Numerical Solution to the Optimal Portfolio Selection Problem for Power Utility Function

by

Kenrick Raymond So
kenrickrayso@gmail.com

Chapter 1

Numerical Solution to the Optimal Portfolio Selection Problem for Power Utility Function

1.1 Problem Setup

Consider a financial market containing two financial assets that are traded continuously on a finite horizon $[0, T]$. The first asset is a risk-free bond P that evolves according to the following ordinary differential equation

$$dP_t = rP_t dt, \quad t \in [0, T], \tag{1.1}$$

where r is a risk-free rate. The second asset ia risky stock that evolves according to the following stochastic differential equation

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \quad t \in [0, T] \tag{1.2}$$

where $\mu \in \mathbb{R}$ is a constant drift term, $\sigma \geq 0$ is a constant volatility term, and W_t is a standard Brownian motion.

The investor is interested in determining an efficient strategy for her final wealth. Define the wealth process X_t, at any time $t < T$, the investor needs to decide what proportion u_t of her wealth to invest in the risky asset S_t and invest the remaining proportion $1 - u_t$ to the risk-free bond. The wealth process must evolve according to the following stochastic differential equation

$$\begin{aligned} dX_t &= u_t X_t \frac{dS_t}{S_t} + (1 - u_t) X_t r dt \\ &= u_t X_t (\mu dt + \sigma dW_t) + (1 - u_t) X_t r dt \\ &= u_t X_t \mu dt + u_t X_t \sigma dW_t + X_t r dt - u_t X_t r dt \\ &= [r + u_t(\mu - r)] X_t dt + X_t u_t \sigma dW_t. \end{aligned} \tag{1.3}$$

Assume that the initial wealth is positive i.e. $X_0 = x_0 > 0$. The investor is interested in finding an optimal investment strategy u_t^* such that the expected utility of the terminal wealth X_T is maximized. The objective function is given by,

$$\max_{u(\cdot)\in\mathcal{A}} \mathbb{E}[U(X_T^u)]. \tag{1.4}$$

Here, $\mathcal{A}$ is the set of admissible controls if the investor had an initial endowment of x_0, and $U(x)$ is a utility function that is strictly increasing and concave up. We apply the power utility function such that the objective function becomes,

$$\max_{u(\cdot)\in\mathcal{A}} \mathbb{E}\left[\frac{(X_T^u)^\gamma}{\gamma}\right]. \quad \gamma \in (0,1) \tag{1.5}$$

By the martingale optimality principle and the verification theorem for Hamilton-Jacobi-Bellman (HJB) equation for the Merton problem, the portfolio problem (1.4) can be evaluated by solving the following HJB-equation

$$\begin{cases} \frac{\partial V}{\partial t} + \max_{u_t\in\mathcal{A}}\left\{[u_t(\mu-r)+r]X_t\frac{\partial V}{\partial X_t} + \frac{1}{2}X_t^2\sigma^2u_t^2\frac{\partial^2 V}{(\partial X_t)^2}\right\} = 0 \\ V(t,X_t) = \frac{(X_t^u)^\gamma}{\gamma}, \quad \forall (t,X_t)\in[0,T]\times\mathbb{R}. \end{cases} \tag{1.6}$$

1.2 Numerical Scheme for the HJB Equation

The PDE (1.6) can be written as,

$$\begin{cases} \frac{\partial V}{\partial t} + \max_{u_t\in\mathcal{A}}\{\mathcal{L}^u V\} = 0 \\ V(t,X_t) = \frac{(X_t^u)^\gamma}{\gamma}, \quad \forall (t,X_t)\in[0,T]\times\mathbb{R}. \end{cases} \tag{1.7}$$

Here,

$$\begin{cases} \mathcal{L}^u V = a(X_t,u_t)\frac{\partial^2 V}{(\partial X_t)^2} + b(X_t,u_t)\frac{\partial V}{\partial X_t} \\ a(X_t,u_t) = \frac{1}{2}X_t^2\sigma^2u_t^2 \\ b(X_t,u) = [r+u_t(\mu-r)]X_t. \end{cases} \tag{1.8}$$

$\mathcal{L}$ is the infinitesimal generator of the process. The domain of the PDE is $\Omega = \{(t,X_t)\in[0,T]\times\mathbb{R}\}$.

1.3 Boundary Conditions

In order to evaluate the HJB PDE numerically, boundary conditions need to be established. At the boundary $X_t = 0$, the PDE reduces to

$$\frac{\partial V}{\partial t} = 0 \tag{1.9}$$

Furthermore, for large X_t, we make the assumption that the optimal control is $u = 0$. The intuition for this assumption is that for a sufficiently large wealth, the marginal gain an investor generates from investing in the risky asset is small compared to investing in the risk-free asset. The PDE (1.6) reduces to

$$\frac{\partial V}{\partial t} + rX_t \frac{\partial V}{\partial X_t} = 0. \tag{1.10}$$

The terminal condition is given by

$$V(t, X_t) = \frac{(X_t^u)^\gamma}{\gamma}, \quad \forall (t, X_t) \in [0, T] \times \mathbb{R}. \tag{1.11}$$

Define $\tau = T - t$, the solution (See Appendix A) to the PDE above is given by,

$$V(T - \tau, X_t) = \delta^2(\tau) X_t^2 + \delta(\tau) X_t, \tag{1.12}$$

where,

$$\delta(\tau) = \exp(2r\tau) \tag{1.13}$$

1.4 Finite Difference

The first step is to localize the PDE to a finite interval $[0, X_{max}]$. Suppose that the price process can be discretized into M nodes, and the time process can be discretized into N nodes. Define $\Delta x = \frac{X_{max}}{M-1}$ and $\Delta \tau = \frac{T}{N-1}$,

$$\begin{aligned} x_i &= (i-1)\Delta x, \quad i = 1, 2, \ldots, M \\ \tau_n &= (n-1)\Delta \tau, \quad n = 1, 2, \ldots, N. \end{aligned} \tag{1.14}$$

The values at each grid point is denotes as $V_i^n = V(\tau_n, x_i)$. The infinitisimal generator $\mathcal{L}^u V$ can be approximated as,

$$(\mathcal{L}_{\Delta x}^u V)_i = \alpha_i V_{i-1} + \beta_i V_{i+1} - (\alpha_i + \beta_i) V_i. \tag{1.15}$$

The PDE (1.6) can be approximated using an explicit timestepping as,

$$\begin{aligned} &\frac{V_i^{n+1} - V_i^n}{\Delta t} + \max_{u_t \in \mathbb{R}} [(\mathcal{L}_{\Delta x}^u V^n)_i] = 0 \\ &\frac{V_i^{n+1} - V_i^n}{\Delta t} + \alpha_i^n V_{i-1}^n + \beta_i^n V_{i+1}^n - (\alpha_i^n + \beta_i^n) V_i^n = 0 \end{aligned} \tag{1.16}$$

From (1.8), note that $a(X_t, u)$ is always positive, while $b(X_t, u)$ is not always positive. When either α_i or β_i is negative, oscillations may appear in the numerical solution. The numerical scheme presented involves the maximal use of central differences to improve the convergence rate and accuracy. In cases that either α_i or β_i is negative, we adopt forward or backward differences to ensure that the new coefficient is positive. The central difference is given by,

$$\begin{aligned}\alpha^n_{i,central} &= \frac{a(x_i, u^n)}{\Delta x^2} - \frac{b(x_i, u^n)}{2\Delta x} \\ \beta^n_{i,central} &= \frac{a(x_i, u^n)}{\Delta x^2} + \frac{b(x_i, u^n)}{2\Delta x}\end{aligned} \tag{1.17}$$

If $\alpha^n_{i,central} < 0$, we apply a forward difference to $\frac{\partial V}{\partial X_t}$,

$$\begin{aligned}\alpha^n_{i,forward} &= \frac{a(x_i, u^n)}{\Delta x^2} \\ \beta^n_{i,forward} &= \frac{a(x_i, u^n)}{\Delta x^2} + \frac{b(x_i, u^n)}{\Delta x}.\end{aligned} \tag{1.18}$$

If $\beta^n_{i,central} < 0$, we apply a backward difference to $\frac{\partial V}{\partial X_t}$,

$$\begin{aligned}\alpha^n_{i,backward} &= \frac{a(x_i, u^n)}{\Delta x^2} - \frac{b(x_i, u^n)}{\Delta x} \\ \beta^n_{i,backward} &= \frac{a(x_i, u^n)}{\Delta x^2}\end{aligned} \tag{1.19}$$

More detail is provided in Appendix B.

1.5 Numerical Algorithm

Define the matrix,

$$A^n\left(u^n\right) = \begin{pmatrix} -\left(\alpha^n_1 + \beta^n_1\right) & \beta^n_1 & & \\ \alpha^n_2 & -\left(\alpha^n_2 + \beta^n_2\right) & \beta^n_2 & \\ & \ddots & \ddots & \ddots \\ & \alpha^n_{M-2} & -\left(\alpha^n_{M-2} + \beta^n_{M-2}\right) & \beta^n_{M-2} \\ & & \alpha^n_{M-1} & -\left(\alpha^n_{M-1} + \beta^n_{M-1}\right) \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Then,

$$\begin{aligned}(\mathcal{L}^u_{\Delta x} V)_i &= \alpha_i V_{i-1} + \beta_i V_{i+1} - (\alpha_i + \beta_i) V_i \\ &= \left[A^n\left(u^n\right) V^n\right]_i.\end{aligned} \tag{1.20}$$

Define the boundary condition vector $G^n = [0, 0, \ldots, 0, G^n_M]$, where the entries are populated using the equation for the boundary condition for large X_t,

$$\begin{aligned}V(T-\tau, X_{max}) &= \delta(\tau) X^2_{max} + \delta(\tau) X_{max} \\ \rightarrow V(t, X_{max}) &= \delta(T-t) X^2_{max} + \delta(T-t) X_{max}\end{aligned} \tag{1.21}$$

such that

$$
\begin{aligned}
G_u^M &:= V(n\Delta t, X_{max}) \\
&= \delta(T - n\Delta t)X_{max}^2 + \delta(T - n\Delta t)X_{max}
\end{aligned}
\tag{1.22}
$$

Then the discretized PDE (1.16) can be written in terms of the control vector u^n

$$
\begin{aligned}
&\frac{V_i^{n+1} - V_i^n}{\Delta t} + max_{u_t \in \mathbb{R}}[\mathcal{L}^u V^n]_i = 0 \\
&\frac{V_i^{n+1} - V_i^n}{\Delta t} + [A^n\left(u^n\right)V^n]_i = 0 \\
&V_i^{n+1} - V_i^n + A^n\left(u^n\right)V^n\Delta t = 0 \\
&[I - A^n(u^n)\Delta t]V^n = V^{n+1} + (G^{n+1} - G^n)
\end{aligned}
\tag{1.23}
$$

Here, $(G^{n+1} - G^n)$ enforces the boundary condition at X_{max}.

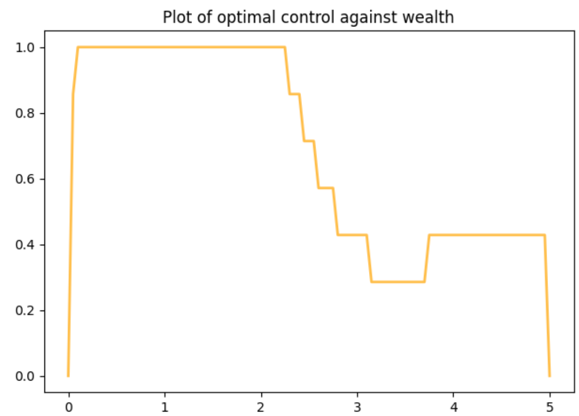
1.6 Simulation

The scheme described above is implemented in Python with the following parameter values:

- r = 0.03
- $\mu = 0.02$
- $\sigma = 0.15$
- T = 20.0
- $\gamma = 0.5$
- $X_{max} = 5.0$
- M = 100
- N = 1600
- $U_{max} = 1$

A plot of the value function $V(X_t, t)$ versus the wealth process X_t and a plot of the optimal control u against wealth X_t derived from the numerical scheme above is presented below:


Plot of $V(X = x_0, t = 0)$ against wealth
0
2
4
6
8
0
1
2
3
4
5


Plot of optimal control against wealth
0.0
0.2
0.4
0.6
0.8
1.0
0
1
2
3
4
5

1.7 Solution to the PDE (Appendix A)

Assume that the solution to the PDE is given by,

$$V(T-\tau, X_t) = \delta(\tau)X_t^2 + \delta(\tau)X_t \tag{1.24}$$

where

$$\delta(\tau) = \exp(2r\tau). \tag{1.25}$$

It follows that

$$\begin{aligned}\frac{\partial V}{\partial t} &= \frac{\partial V}{\partial \tau}\frac{\partial \tau}{\partial t} \\ &= -\frac{\partial V}{\partial \tau} \\ &= -\delta'(\tau)X_t^2 - \delta'(\tau)X_t \\ &= -2r\exp(2r\tau)X_t^2 - 2r\exp(2r\tau)X_t \\ &= -2r\exp(2r\tau)(X_t^2 + X_t).\end{aligned} \tag{1.26}$$

Also,

$$\begin{aligned}rX_t\frac{\partial V}{\partial X_t} &= rX_t(2X_t\delta(\tau) + \delta(\tau)) \\ &= rX_t(2X_t\exp(2r\tau) + \exp(2r\tau)) \\ &= 2r\exp(2r\tau)(X_t^2 + X_t).\end{aligned} \tag{1.27}$$

Hence,

$$\begin{aligned}\frac{\partial V}{\partial t} + rX_t\frac{\partial V}{\partial X_t} &= -2r\exp(2r\tau)(X_t^2 + X_t) + 2r\exp(2r\tau)(X_t^2 + X_t) \\ &= 0.\end{aligned} \tag{1.28}$$

Therefore, $V(t, X_t) = \delta(\tau)X_t^2 + \delta(\tau)X_t$ is a solution to the PDE.

1.8 Discretization Scheme (Appendix B)

We discretize the operator $\mathcal{L}^u V$ and derive the coefficients α_i^{n+1} and β_i^{n+1},

$$\mathcal{L}^u V = a(X_t, u_t)\frac{\partial^2 V}{(\partial X_t)^2} + b(X_t, u_t)\frac{\partial V}{\partial X_t} a(X_t, u_t) = \frac{1}{2}X_t^2\sigma^2 u_t^2. \tag{1.29}$$

Suppose we apply central differencing to $\frac{\partial V}{\partial X_t}$ and $\frac{\partial^2 V}{(\partial X_t)^2}$, then

$$\begin{aligned} \frac{\partial V}{\partial X_t} &\approx \frac{V_{i-1} + V_{i+1}}{\Delta x} \\ \frac{\partial^2 V}{(\partial X_t)^2} &\approx \frac{V_{i-1} + V_{i+1} - 2V_i}{(\Delta x)^2} \end{aligned} \tag{1.30}$$

Then, (1.29) results in,

$$\mathcal{L}^u V \approx a(X_t, u_t)\frac{V_{i-1} + V_{i+1} - 2V_i}{(\Delta x)^2} + b(X_t, u_t)\frac{V_{i-1} + V_{i+1}}{\Delta x} \tag{1.31}$$

Matching the coefficient with (1.15) results in:

$$\alpha_i V_{i-1} + \beta_i V_{i+1} - (\alpha_i + \beta_i)V_i = a(X_t, u_t)\frac{V_{i-1} + V_{i+1} - 2V_i}{(\Delta x)^2} + b(X_t, u_t)\frac{V_{i-1} + V_{i+1}}{\Delta x} \tag{1.32}$$

For the above equation to hold,

$$\begin{aligned} \alpha_{i,central}^n &= \frac{a(x_i, u^n)}{\Delta x^2} - \frac{b(x_i, u^n)}{2\Delta x} \\ \beta_{i,central}^n &= \frac{a(x_i, u^n)}{\Delta x^2} + \frac{b(x_i, u^n)}{2\Delta x} \end{aligned} \tag{1.33}$$

Suppose instead that $\alpha_{i,central}^n < 0$ and we apply forward difference to $\frac{\partial V}{\partial X_t}$, and central difference to $\frac{\partial^2 V}{(\partial X_t)^2}$, then

$$\begin{aligned} \frac{\partial V}{\partial X_t} &\approx \frac{V_{i+1} - V_i}{\Delta x} \\ \frac{\partial^2 V}{(\partial X_t)^2} &\approx \frac{V_{i+1} + V_{i-1} - 2V_i}{(\Delta x)^2} \end{aligned} \tag{1.34}$$

Matching the coefficient with (1.15) results in:

$$\alpha_i V_{i-1} + \beta_i V_{i+1} - (\alpha_i + \beta_i)V_i = a(X_t, u_t)\frac{V_{i+1} + V_{i-1} - 2V_i}{(\Delta x)^2} + b(X_t, u_t)\frac{\partial V}{\partial X_t}\frac{V_{i+1} - V_i}{\Delta x} \tag{1.35}$$

For the above equation to hold,

$$\begin{aligned}
\alpha_{i,forward}^n &= \frac{a(x_i, u^n)}{\Delta x^2} \\
\beta_{i,forward}^n &= \frac{a(x_i, u^n)}{\Delta x^2} + \frac{b(x_i, u^n)}{\Delta x}.
\end{aligned} \tag{1.36}$$

Suppose instead that $\beta_{i,central}^n < 0$ and we apply backward difference to $\frac{\partial V}{\partial X_t}$, and central difference to $\frac{\partial^2 V}{(\partial X_t)^2}$, then

$$\begin{aligned}
\frac{\partial V}{\partial X_t} &\approx \frac{V_i - V_{i-1}}{\Delta x} \\
\frac{\partial^2 V}{(\partial X_t)^2} &\approx \frac{V_{i+1} + V_{i-1} - 2V_i}{(\Delta x)^2}
\end{aligned} \tag{1.37}$$

Matching the coefficient with (1.15) results in:

$$\alpha_i V_{i-1} + \beta_i V_{i+1} - (\alpha_i + \beta_i) V_i = a(X_t, u_t) \frac{V_{i+1} + V_{i-1} - 2V_i}{(\Delta x)^2} + b(X_t, u_t) \frac{V_i - V_{i-1}}{\Delta x} \tag{1.38}$$

For the above equation to hold,

$$\begin{aligned}
\alpha_{i,backward}^n &= \frac{a(x_i, u^n)}{\Delta x^2} - \frac{b(x_i, u^n)}{\Delta x} \\
\beta_{i,backward}^n &= \frac{a(x_i, u^n)}{\Delta x^2}
\end{aligned} \tag{1.39}$$