Variance, for large levels of required risk (slow trading)

Cons:

- If performance is measured by post-trade data (mean and variance)
 - $\rightarrow~$ This is not the optimal strategy
 - \rightarrow Significantly sub-optimal for low levels of risk (fast trading)