Trading Regions Under Proportional Transaction Costs

Karl Kunisch[1] and Jörn Sass[2]

[1] Institute of Mathematics, Karl-Franzens University Graz, Heinrichstr. 36, 8010 Graz, Austria karl.kunisch@uni-graz.at
[2] RICAM, Austrian Academy of Sciences, Altenberger Str. 69, 4040 Linz, Austria joern.sass@oeaw.ac.at

Summary. In the Black-Scholes model optimal trading for maximizing expected power utility under proportional transaction costs can be described by three intervals B, NT, S: If the proportion of wealth invested in the stocks lies in B, NT, S, then buying, not trading and selling, respectively, are optimal. For a finite time horizon, the boundaries of these trading regions depend on time and on the terminal condition (liquidation or not). Following a stochastic control approach, one can derive parabolic variational inequalities whose solution is the value function of the problem. The boundaries of the active sets for the different inequalities then provide the boundaries of the trading regions. We use a duality based semi-smooth Newton method to derive an efficient algorithm to find the boundaries numerically.

1 Trading Without Transaction Costs

The continuous-time Black Scholes model consists of one bond or bank account and one stock with prices $(P_0(t))_{t\in[0,T]}$ and $(P_1(t))_{t\in[0,T]}$ which for interest rate $r \geq 0$, trend $\mu \in \mathbb{R}$, and volatility $\sigma > 0$ evolve according to

$$\mathrm{d}P_0(t) = P_0(t)\, r\, \mathrm{d}t\ , \quad \mathrm{d}P_1(t) = P_1(t)\, (\mu\, \mathrm{d}t + \sigma\, \mathrm{d}W(t))\ , \quad P_0(0) = P_1(0) = 1\ ,$$

where $W = (W(t))_{t\in[0,T]}$ is a Brownian motion on a probability space $(\Omega, \mathcal{A}, P)$. Let $\mathcal{F} = (\mathcal{F}_t)_{t\in[0,T]}$ denote the augmented filtration generated by W.

Without transaction costs the trading of an investor may be described by initial capital $\underline{x} > 0$ and risky fraction process $(\eta(t))_{t\in[0,T]}$, where $\eta(t)$ is the fraction of the portfolio value (wealth) which is held in the stocks at time t. The corresponding *wealth process* $(X(t))_{t\in[0,T]}$ is defined self-financing by

$$\mathrm{d}X(t) = (1-\eta(t))X(t)r\, \mathrm{d}t + \eta(t)\, X(t)\, (\mu\, \mathrm{d}t + \sigma\, \mathrm{d}W(t))\ , \quad X(0) = \underline{x}\ .$$

The utility of terminal wealth $x > 0$ is given by power utility $\frac{1}{\alpha}\, x^\alpha$ for any $\alpha < 1$, $\alpha \neq 0$. The parameter α models the preferences of an investor. The limiting case $\alpha \to 0$ corresponds to logarithmic utility i.e. maximizing the expected rate of return, $\alpha > 0$ corresponds to less risk averse and $\alpha < 0$ to more risk averse utility functions. Merton showed that for logarithmic ($\alpha = 0$) and power utility the optimal trading strategy is given by a constant optimal risky fraction

$$\eta(t) = \hat{\eta}\,, \quad t \in [0, T], \quad \text{where} \quad \hat{\eta} = \frac{1}{1-\alpha}\frac{\mu - r}{\sigma^2}\,. \tag{1}$$

2 Proportional Transaction Costs

To keep the risky fraction constant like in (1) involves continuous trading which, under transaction costs, is no longer adequate. For possible cost structures see e.g. [5, 7, 8]. We consider proportional costs $\gamma \in (0, 1)$ corresponding to the proportion of the traded volume which has to be paid as fees.

For suitable infinite horizon criteria solution of the corresponding Hamilton-Jacobi-Bellmann equation (HJB) leads to a characterization of the optimal wealth process as a diffusion reflected at the boundaries of a cone, see [3, 9]. When reaching the boundaries of the cone, infinitesimal trading occurs in such a way that the wealth process just stays in the cone. The cone corresponds to an interval for the risky fraction process. The existence of a viscosity solution for the HJB equation for finite time horizon is shown in [1] and numerically treated in [10] using a finite difference method.

Now let us fix costs $\gamma \in (0, 1)$ and parameters $\alpha < 1$, $\alpha \neq 0$, r, μ, σ such that $\hat{\eta} \in (0, 1)$. The trading policy can be described by two increasing processes $(L(t))_{t\in[0,T]}$ and $(M(t))_{t\in[0,T]}$ representing the cumulative purchases and sales of the stock. We require that these are right-continuous, $\mathcal{F}$-adapted, and start with $L(0-) = M(0-) = 0$. Transaction fees are paid from the bank account. Thus the dynamics of the controlled wealth processes $(X_1(t))_{t\in[0,T]}$ and $(X_0(t))_{t\in[0,T]}$, corresponding to the amount of money on the bank account and the amount invested in the stocks, are

$$\begin{aligned}
\mathrm{d}X_0(t) &= rX_0(t)\,\mathrm{d}t - (1+\gamma)\,\mathrm{d}L(t) + (1-\gamma)\,\mathrm{d}M(t)\,,\\
\mathrm{d}X_1(t) &= \mu X_1(t)\,\mathrm{d}t + \sigma X_1(t)\,\mathrm{d}W(t) + \mathrm{d}L(t) - \mathrm{d}M(t)\,.
\end{aligned}$$

The objective is the maximization of expected utility at the terminal trading time T, now over all control processes $(L(t))_{t\in[0,T]}$ and $(M(t))_{t\in[0,T]}$ which satisfy the conditions above and for which the wealth processes X_0 and X_1 stay positive and the total wealth strictly positive i.e. $(X_0(t), X_1(t)) \in \mathcal{D} := \mathbb{R}_+^2 \setminus \{(0, 0)\}$, $t \in [0, T]$. So suppose $(\underline{x}_0, \underline{x}_1) = (X_0(0-), X_1(0-)) \in \mathcal{D}$. We distinguish the maximization of expected utility for the terminal total wealth,

$$\tilde{J}(t, x_0, x_1) = \sup_{(L,M)} \mathrm{E}[\tfrac{1}{\alpha}\,(X_0(T) + X_1(T))^\alpha \,|\, X_0(t) = x_0, X_1(t) = x_1]\,,$$

and of the terminal wealth after liquidating the position in the stocks,

$$J(t,x_0,x_1) = \sup_{(L,M)} \mathrm{E}[\tfrac{1}{\alpha}\,(X_0(T)+(1-\gamma)X_1(T))^\alpha \mid X_0(t)=x_0, X_1(t)=x_1]\ .$$

We always assume $\tilde{J}(0,x_0,x_1) < \infty$ for all $(x_0,x_1) \in \mathcal{D}$.

Theorem 1. *J is concave, continuous, and a viscosity solution of*

$$\max\{J_t + \mathcal{A}J, -(1+\gamma)J_{x_0} + J_{x_1}, (1-\gamma)J_{x_0} - J_{x_1}\} = 0 \qquad (2)$$

on $[0,T)\times\mathcal{D}$ with $J(T,x_0,x_1) = \frac{1}{\alpha}(x_0+(1-\gamma)x_1)^\alpha$. $J_t, J_{x_0}, J_{x_1}, J_{x_1,x_1}$ denote the partial derivatives of $J = J(t,x_0,x_1)$ and the differential operator $\mathcal{A}$ (generator of (X_0,X_1)) is defined by

$$\mathcal{A}h(x_0,x_1) = r\,x_0 h_{x_0}(x_0,x_1) + \mu\,x_1 h_{x_1}(x_0,x_1) + \tfrac{1}{2}\sigma^2 x_1^2 h_{x_1,x_1}(x_0,x_1)$$

for all smooth functions h. Further J is unique in the class of continuous functions satisfying $|h(t,x_0,x_1)| \le K(1+(x_0^2+x_1^2)^\alpha)$ for all $(x_0,x_1)\in\mathcal{D}$, $t\in[0,T]$, and some constant K.

The proof in [1] is based on the derivation of a weak dynamic programming principle leading to the HJB (2). The uniqueness is shown following the Ishii technique, see [2]. Using $\tilde{J}(T,x_0,x_1) = \frac{1}{\alpha}$, the argument is the same for $\tilde{J}$.

The variational inequalities in (2) are active, if it is optimal not to trade, to buy stocks, and to sell stocks, respectively. At time t, $\mathcal{D}$ can be split into the buy region $B'(t)$, the sell region $S'(t)$, and the no trading region $NT'(t)$,

$$\begin{aligned} B'(t) &= \{(x_0,x_1)\in\mathcal{D} \,:\, -(1+\gamma)J_{x_0}(t,x_0,x_1) + J_{x_1}(t,x_0,x_1) = 0\}\ , \\ S'(t) &= \{(x_0,x_1)\in\mathcal{D} \,:\, (1-\gamma)J_{x_0}(t,x_0,x_1) - J_{x_1}(t,x_0,x_1) = 0\}\ , \\ NT'(t) &= \mathcal{D}\setminus(B'(t)\cup S'(t))\ . \end{aligned}$$

If $x_0 = 0$ ($x_1 = 0$) we should exclude the second (third) inequality in (2) since buying (selling) is not admissible. But due to the $\hat{\eta}\in(0,1)$ we expect that (x_0,x_1) lies for all t in $NT'(t)\cup S'(t)$ if $x_0 = 0$ and in $NT'(t)\cup B'(t)$ if $x_1 = 0$, cf. [9]. Thus we did not specify the different cases in Theorem 1.

3 Reduction to a One-Dimensional Problem

We use a transformation to the risky fractions, different to e.g. [10] where the fractions modified by the transaction costs are used. From the definition of J and $\tilde{J}$ one verifies directly that on $[0,T]\times\mathcal{D}$

$$J(t,x_0,x_1) = x^\alpha J(t,\tfrac{x_0}{x},\tfrac{x_1}{x})\ ,\quad \tilde{J}(t,x_0,x_1) = x^\alpha \tilde{J}(t,\tfrac{x_0}{x},\tfrac{x_1}{x})\ ,\quad x = x_0 + x_1\ .$$

So it is enough to look at the risky fractions $y = \frac{x_1}{x}$. Introducing

$$V(t,y) = J(t,1-y,y)\ ,\quad \tilde{V}(t,y) = \tilde{J}(t,1-y,y)\ ,\quad y\in[0,1]$$

we get for $J(t,x_0,x_1) = (x_0+x_1)^\alpha V(t,\frac{x_1}{x_0+x_1})$ (writing $x = x_0+x_1$)

$$J_t = x^\alpha V_t\ ,\quad J_{x_0} = x^{\alpha-1}\,(\alpha V - y\,V_y)\,,\quad J_{x_1} = x^{\alpha-1}\,(\alpha V + (1-y)\,V_y)\,,$$
$$J_{x_1,x_1} = x^{\alpha-2}\left((1-y)^2 V_{y,y} - 2(1-y)(1-\alpha)V_y - \alpha(1-\alpha)V\right).$$

Plugging this into (2) we have to solve

$$\max\{V_t + \mathcal{L}V,\ \mathcal{L}_B V,\ \mathcal{L}_S V\} = 0 \tag{3}$$

with $V(T,y) = \frac{1}{\alpha}(1-\gamma\,y)^\alpha$, $y\in[0,1]$, and operators

$$\begin{aligned}\mathcal{L}h(y) &= \alpha\left(r + (\mu-r)y - \tfrac{1}{2}(1-\alpha)\sigma^2y^2\right)\\ &\quad + \left((\mu-r)(1-y) - (1-\alpha)\sigma^2 y(1-y)\right) y h_y(y) + \tfrac{1}{2}\sigma^2y^2(1-y)^2 h_{y,y}(y)\ ,\\ \mathcal{L}_B h(y) &= (1+\gamma y)h_y - \alpha\gamma h\ ,\quad \mathcal{L}_S h(y) = -(1-\gamma y)h_y - \alpha\gamma h\ .\end{aligned}$$

The same applies to $\tilde V$ using the terminal condition $\tilde V(T,y) = \frac{1}{\alpha}$ instead. The trading regions are now given by $B(t) = \{y\in[0,1] : \mathcal{L}_B V(t,y) = 0\}$, $S(t) = \{y\in[0,1] : \mathcal{L}_S V(t,y) = 0\}$ and $NT(t) = [0,1]\setminus(B(t)\cup S(t))$ corresponding to buying, selling and not trading, respectively. On $B(t)$ and $S(t)$ we hence know that V satisfies as a solution of $\mathcal{L}_B V = 0$ and $\mathcal{L}_S V = 0$

$$V(t,y) = C_B(t)(1+\gamma y)^\alpha \quad\text{and}\quad V(t,y) = C_S(t)(1-\gamma y)^\alpha\ . \tag{4}$$

As proven in many cases we assume that $B(t)$, $N(t)$, $S(t)$ are intervals. So they can be described by their boundaries

$$a(t) = \inf NT(t)\ ,\quad b(t) = \sup NT(t)\ . \tag{5}$$

From the condition $\hat\eta\in(0,1)$ it is reasonable to expect that $NT(t)\neq\emptyset$ for all $t\in[0,T)$. But since borrowing and short selling are not allowed, this might not be true for all $B(t)$ and $S(t)$. If that happens we need boundary conditions different from (4) to solve for V on $NT(t)$. These are

$$V_t(t,0) + \alpha\,rV(t,0) = 0\ ,\ \text{if}\quad 0\in NT(t)\ , \tag{6}$$
$$V_t(t,1) + \alpha(\mu - \tfrac{1}{2}(1-\alpha)\sigma^2)V(t,1) = 0\ ,\ \text{if}\quad 1\in NT(t)\ . \tag{7}$$

4 A Semi-Smooth Newton Method

The algorithm we present to solve (3) is based on a primal-dual active set strategy, see e.g. [4] where the relationship to semismooth Newton methods is explored, or for its parabolic version cf. [6] where it is also applied to find the exercise boundary for an American option. Here we face two free boundaries and a different type of constraints and hence have to adapt the algorithm. A more detailed analysis including convergence and existence of the Lagrange multipliers is work in progress and deferred to a future publication. However, Example 1 below shows that the algorithm can work efficiently.

Problem (3) is equivalent to solving

$$V_t + \mathcal{L}V + \lambda_B + \lambda_S = 0\,, \tag{8}$$
$$\mathcal{L}_B V \le 0,\ \lambda_B \ge 0,\ \lambda_B \mathcal{L}_B V = 0\,, \quad \mathcal{L}_S V \le 0,\ \lambda_S \ge 0,\ \lambda_S \mathcal{L}_S V = 0\,. \tag{9}$$

The two complementarity problems in (9) can be written as

$$\lambda_B = \max\{0, \lambda_B + c\,\mathcal{L}_B V\}\,, \quad \lambda_S = \max\{0, \lambda_S + c\,\mathcal{L}_S V\} \tag{10}$$

for any constant $c > 0$. So we have to solve (8), (10). At T the trading regions are given by $S(T) = [0,1]$ for V and $NT(T) = [0,1]$ for $\tilde{V}$. We split $[0,T]$ in N intervals and go backwards in time with $t_N = T$, $t_n = t_{n+1} - \Delta t$, $\Delta t = T/N$. Having computed $V(t_{n+1}, \cdot)$ and the corresponding regions we use the following algorithm to compute $v = V(t_n, \cdot)$ and $NT(t_n)$:

0. Set $\overline{v} = V(t_{n+1}, \cdot)$, $k = 0$, choose an interval NT_0 in $[0,1]$, constant $c > 0$.
1. Define the boundaries a_k and b_k of NT_k as in (5).
2. On $[a_k, b_k]$ solve the elliptic problem $\frac{1}{\Delta t}(\overline{v} - v) + \mathcal{L}v = 0$ using the boundary conditions $\mathcal{L}_B v = 0$ if $a_k \notin NT_k$, (6) if $a_k \in NT_k$ (implying $a_k = 0$) and $\mathcal{L}_S v = 0$ if $b_k \notin NT_k$, (7) if $b_k \in NT_k$ (implying $b_k = 1$).
3. If $a_k \neq 0$ define v on $[0, a_k)$ by the first equation in (4). If $b_k \neq 1$ define v on $(b_k, 1]$ by the second equation in (4). Choose C_B and C_S such that v is continuous in a_k and b_k. So $v_{k+1} = v$ is continuously differentiable.
4. Set $\lambda_B^{k+1} = -\frac{1}{\Delta t}(\overline{v} - v_{k+1}) - \mathcal{L}v_{k+1}$ on $[0, a_k]$ and $\lambda_S^{k+1} = -\frac{1}{\Delta t}(\overline{v} - v_{k+1}) - \mathcal{L}v_{k+1}$ on $[b_k, 1]$ and set them to 0 otherwise.
5. Introduce the active sets
$$B_{k+1} = \{y \in [0,1] \,:\, \lambda_B^{k+1}(y) + c\,\mathcal{L}_B v_{k+1}(y) > 0\}\,,$$
$$S_{k+1} = \{y \in [0,1] \,:\, \lambda_S^{k+1}(y) + c\,\mathcal{L}_S v_{k+1}(y) > 0\}$$
and set $NT_{k+1} = [0,1] \setminus (B_{k+1} \cup S_{k+1})$. Verify that the interval structure holds and define the boundaries a_{k+1} and b_{k+1} by (5).
6. If $a_{k+1} = a_k$ and $b_{k+1} = b_k$ then STOP; otherwise increase k by 1 and continue with step 1.

Example 1. We consider a money market and a stock with parameters $r = 0$, $\mu = 0.096$, $\sigma = 0.4$, and horizon $T = 1$. We use mesh sizes $\Delta t = 0.01$ and $\Delta y = 0.001$, choose $c = 1$, and at t_{N-1} use $NT_0 = (0.1, 0.8)$, and at all other time steps t_n use $NT_0 = NT(t_{n+1})$. For the utility function we consider both $\alpha = 0.1$ and the more risk averse parameter $\alpha = -1$. These yield without transaction costs optimal risky fractions 0.667 and 0.3 (dotted lines in Figs. 1 and 2). We consider proportional costs $\gamma = 0.01$. In Fig. 1 we look at $\alpha = 0.1$, left-hand at V with liquidation at the end, right-hand at $\tilde{V}$. We see that the liquidation costs we have to pay at T imply that we also trade close to the terminal time, while without liquidation this is never optimal. In Fig. 2 we plotted the trading regions for the more risk averse parameter $\alpha = -1$, which leads to less holdings in the stock.

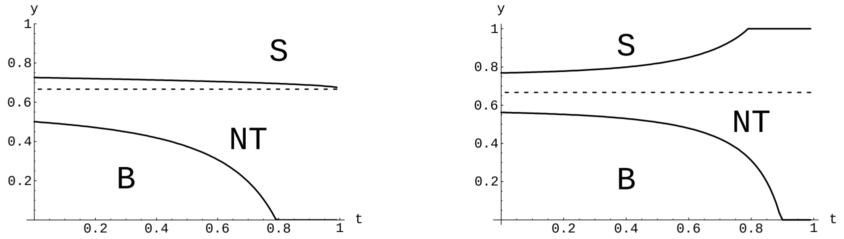

Fig. 1. Trading regions for $\alpha = 0.1$ for V and $\tilde{V}$

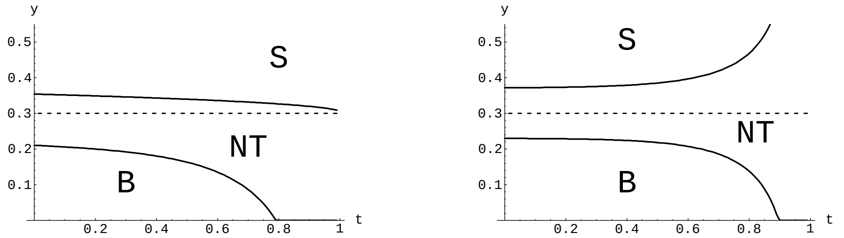

Fig. 2. Trading regions for $\alpha = -1$ for V and $\tilde{V}$

References

1. Akian M, Sulem A, Séquier P (1996) A finite horizon multidimensional portfolio selection problem with singular transactions. Proceedings of the 34th Conference on Decisions & Control, New Orleans: 2193–2197
2. Crandall MG, Ishii H, Lions, P-L (1992) User's guide to viscosity solutions of second order partial differential equations. Bulletin of the American Mathematical Society 27: 1–67
3. Davis MHA, Norman AR (1990) Portfolio selection with transaction costs. Mathematics of Operations Research 15: 676–713
4. Hintermüller M, Ito K, Kunisch K (2003) The primal-dual active ste stratgey as a semismooth Newton method. SIAM Journal on Optimization 13: 865–888
5. Irle A, Sass J (2005) Good portfolio strategies under transaction costs: A renewal theoretic approach. In: do Rosario Grossinho M, Shiryaev AN, Esquivel ML, Oliveira PE (eds) Stochastic Finance, Springer, New York: pp. 321–341
6. Ito K, Kunisch K (2006) Parabolic variational inequalities: The Lagrange multiplier approach. Journal de Mathḿatiques Pures et Appliquées (9) 85: 415–449
7. Korn R (1997) Optimal Portfolios: Stochastic Models for Optimal Investment in Continuous Time. World Scientific, Singapore
8. Sass J (2005) Portfolio optimization under transaction costs in the CRR model. Mathematical Methods of Operations Research 61: 239–259
9. Shreve SE, Soner HM (1994) Optimal investment and consumption with transaction costs. Annals of Applied Probability 4, 609–692
10. Sulem A (1997) Dynamic optimization for a mixed portfolio with transaction costs. In: Rogers LCG, Talay D (eds) Numerical Methods in Finance. Cambridge University Press, Cambridge: 165–180