University Of Waterloo

ACTSC 973

Portfolio Optimization

Report on Solving Multi-period Mean Variance Model through Hamilton-Jacobi-Bellman Equations & Monte Carlo Simulation

---

Name:

Chao Qian, Yifei Deng

12 August, 2020

UNIVERSITY OF WATERLOO

Contents

1 Abstract

The report is involved with solving a continuous-time mean-variance asset allocation problem, which can define as a bi-criteria optimization problem. The objective is to minimize the variance of the terminal wealth while maximizing the expected terminal wealth. By imposing weights on both criteria, one can obtain an optimization problem subject to a stochastic process. However, due to the presence of variance terms in the objective, the problem is not in a standard form for dynamic programming. It is shown that in [6] [11], also explicitly explained in the report, to resolve this problem, the original non-standard problem can be embedded into a class of auxiliary stochastic linear-quadratic (LQ) Problem. In this report, we implement a finite difference method suggested in [8] to solve the resulting nonlinear Hamilton-Jacobi-Bellman (HJB) PDE. The numerical solutions of the problem are presented in the report in terms of an efficient frontier and an optimal control strategy under different realistic constraints imposed on the optimal iterative policy algorithm. Later in this report, to deal with the curse of dimensionality when implementing the numerical method to solve the HJB equations, we introduce a simulated-based approach for solving this constraint dynamic mean-variance optimization problem using Monte-Carlo simulation. We define and discuss two simulation methods in general: a multi-stage strategy implements in a forward fashion and backward recursive programming that is an extension from the multi-stage with a guaranteed convergent solution. The numerical solutions are presented in terms of the efficient frontier with detailed analysis. The main reference paper for this project was [8] written by Wang and Peter in 2010, and the supporting reference was [2] written by Cong and Cornelis in 2016.

2 Introduction

Portfolio selection is a process to find the best allocation of wealth among the collection of selected securities that provides an optimal trade-off between expected returns and risk to an investor. Ever since the first publish of Markowitz's pioneering work [7], the mean-variance portfolio optimization problem has been widely used in the financial industry. It has become the foundation for portfolio selection in single-period investment. As an extension to Markowitz's work, multi-period mean-variance asset allocation problem have also been well studied over the years [6] [11] [4] [2]. Many financial applications are using the continuous-mean variance model including pension asset allocation [4] (later revealed in the report), insurance [10], and even for hedging derivatives [3]. Among all the formulations under different realistic circumstances, the simplest case is by considering two assets model with one risky and one risk-free asset. From such a setting, we can formulate a single objective stochastic control problem, where for each asset, we alter the proportion of wealth dynamically to find the mean-variance efficient one.

However, the original continuous mean-variance problem under such a setting does not meet the standard form of a dynamic programming formulation due to the non-linearity of variance. To resolve this problem, according to [8], one can embed the original non-standard problem into a class of auxiliary stochastic linear-quadratic (LQ) problem. The problem can then be solved through dynamic programming. In the implementation of the LQ method based on [8], one can implement the finite difference method to solve the Hamilton-Jacobi-Bellman (HJB) PDE iteratively and use analytic techniques to derive analytic solution under specific conditions.

Using a finite difference method to solve for HJB equations will give us accurate results of the mean-variance strategy. However, when it comes to the problem with multiple risky assets, a higher-dimension problem, it is rather computational expensive to implement this algorithm. Due to this curse of dimensionality, the authors in [2] proposed a better method of applying the Monte-Carlo Simulation under the same transformed problem using the stochastic LQ method, but a forward fashion multi-stage strategy and an improved backward recursive strategy based from the former policy via Monte-Carlo simulation. Throughout this report, we will discuss and develop both methods: the method based on the numerical solution of HJB equations and the implementation of the Monte-Carlo simulation under different scenarios.

A worthy-note fact is that the final mean-variance strategy using methods from [8] [11] is called the pre-commitment policy. When the initial scheme is determined as a function of wealth at the initial time, the investor follows the pre-commitment tactic and will commit to this strategy regardless of the control computed at later times. In contrast with the time-consistent policy where the investor changes his/her optimal strategy at each instant in time in intermediate steps. In this report, however, we mainly focus on the pre-commitment strategy on [8] while focus on the time-consistent policy on [2].

3 General Setting

This section provides the general settings for the dynamic portfolio optimization problem under a well-defined pension allocation plan for both papers written by Wang and Forsyth (2010) [8] and Cong and Cornelis (2016) [2].

Let Ω be the set of all possible realizations in the financial market within the time horizon [0, T].

Let $\mathcal{F}$ be the sigma algebra of events at time T i.e. $\mathcal{F} = \mathcal{F}_T$

Let $\mathcal{P}$ be the probability measure define on $\mathbb{F}$

Let $\mathbb{D}$:= the set of all admissible wealth w(t), for 0≤ t ≤ T, and $\mathbb{P}$:= the set of all admissible controls p(t,w), for 0≤ t ≤ T and w ∈ $\mathbb{D}$

We assume that the financial market is defined on the probability space (Ω, $\{\mathbb{F}_t\}_{0\leq t\leq T}$, $\mathcal{P}$), where filtration $\{\mathcal{F}_t\}_{0\leq t\leq T}$ is generated by the price process of the financial market and $p_t \in \mathcal{P}$ is an $\mathcal{F}_t$ measurable control. Now, we have the general setting of the problem defined for both papers. We will begin our discussion for the first paper.

4 Wang and Forsyth (2010)

4.1 Mean variance efficient wealth case

We first consider the problem in terms of determining the mean-variance efficient strategy for the investor's final wealth. Since our goal is to maximize the terminal wealth, we refer to this as the wealth case. For simplicity, according to the paper[8], we consider the market with two assets: one risky asset and one risk-free asset. The dynamic of the portfolio's wealth is given by,

$$\begin{aligned} dW_t &= \frac{pW_t}{S_t} dS_t + \pi dt + (1-p)W_t r\, dt & (1) \\ &= \left[(r + p\xi_1\sigma_1)W_t + \pi\right] dt + pW_t\sigma_1 dZ_t & (2) \end{aligned}$$

where, for unit increment of wealth around time t is contributing by investing **p** proportion of wealth in the risky asset per unit time plus constant contribution rate $\boldsymbol{\pi}$ paid by the pension plan members per unit time plus investing **(1-p)** proportion of wealth in the risk-free asset per unit time. With careful re-arrangement, we will have the resultant dynamic for our portfolio of wealth as equation (2), where

1. σ_1 is the volatility.
2. $\xi_1 = \frac{\mu_s - R_f}{\sigma_1}$ is the market price of risk, $r = R_f$ is the risk-free rate
3. The risky asset follows dynamic
$$dS_t = (r + \xi_1\sigma_1)S_t dt + \sigma_1 S_t dZ_t$$
Z_t is the increment of the standard BM, ξ_1 is the market price of risk.

Let $W_T = W(t = T)$, given an expected wealth at terminal time T (i.e. $\mathrm{E}^{t=0}[W_T]$), an investor wishes to minimized his/her risk, measured by variance of terminal wealth $\mathrm{Var}^{t=0}[W_T]$, and thereby the objective function becomes the stochastic control problem to determine the control p(t, W(t) = w), such that p(t,w) maximize,

$$\max_{p(t,w)\in\mathbb{P}} \{\mathrm{E}^{t=0}[W_T] - \lambda \mathrm{Var}^{t=0}[W_T]\} \tag{3}$$

subject to (2), a stochastic dynamic generated by the portfolio of wealth. Here, λ (> 0) measures the level of the risk aversion and W(t) is the path of wealth with a control p(t,w) at time t. Note that p ∈ $(-\infty, \infty)$ in general, an negative p value means shorting the risky asset and a positive p means shorting the risk-less assets to borrow.

According to [11], since the presence of the variance term in problem (3) cause difficulty when using dynamic programming to determine the mean-variance efficient strategy, one can avoid this problem with the help of the following theorem,

Theorem 4.1. *if $p^*(t,w) \in$ the optimal control of problem (3), then $p^*(t,w)$ is also an optimal control of*

$$\max_{p(t,w)\in\mathbb{P}} \{E^{t=0}[\mu W_T - \lambda W_T^2]\} \tag{4}$$

Where

$$\mu = 1 + 2\lambda E_{p^*}^{t=0}[W_T] \tag{5}$$

Proof. see Appendix A.3 □

After we proved the equivalence of (3) and (4), to find control p(t,w) that maximize this new objective function (4), one has to reduce the objective function to the stochastic linear-quadratic problem and utilize the results in dynamic-programming. To achieve this, we first let $\gamma = \frac{\mu}{\lambda}$, then from equation (5),

$$\gamma = \frac{1}{\lambda} + 2\mathrm{E}_{p^*}^{t=0}[W_T] \tag{6}$$

Observe that, for fixed γ, with $\lambda > 0$, equation (5) is equivalent to (See Appendix A.2)

$$\inf_{p\in\mathbb{P}} E^{t=0}\left[(W_T - \frac{\gamma}{2})^2\right] \tag{7}$$

Based on the theoretical results in [11], we embedded the improved objective into a stochastic LQ problem by defining

$$V(w,t) = \inf_{p\in\mathbb{P}} E^{t=0}\left[(W_T - \frac{\gamma}{2})^2 \middle| W_t = w\right] \tag{8}$$

$$= \inf_{p\in\mathbb{P}} J(t,w,p) \tag{9}$$

where J(t, w, p) is called the associated cost function in a dynamic programming setting. Then, by Ito's Lemma and using the dynamics W_t previously derived, we have,

$$dV = V_t + V_w dW_t + \frac{1}{2}V_{ww}d[W,W]_t \tag{10}$$

$$= V_t dt + V_w\left[(r + p\xi_1\sigma_1)w + \pi\right]dt + pw\sigma_1 dZ_t + \frac{1}{2}(pw\sigma_1)^2 V_{ww}dt \tag{11}$$

$$= \left(V_t + V_w \underbrace{[(r + p\xi_1\sigma_1)w + \pi]}_{\mu_w^p} + \frac{1}{2}\underbrace{(pw\sigma_1)^2}_{(\sigma_w^p)^2} V_{ww}\right)dt + \underbrace{pw\sigma_1 dZ_t}_{\text{local martingale}} \tag{12}$$

Since V is a martingale, then no dt term, we know that V satisfies the HJB equation (See Appendix A.1),

$$V_t + \inf_{p\in\mathbb{P}} \underbrace{\left(\mu_w^p V_w + \frac{1}{2}(\sigma_w^p)^2 V_{ww}\right)}_{L^pV} = 0 \tag{13}$$

$$\Rightarrow V_t + \inf_{p\in\mathbb{P}} L^pV = 0 \tag{14}$$

with terminal condition,

$$V(w, t = T) = (w - \frac{\gamma}{2})^2 \tag{15}$$

To further trace out the efficient frontier of the original problem (3), we can proceed by picking an arbitrary γ so that the optimal $p^*(t,w)$ is determined by solving (14).

However, given $p^*(t,w)$, we still need to know $\mathrm{E}_{p^*}^{t=0}[W_T]$ to draw the efficient frontier, let U = U(w,t) = $\mathrm{E}[W_T \mid W(t) = w, p(t,w) = p^*(t,w)]$. Then, applies Ito's lemma to U, so U is given from the solution to

$$U_t = -\left(\mu_w^p U_w + \frac{1}{2}(\sigma_w^p)^2 U_{ww}\right)_{W(t)=w,p(t,w)=p^*(t,w)} \tag{16}$$

with terminal condition,

$$U(w, t = T) = w \tag{17}$$

Assume that W = $\hat{w}_0$ at t = 0, i.e. known initial wealth for the pension plan member, and assuming that V($\hat{w}_0$, t = 0) and U($\hat{w}_0$, t = 0) are known, then

$$V(\hat{w}_0, t = 0) = \mathrm{E}_{p^*}^{t=0}[W_T] - \gamma \mathrm{E}_{p^*}^{t=0}[W_T] + \frac{\gamma^2}{4} \tag{18}$$
$$U(\hat{w}_0, t = 0) = \mathrm{E}_{p^*}^{t=0}[W_T] \tag{19}$$

For each given γ, where $\gamma = \frac{1}{\lambda} + 2\mathrm{E}_{p^*}^{t=0}[W_T]$, we can compute pair ($\mathrm{Std}_{p^*}^{t=0}[W_T]$, $\mathrm{E}_{p^*}^{t=0}[W_T]$) for the optimal control $p^*(t,w)$ which solves (3) and λ from (6). Thus, we can vary the parameter γ to effectively trace out the efficient frontier. Since $\lambda > 0$, from (6), we must have,

$$\gamma = \frac{1}{\lambda} + 2\mathrm{E}_{p^*}^{t=0}[W_T] \tag{20}$$
$$\Rightarrow \frac{1}{2\lambda} = \frac{\gamma}{2} - \mathrm{E}_{p^*}^{t=0}[W_T] > 0 \tag{21}$$

for a valid point on the efficient frontier. Now, we have finished the theoretical part to determine the mean-variance efficient strategy for the investor's terminal wealth. In the next section, we will consider the boundary conditions for large w and the constraints for control and asset allocation in realistic settings.

4.1.1 Localization

In this section, to numerically solve the HJB equation derived from the previous section, we need to localize our computational domain and discuss the domain into different cases. Let

$$\hat{\mathbb{D}} := \text{a finite computational domain which approximates the set } \mathbb{D} \tag{22}$$

To start, we have to figure out the general solutions of V and U under boundary conditions for large $|w|$. We use a finite computational domain $\hat{\mathbb{D}} = [w_{min}, w_{max}]$. When w $\to \pm\infty$, based from (15) & (17) we assume that

$$V(w \to \pm\infty, \tau = T - t) \approx H_1(\tau = T - t)w^2 + H_2(\tau = T - t)w + H_2(\tau = T - t) \tag{23}$$
$$U(w \to \pm\infty, \tau = T - t) \approx J_1(\tau = T - t)w + J_2(\tau = T - t) \tag{24}$$

Let $\tau = T - t$, we perform change of variable for equations (14) & (16) and consider the initial condition determined in (15) & (17), we can solve for the approximated general solution of U and V, (see Appendix A.4)

$$V(\tau) \approx e^{(2k_1+k_2)\tau} w^2 \tag{25}$$
$$U(\tau) \approx e^{k_1\tau} w \tag{26}$$

Where $k_1 = r + p\sigma_1\xi_1$ and $k_2 = (p\sigma_1)^2$ Now, according to [8], we only consider the following three cases for boundary constraints.

4.1.2 Boundary constraints

From the aspects of real-life applications, imposing constraints on asset allocations is crucial. Consider when an investor has large enough wealth, he/she prefers to choose the risk-less assets, as the marginal gain from investing in the risky asset is relatively small compare to invest in the risk-less asset. Thus, in this section, to meets such practical circumstances, we plan to impose boundary conditions on asset allocations and wealth.

Allowing Bankruptcy, unbounded controls In this case, we assume there are no constraint on W(t) or on the control p(t,w) i.e. $\mathbb{D} = (-\infty, +\infty)$ and $\mathbb{P} = (-\infty, +\infty)$.

1. Since $W(t) < 0$ is allowed, then we called this case allowing bankruptcy.
2. Numerically, we use $\hat{\mathbb{D}} = [w_{min}, w_{max}]$ (like Dirichlet conditions) to approximate $\mathbb{D} = (-\infty, +\infty)$. Noted that by using the artificial boundary condition (i.e. truncated the domain of interest) will introduce some numerical error. Thus, we need to choose sufficiently large $|w_{min}|$ and w_{max} to avoid this numerical error.

No Bankruptcy, No short sales In this case, we assume bankruptcy is prohibited (i.e. W(t)≥0) and the investor cannot short sell the risky assets (i.e. p(t,w)≥0) i.e. $\mathbb{D} = [0, +\infty)$ and $\mathbb{P} = [0, +\infty)$.

1. Since $W(t) \geq 0$ & p(t,w)≥0, then we called this case no bankruptcy and no short sales.
2. Numerically, we use $\hat{\mathbb{D}} = [0, w_{max}]$ to approximate $\mathbb{D} = [0, +\infty)$.
3. The author make assumption that $p^*(t, w_{max}) \approx 0$, as investor has large enough wealth, he/she prefers the risk-less asset
4. The boundary conditions of V and U at w_{max} are given by (25)&(26) with p = 0 and w_{max}. Whereas, at $w_{min} = 0$, to prohibit bankruptcy, we have $\lim_{w\to 0}(pw) = 0$ so that equations (14) & (16) reduce to

$$V_\tau(w_{min} = 0, \tau) = \pi V_w$$
$$U_\tau(w_{min} = 0, \tau) = \pi U_w \tag{27}$$

No Bankruptcy, bounded control In this case, we assume that bankruptcy is prohibited (i.e. W(t)≥0) and infinitely borrowing is not allowed (i.e. p(t,w) = p_{max}) i.e. $\mathbb{D} = [0, +\infty)$ and $\mathbb{P} = [0, p_{max})$.

1. Numerically, we use $\hat{\mathbb{D}} = [0, w_{max}]$ to approximate $\mathbb{D} = [0, +\infty)$
2. Note that w_{max} is an approximation of infinity boundary condition for sufficiently large w_{max} and the boundary condition of V and U at w_{max} are given by (25),(26) with p = p_{max} and w_{max}
3. Other assumptions and the boundary conditions for V and U are the sames as those of no bankruptcy case

In summary, we have Table 1.

Table 1
Summary of cases.

Case	$\hat{\mathbb{D}}$	$\mathbb{P}$
Bankruptcy	$[w_{\min}, w_{\max}]$	$(-\infty, +\infty)$
No bankruptcy	$[0, w_{\max}]$	$[0, +\infty)$
Bounded control	$[0, w_{\max}]$	$[0, p_{\max}]$

Analytic solution for unconstrained control

According to [8], this is the case of allowing bankruptcy from previous section. By using analytic techniques, we can derive an analytic solution to this case, where the solution is given by [4],

$$\begin{cases} \mathrm{Var}^{t=0}[W_T] = \frac{e^{\xi_1^2 T} - 1}{4\lambda^2} \\ \mathrm{E}^{t=0}[W_T] = \hat{w}_0 e^{rT} + \pi \frac{e^{rT}-1}{r} + \sqrt{e^{\xi_1^2 T} - 1} Std(W_T) \end{cases} \tag{28}$$

and the optimal control at each time t ∈ [0,T] is (See Appendix A.5)

$$p^*(t, w) = -\frac{\xi_1}{\sigma_1 w}[w - (\hat{w}_0 e^{rt} + \frac{\pi}{r}(e^{rt} - 1)) - \frac{e^{-r(T-t)+\xi_1^2 T}}{2\lambda}] \tag{29}$$

We are observed that the larger the weight is given to the minimization of variance, the lower the amount invested in the risky assets. Now, we have done the localization needed for various cases for the boundary conditions. In the later section, we will discretize and solve the HJB equations dynamically with different discretize methods.

4.2 Discretization of the HJB PDE

In this section, we discretize the HJB PDE using fully implicit time-stepping and "maximal use of central difference" suggested by [9]. Firstly, we need to discretize both the wealth and time step with uniform grid,

- For time horizon, we divides [0, T] into N equally sized sub-interval with step size Δt, where n = 0,. . .,N i.e. N = the number of time steps
- For wealth, we divides $[w_{min}, w_{max}]$ into M equally sized sub-interval with step size Δw, where i = 0,. . .,M i.e. M = the number of wealth steps
- For V(w_i, t^n), V_i^n is a discrete approximation to V(w_i, t^n); U_i^n is a discrete approximation to U(w_i, t^n)

Base on above discretization on both wealth and time horizon, let $(L_p^h V^n)_i$ be the discrete approximation of the operator L_p. The differential operator L^p can be discretized using forward, backward or central differencing to give the general form as follow,

$$(L_p^h V^n)_i = \alpha_i^n V_{i-1}^n + \beta_i^n V_{i+1}^n - (\alpha_i^n + \beta_i^n) V_i^n \tag{30}$$

Where α_i & β_i are defined in the Appendix A.7 for different finite difference schemes.

Given different schemes in Appendix A.7, we now can dicretize equation (30),

$$\frac{V_i^{n+1} - V_i^n}{\Delta \tau} = \inf_{p^n \in \mathbb{P}_h} \{(L_{p^n}^h V^n)_i\} \tag{31}$$

where we approximate set $\mathbb{P}$ by equally sized space set $\mathbb{P}_h = \{p_1, p_2, \ldots, p_q = p_{max}\}$ with p^n calculate from equation (30), then U can be dicretized as, (recall that $V_\tau = \inf_{p \in \mathbb{P}} \{L^p V\}$ & $U_\tau = \{L^p U\}_{p=p^*}$)

$$\frac{U_i^{n+1} - U_i^n}{\Delta \tau} = \{(L_{p^n}^h U^n)_i\} \tag{32}$$

note that both $\alpha_i^n = \alpha_i^n(p^n)$ and $\beta_i^n = \beta_i^n(p_i^n)$ are functions of the local optimal control p_i^n, which makes equation (31) nonlinear. Besides the discretization scheme developed from above, we also need to be careful about the grid choices under certain cases. (see Appendix A.9)

Positive coefficient conditions From Condition 4.1 in [8], the authors shows that if α_i^n and β_i^n satisfy the following condition,

$$\alpha_i^n \geq 0, \beta_i^n \geq 0, i = 0, \ldots, M \tag{33}$$

the discretization from above is convergent [1]. Therefore, we need to use forward, backward and central discretiations to ensure that the coefficients in (31) and (32) are positive.

According to [9], the authors develop and implement a positive coefficient scheme for HJB PDEs in finance called "maximal use of central difference". The implementation of the method uses central differences (higher order of accuracy shown in Appendix A.6) as much as possible by changing between forwarding, backward and central schemes at each node to preserve the positive coefficient condition. Hence, we will always have the coefficients satisfy the requirement (33), where our algorithm has guaranteed convergent solutions at all times and space.

Policy-iteration Algorithm In this section, we introduce the Policy-iteration algorithm to calculate U and V using the discretization method on HJB PDE. (See Figure 1.)Followed by the matrix transformations in Appendix A.8, we have the matrix form for equation (31) and (32) as below,

$$[I - \Delta\tau A^n(P^n)]\, V_i^n = V_i^{n+1} + (G^{n+1} - G^n)\ \text{, with } p_i^n = \inf_{p \in P_h} \{(L_p^h V^n)_i\} \tag{34}$$

$$[I - \Delta\tau A^n(P^n)]\, U_i^n = U_i^{n+1} + (H^{n+1} - H^n)\ \text{, with each } p_i^n \text{ from (34)} \tag{35}$$

Thus, we can formulate the final policy-iteration based on (34) & (35) according to [8] in Figure 1 as below,

Algorithm 1. Iterative solution of the discrete equations.

Let $(V^{n+1})^0 = V^n$
Let $\hat{V}^k = (V^{n+1})^k$
For $k = 0, 1, 2, \ldots$ until convergence
  Solve
$$[I - \Delta\tau A^{n+1}(P^k)]\hat{V}^{k+1} = V^n + (G^{n+1} - G^n)$$
$$p_i^k = \underset{P \in \hat{P}}{\operatorname{argmin}}\{(A^{n+1}(P)\hat{V}^k)_i\}$$
  If$(k>0)$ and $\left(\max_i \frac{|\hat{V}_i^{k+1} - \hat{V}_i^k|}{\max(scale, |\hat{V}_i^{k+1}|)} < tolerance\right)$ then quit
EndFor

Figure 1: Policy Iteration algorithm

Algorithm for tracing the efficient frontier Plotting the efficient frontier becomes straightfoward after we have the value of U and V. Two things to keep in mind: First, for each γ in the domain, we assemble U and V to get a pair $(Std_{p^*}^{t=0}[Z_T], E_{p^*}^{t=0}[Z_T])$. Second, the efficient frontier is plotted at $Z(t=0) = \hat{Z}_0$ using interpolation between two adjacent nodes. Closely follow the algorithm 2, we could plot the efficient frontiers for each case to do the comparison.

Algorithm 2. Algorithm for constructing the efficient frontier.

For $\gamma = \gamma_{\min}, \gamma_1, \ldots, \gamma_{\max}$
  For timestep $n = 1, \ldots, N$
    Solve Eq. (4.12) by using policy iteration (4.15)
    Solve Eq. (4.14) // P^{n+1} is given from the solution of Eq. (4.12)
  EndFor
  Given the initial $\hat{z}_0$, use interpolation to get the value of
$$\left(E_{p^*}^{t=0}[Z_T], E_{p^*}^{t=0}\left[\left(Z_T - \frac{\gamma}{2}\right)^2\right]\right)_\gamma \text{ at } Z(t=0) = \hat{z}_0$$
  If $\left(\frac{\gamma}{2} - E_{p^*}^{t=0}[Z_T]_\gamma > 0\right)$ // possible valid point $\lambda > 0$
    SolveEq. (2.16) to get $(E_{p^*}^{t=0}[Z_T], E_{p^*}^{t=0}[Z_T^2])_\gamma$
    Calculate the pair $(\mathrm{Std}_{p^*}^{t=0}[Z_T], E_{p^*}^{t=0}[Z_T])_\gamma$
  EndIf
EndFor
Construct the efficient frontier by the points of
$$(\mathrm{Std}_{p^*}^{t=0}[Z_T], E_{p^*}^{t=0}[Z_T])_\gamma, \gamma \in [\gamma_{\min}, \gamma_{\max}]$$

Figure 2: Algorithm for tracing the efficient frontier

5 Cong and Cornelis (2016)

5.1 Problem Reformulation

In this supporting paper [2], the author assumed the same setting as paper [8] including portfolio construction, the dynamic mean-variance problem embedded into the stochastic LQ problem (9), and resultant HJB equation involve with U & V with the same boundary conditions from (14),(15),(16)&(17). The only difference between this paper and the paper in the previous section is that in this case we reformulate the problem such that we traded at discrete opportunities. In particular, at each trading time t $\in [0, \Delta t, \ldots, T - \Delta t]$ before terminal time, an investor desires to maximized the expectation of his or her terminal wealth and to minimized the investment risk. Theoretically, the investor's problem is as the following,

$$\max_{\{p_s\}_{s=t}^{T-\Delta t}} \{\mathrm{E}[W_T|W_t] - \lambda \cdot \mathrm{Var}\,[W_T|W_t]\} \tag{36}$$

subject to the wealth dynamic (2) in discrete form:

$$W_{s+\Delta t} = W_s \cdot (p_s \mathrm{e}^{r_t^e} + (1-p_s)R_f) + \pi \cdot \Delta t = W_s \cdot (p_s R_s^e + R_f) + \pi \cdot \Delta t, \; s = t, t+\Delta t, \ldots, T-\Delta t \tag{37}$$

Here, all parameter are the same as the previous paper [8] except for some parameters. In particular, $R_s^e = \mathrm{e}^{r_t^e} - R_f$ is the excess return of the risky asset and $\mathrm{e}^{r_t^e}$ is the mean return of the risky asset during the period $[s, s+\Delta t]$. We assume that $\{R_t^e\}_{t=0}^{T-\Delta t}$ are independent. In addition, p_s is now denoted as the proportion of wealth investing in the risky asset at period $[s, s+\Delta t]$ and the authors in [2] assume that the admissible p_t is an $\mathcal{F}_t$ measurable control (i.e. we know strategy p_t at time t). For simplicity, the risk-free rate R_f is assumed to be constant and negative π means a constant withdrawal from the the pension plan's wealth portfolio by the plan's member.

Eventually, by Theorem 4.1 and Appendix A.3 again, our objective becomes,

$$V_t(W_t) = J_t(W_t) = \min_{\{p_s\}_{s=t}^{T-\Delta t}} \{\mathrm{E}\left[(W_T - \frac{\gamma}{2})^2 \middle| W_t\right]\}, \text{ where } \gamma = \frac{1}{\lambda} + 2\mathrm{E}_{p^*}[W_T|W_t] \tag{38}$$

However, according to [2], casting constraints on the control in the numerical approach, like authors in paper [8] propose, is not trivial. Specifically, from the computational efficiency aspect, solving the optimally for constraint case at each time step may be extremely hard and computationally expensive. While, from the error aspect, using the value function J to transfer the information between two recursive steps (34) & (35) will introduce accumulated numerical error as the iteration proceeds.

Therefore, the authors in this paper [2] propose two better schemes than the numerical approach proposed by [8] to solve the continuous mean-variance optimization problem. The authors introduce two strategies in general via Monte-Carlo simulation: a multi-stage strategy in a forward fashion and an improved backward recursive dynamic programming calculation based on the solution of the multi-stage strategy.

5.2 The multi-stage strategy in forward fashion

This section describes the formulation of multi-stage strategy by prevention on dealing with the non-smooth value functions J. In this case, we first re-write (40) into the pre-commitment form at time t in an iterative form,

$$J_t^{*pc} := \min_{p_t}\{\mathrm{E}\left[(W_T - \frac{\gamma}{2})^2) \middle| W_t\right]\} \tag{39}$$

$$= \min_{p_t}\{\mathrm{E}\left[J_{t+\Delta t}^{*pc}(W_{t+\Delta t}) \middle| W_t\right]\}, \text{ with } J_t^{*pc}(W_T) = (W_T - \frac{\gamma}{2})^2 \tag{40}$$

where at state (t, W_t), the value function J_t^{*pc} depends solely on optimal p at each successive time paths. Clearly observe that (40) is the algorithm we discuss in the previous paper [8] using the numerical methods to backwardly solve the resultant PDE equations in a recursive form. This pre-commitment algorithm that solve the dynamic programming problem numerically, according to [2][8], are computational inefficient and numerically inaccuracy base on the accumulated error.

Therefore, reflections on these two issues leads the author in the paper [2] to propose a sub-optimal strategy that does not has occurrence of these two issues. We set our target to $\frac{\gamma}{2}$ at the terminal time. Then, state (t, W_t), the sub-optimal strategy is as follow,

$$p_t^{*ms} := \arg\min_{p_t}\{\mathrm{E}\left[W_t \cdot (p_t R_t^e + R_f) + \pi \cdot \Delta t - \delta_{t+\Delta t} \middle| W_t\right]\} \tag{41}$$

where

$$\delta_t = \frac{\frac{\gamma}{2} - \pi \cdot \Delta t \cdot \frac{1-(R_f)^{\frac{T-t}{\Delta t}}}{1-R_f}}{(R_f)^{\frac{T-t}{\Delta t}}} \tag{42}$$

So, in this case, instead of looking at the optimality over a entire constraint in the future, we perform a single stage optimization problem at each instant time step with respect to a given target value. Due to this static optimization algorithm that perform at each stage, we also call this sub-optimal strategy as multi-stage strategy.

The interpretation of δ_t is straightforward according to [2], we simply decide an intermediate target at time t so that as we achieve this target for risky asset, we can deposit all the money into the risk-free asset. Eventually, at the terminal time stage, we get to the final target. Thereby, this δ_t is calculated by discounting the final target via $(R_f)^{\frac{T-t}{\Delta t}}$ term along with the considerations of the contribution rate pay by the pension plan members per unit time Δt. Thus, due to this static optimization feature at each time step, we can generate the steps for this multi-stage algorithm in a forward fashion, (see Appendix B.1)

Under the assumption of $\pi = 0$ i.e. no contribution, we can re-write the pre-commitment problem (42) as:

$$p_t^{*pc}(W_t) = \arg\min_{p_t}\{\mathrm{E}\left[\left(W_t \cdot (p_t R_t^e + R_f) \cdot \prod_{s=t+\Delta t}^{T-\Delta t}(p_s^{*pc} R_s^e + R_f) - \frac{\gamma}{2}\right)^2 \middle| W_t\right]\} \tag{43}$$

Consider the same assumption, we can also obtain the multi-stage optimization as,

$$p_t^{*ms}(W_t) = \arg\min_{p_t}\{\mathrm{E}\left[\left(W_t \cdot (p_t R_t^e + R_f) \cdot \prod_{s=t+\Delta t}^{T-\Delta t}(R_f) - \frac{\gamma}{2}\right)^2 \middle| W_t\right]\} \tag{44}$$

In the sub-optimal strategy, instead of considering the $\{p_s^{*pc}\}_{s=t+\Delta t}^{T-\Delta t}$, we set $\{p_s^{*pc}\}_{s=t+\Delta t}^{T-\Delta t}$ equal to zero, which is the primary reason we call multi-stage approach as sub-optimal strategy. In summary, the multi-stage approach is sub-optimal but highly efficient for solving the mean-variance portfolio optimization problem, which does not generate accumulated error. We can also extend this multi-stage or sup-optimal strategy to high-dimensional problems with complex asset dynamics.

However, the drawbacks of the multi-stage approach is also apparent, as we changed the objective of the problem, the optimal control obtained from this sub-optimal problem can be differ from the pre-commitment one. In fact, the mean-variance strategy pairs using the sub-optimal approach will lies below the optimal efficient frontier by the pre-commitment strategy. However, the authors in [2] argue that under certain constraints (i.e. as $\pi = 0$), the optimal allocation for the multi-stage strategy is exactly same as the pre-commitment policy. The detailed proof and deviations are provided in [2] for further interests.

5.3 The backward recursive programming

This section describes the formulation of a backward recursive strategy that ensures convergence to optimal solutions. One can use the solution from the multi-stage strategy to proceed to this section. In this case, we are given the p_t^{*ms} by multi-stage strategy, an approximation for the real optimal asset allocation p_t^*. If we constraint ourselves in the domain of $A_\eta = [p_t^{*ms} - \eta, p_t^{*ms} + \eta]$ Then the optimal allocation p_t in

$$J_t(W_t) = \min_{p_t \in A_\eta}\{\mathrm{E}\left[J_{t+\Delta t}(W_{t+\Delta t}) | W_t\right]\}$$

should be the same as the one without the truncated domain at state (t,W_t).

To solve this truncated problem, we should first know the value function $J_{t+\Delta t}(W_{t+\Delta t})$ on the domain $D_{t+\Delta t}$. Using the simulation with bundling technique introduced in [5], the domain $D_{t+\Delta t}$ can be approximated as

$$\hat{D}_{t+\Delta t} \coloneqq \{W_{t+\Delta t} | W_{t+\Delta t} = W_t \cdot (p_t^{*ms} \cdot R_t^e + R_f) + \pi \cdot \Delta t,\ p_t \in B_\delta\}$$

We could summarize the backward programming algorithm in to the following four steps and the detailed schematic can be found in Appendix B.2 (we consider letting $x_t = p_t$ to be consistent with the following statements)

1. Start with the optimal allocations,$\{\tilde{x}_t\}_{t=0}^{T-\Delta t}$ given by the multi-stage strategy, simulate the stochastic processes $W_t(i))_i^N, t = 0 \cdots T$, at the terminal time we calculate the value function $J_T(W_T)$(N of them in total, where N is the number of the simulations.) Then perform the following three steps iteratively backward in time for $t = T - \Delta t, \cdots, \Delta t, 0$.

2. At each time step, sort the wealth in decreasing order, bundle the paths(simulations) in to N_B partitions, indexes can be saved for each path in bundle for the convenience of programming. For each path in bundles, we perform

   (a) Using least square regression to find a local function $f^b_{t+\Delta t}(\cdot)$ using up to second order basis functions to regress the continuation values $J^b_{t+\Delta t}(i)^{N_b}_{i=1}$ on the wealth values $W^b_{t+\Delta t}(i)^{N_b}_{i=1}$ at time $t+\Delta t$.

   (b) Calculate the new optimal allocation $\hat{x}^b_t(i)^{N_b}_{i=1}$ by solving the first order condition of the value function $f^b_{t+\Delta t}(W^b_{t+\Delta t})$.

   (c) Compute the new continuation value,$\hat{J}^b_{t+\Delta t}(i)^{N_b}_{i=1}$, conditional on the known values, $W^b_{t+\Delta t}(i)^{N_b}_{i=1}$ and $\hat{x}^b_t(i)^{N_b}_{i=1}$.

3. Calculate the old continuation value $\tilde{J}^B_t(i)$. Update the allocation $x^b_t(i)v^{N_b}_{i=1}$ based on the condition if the new continuation value,$\hat{J}^b_t(i)$, is less than the original one, $\tilde{J}^b_t(i)$.

4. Once the updated allocations are obtained, again by regression we can calculate the "updated" continuation value $J^b_t(i)^{N_B}_{i=1}$and proceed with the backward recursion.

In the algorithm, at each time step and inside each bundle, four regression steps are performed. The last three regression steps are utilized for calculating value functions. Since the value function is used to evolve information between time steps, an error in calculating them will accumulate due to recursion. In general, as suggested by the author, we can ease this problem by using large number of simulations, yielding a satisfiable result, yet computationally expensive.

5.4 Constraints on the asset allocations

In practice, imposing constraints on asset allocation is as important as the algorithm itself. For example, when someone gets bankrupt, realistically, he or she should not have control over the portfolio anymore. Furthermore, according to the banks' regulations, "Basel III" for example, the bounded leverage is required to avoid excessive borrowing. Therefore, this section presents some realistic constraints on asset allocations.

5.4.1 No bankruptcy constraint

In this case, no bankruptcy constraint refers to the zero probability to have wealth below zero. After careful derivation, we have (See Appendix A.10)

$$0 \leq p_t \leq 1 + \frac{\pi \Delta t}{W_t \cdot R_f} \tag{45}$$

A worthy-note point is that no bankrupt constraint (45) implies $\lim_{W_t \to 0}(p_t \cdot W_t) = 0$. The financial interpretation for this constraint is as investor has wealth close to zero, he or she should not invest in the risky asset.

5.4.2 No bankruptcy constraint with 1-2α% certainty

In this case, when we have a large portfolio, the upper bound for constraint (45) in no bankruptcy case will reach to 1, which is quite exhaustive. Since the upper bound in this case that prevent bankruptcy only happens in a rare event that the risky asset generates zero return. According to paper [2], we can consider using the possibility of bankruptcy to relieve this extreme upper bound. Assume the (α) and the $(1-\alpha)$ quantiles for the excess return R^e_t are $R^{e,\alpha}_t$ and $R^{e,1-\alpha}_t$ correspondingly. It follows by paper [2], with certainty $1-2\alpha\%$ and the constraint from below, The bound for the control p_t can be computed as,

$$\frac{-\pi \Delta t - W_t \cdot R_f}{W_t \cdot R^{e,1-\alpha}_t} \leq p_t \leq \frac{-\pi \Delta t - W_t \cdot R_f}{W_t \cdot R^{e,\alpha}_t} \tag{46}$$

5.4.3 Bounded leverage

In this case, we simply impose $[p_{min}, p_{max}]$ boundary on asset allocation to ban an investor from gambling when he or she is almost goes bankrupt.

6 Results of Replication

6.1 Wang and Forsyth (2010)

Before looking at the details, let have a look at how V, U, and optimal control p^* evolve w.r.t. the value of wealth with the parameters given in the following table.

Table 1: Parameters for Numerical results

r	0.03	M	728
σ_1	0.15	N	160
W(t=0)	1	numOfReb	10
T	20 years	γ	14.47
ξ_1	0.1	P_{max}	1.5
π	0.1	P_{min}	0.0

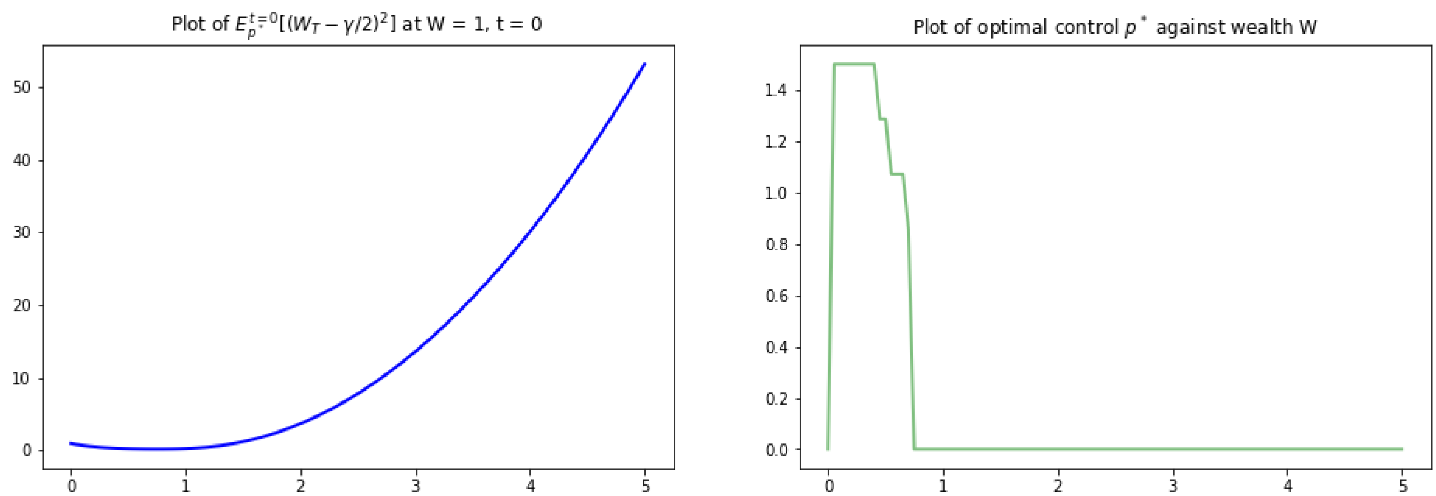

Figure 3: Plot of V, p^* versus wealth ($\gamma = 9.125$)

For this experiment, we chose the number of rebalancing periods as 10 to avoid the computational burden, even though the time steps are 160. As one could observe from the plot that the portion that invests in the risky asset increases first until it reaches the upper bound(1.5 in this case) and then gradually decrease as the wealth accumulate, which verifies the assumption that made in [8] where people would invest all money in the non-risky asset when the wealth is relatively large.

6.1.1 Allowing bankruptcy

On top of the fist trial, we implement two algorithms in the paper [8] and closely follow the process to trace out the efficient frontier, with the generic choice of $\gamma \in [0.1, 12]$, as shown in the first plot of figure 4. In the allow bankruptcy case, the frontier is a straight line as expected, shown in figure 4 second panel. However, in contrast with the bounded control case, two plots have quite different horizons on the y-axis, which is not in alignment with the results in the original paper. One possible explanation for this error is that the rebalancing periods that

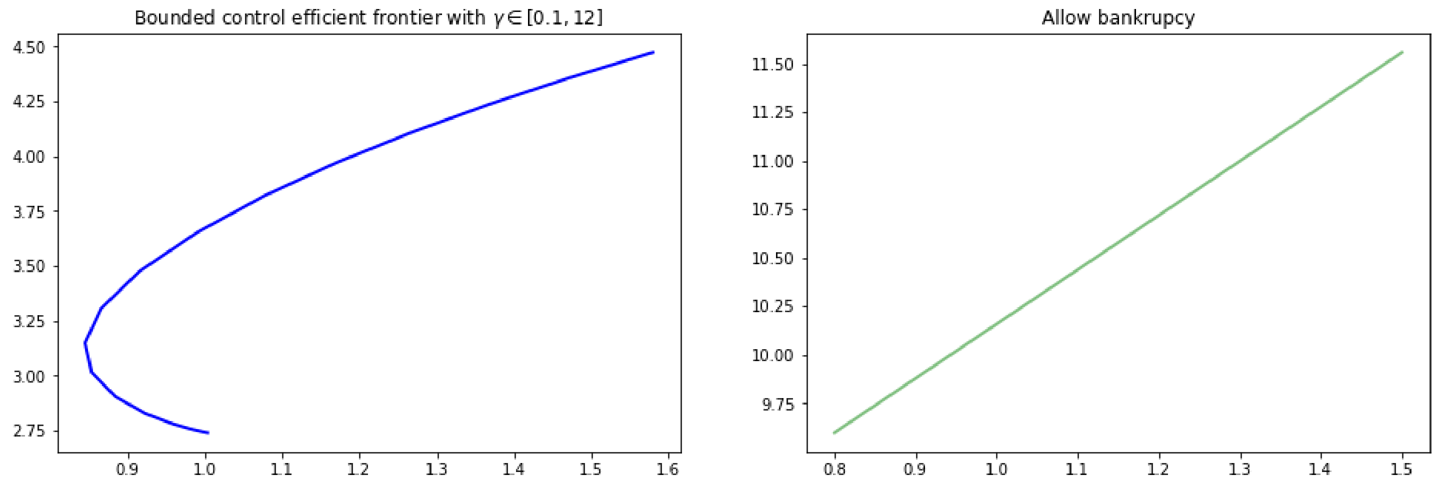

Figure 4: Plot of V, p^* versus wealth

we chose as 10 is far smaller than in the paper as 160, which results in the slightly different efficient frontier in the bounder control case. Indeed, after increasing the number of rebalancing periods throughout the time horizon, we observe the efficient frontier entrenched in the sense that it moves closer to the allow bankruptcy case.

To further testify the impact of number of nodes, number of time steps on the $E_{p^*}^{t=0}[(W_T-\gamma/2)^2]$ and $E_{p^*}^{t=0}[W_T]$, we reproduce the following tables. ($P_{max} = 1.5$ is maximum borrowing of 50% of net wealth)
Table 3,4 are reproduced in the context of $\gamma = 9.125$, and $|Wmin| = |Wmax| = 5925$ to approximate the

Table 2: Parameters used in the pension plan examples

r	0.03	M	1600
σ_1	0.15	N	100
T	20 years	tol	1e-6
ξ_1	0.1	P_{max}	1.5
π	0.1	scale	1.0
W(t=0)	1		

Table 3: Value Function J

Nodes (W)	Timesteps	Nonlinear iterations	CPU time(s)	$E_{p^*}^{t=0}[(W_T-\frac{\gamma}{2})^2]$
728	160	480	0.87	49.4670
1456	320	960	1.74	30.9555
2912	640	1295	4.28	21.6984
5824	1280	2561	9.59	17.06992
11648	5120	5120	21.50	14.7556

$\gamma = 9.125, |Wmin| = |Wmax| = 5925$. At each node, use analytic solution for optimal control. Apply fully implicit time stepping and use constant time steps

infinity wealth boundary. For this part of the convergence study, we do not see that the $\mathrm{Std}_{p^*}^{t=0}[W_T]$ and $\mathrm{E}_{p^*}^{t=0}[W_T]$ converge monotonically to the analytic solution, but rather in a bumpy fashion. The possible explanation for this mismatch with the results in [8] are in two folds: First, the number of non-linear iteration that we chose is 2 which is smaller than suggested in the paper; second, the discretized control we chose is 10. All these two reasons can lead to different control processes, final wealth, and hence, different scale of the efficient frontier.

Similarly, when setting the $\gamma = 14.47, |Wmin| = |Wmax| = 5925$, we get the numerical results in table 5, 6.

Table 4: Convergence Study: Std and Mean

Nodes (W)	Timesteps	Nonlinear iterations	$\text{Std}_{p^*}^{t=0}[W_T]$ [$\star$]	$\text{E}_{p^*}^{t=0}[W_T]$ [$\bigcirc$]	Ratio for [$\star$]	Ratio for [$\bigcirc$]
728	160	480	162.9951	3.4955		
1456	320	960	151.8224	3.0656		
2912	640	1295	274.8677	9.4465	0.5929	0.3700
5824	1280	2561	221.3222	6.2596	0.7364	0.5584
11648	2560	5120	406.4779	20.3609	0.4009	0.1716

$\text{Std}_{p^*}^{t=0}[W_T]$ and $\text{E}_{p^*}^{t=0}[W_T]$ are evaluated at (W = 1, t = 0). Analytic solution is ($\text{Std}_{p^*}^{t=0}[W_T]$, $\text{E}_{p^*}^{t=0}[W_T]$) = (0.0, 4.5625)

Table 5: Value Function J

Nodes (W)	Timesteps	Nonlinear iterations	CPU time (s)	$E_{p^*}^{t=0}[(W_T - \frac{\gamma}{2})^2]$	Ratio	p^*	Ratio for p^*
728	160	480	0.868	75.6533		2.9022	
1456	320	960	1.76	57.1392		2.8682	
2912	640	1295	4.4	47.8822		2.8015	
5824	1280	2561	9.58	43.2537		2.6734	
11648	2560	5120	25.9	40.9394		2.4358	

$\gamma = 14.47, |Wmin| = |Wmax| = 5925$. At each node, use analytic solution for optimal control. Apply fully implicit time stepping and use constant time steps

Table 6: Convergence Study: Std and Mean

Nodes (W)	Timesteps	Nonlinear iterations	$\text{Std}_{p^*}^{t=0}[W_T]$ [$\star$]	$\text{E}_{p^*}^{t=0}[W_T]$ [$\bigcirc$]	Ratio for [$\star$]	Ratio for [$\bigcirc$]
728	160	480	162.9945	3.4955		
1456	320	960	151.8204	3.0656		
2912	640	1295	274.8671	9.4466	0.5930	0.3700
5824	1280	2561	221.3216	6.2597	0.7365	0.5584
11648	2560	5120	406.4771	20.3663	0.4009	0.1716

$\text{Std}_{p^*}^{t=0}[W_T]$ and $\text{E}_{p^*}^{t=0}[W_T]$ are evaluated at (W = 1, t = 0). Analytic solution is ($\text{Std}_{p^*}^{t=0}[W_T]$, $\text{E}_{p^*}^{t=0}[W_T]$) = (0.8307, 6.9454)

6.1.2 Bounded Control

Followed by the procedure described in the paper, we can visualize the effect of putting constraints on the optimal control (wealth) by comparing their efficient frontiers. We achieve the same result as claimed in [8]: The bounded control cases are less optimal compared with no bankruptcy case (recall that in the no bankruptcy case, there is no constraint on the control). The result also verifies the common sense, if there is no limit on the position that one can take, the efficient frontier should also be more "efficient".

6.2 Cong and Oosterlee (2016)

6.2.1 A forward solution: the multi-stage strategy

Since the multi-stage method merely depends on solving a single-stage optimization problem at each time point, the problem can be implemented in the following straightforward fashion.

1. Generate the intermediate target values at each re-balancing time.
2. Compute the optimal allocation step by step starting from the initial state.

Using the parameters given in the table 7, we plot the efficient frontier for $\gamma \in [9.125, 81.25]$, shown in Figure 6.

6.2.2 A backward solution: Backward recursion approach

We choose the number of paths in the simulation experiment of geometric brownian motion processes to be 5000. The number of the paths are one tenth of the paths suggests in [2], as we find out that the simple five thousands

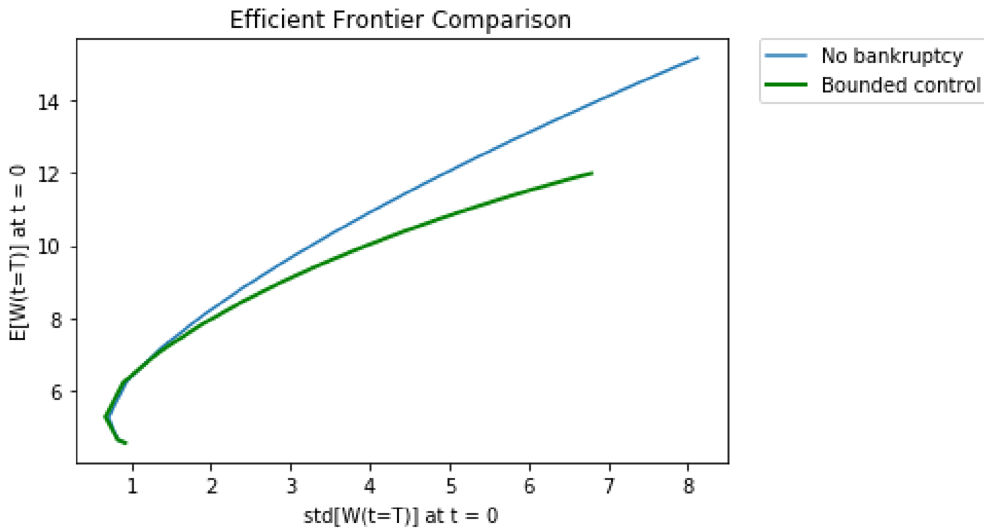

Figure 5: Comparison between Bounded control and No bankruptcy case

Table 7: Parameters used in Multi-staged Strategy

r	0.03	M	80
σ_1	0.15	NumOfPath	5000
T	20 years		
ξ_1	0.1	P_{max}	1.5
π	0.1	P_{min}	0.0
W(t=0)	1		

forward simulations for each final target $\gamma/2$, is very time consuming - it takes nearly thirty minutes to plot the efficient frontier on a 2.7 Ghz CPU. With this, the frontier is rather non-smooth compared with the one in [2]. We also tried to increase the number of paths in the simulation experiment, this results in smoother frontier indeed.

Implementing the backward recursion strategy is less straightforward compared with the forward Multi-stage strategy. As mentioned in the algorithm introduction, since the value function is used to evolve information between time steps, an error in calculating them will accumulate due to recursion, which is one of the significant disadvantages of this algorithm.

We also make a comparison between the efficient frontiers generated by two strategies in figure 7. One could observe that, for the multi-stage strategy and the backward recursive strategy, they do not intersect as the left endpoint, as shown in figure 5, the efficient frontier comparison for the numerical PDE method. In our opinion, the problems could be caused by, first, in the implementation, we do not use the wealth process generated by the multi-stage strategy as the input for the Backward recursion strategy - we generate the wealth processes based on an initial guess of the "optimal controls" within [0,1.5] to verify the author's conclusion - "The backward recursive programming is initiated with a reasonable guess for the asset allocations, which can be, but is not restricted to, the one generated by the multi-stage strategy." - can not be proved at the moment; second, the number of simulation is inadequate, so the range of the backward frontier looks the same as in [2].

The backward efficient frontier takes about ten minutes to plot on our 2.7 GHz CPU. This outcome does not meet our expectations of this algorithm - It saves the computational time from the PDE method.

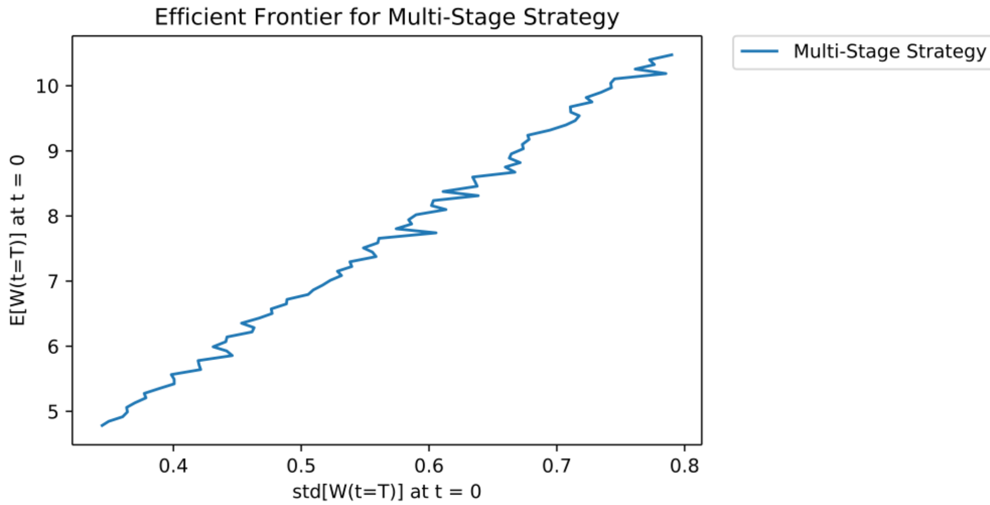

Figure 6: Efficient Frontier for Multi-stage Strategy

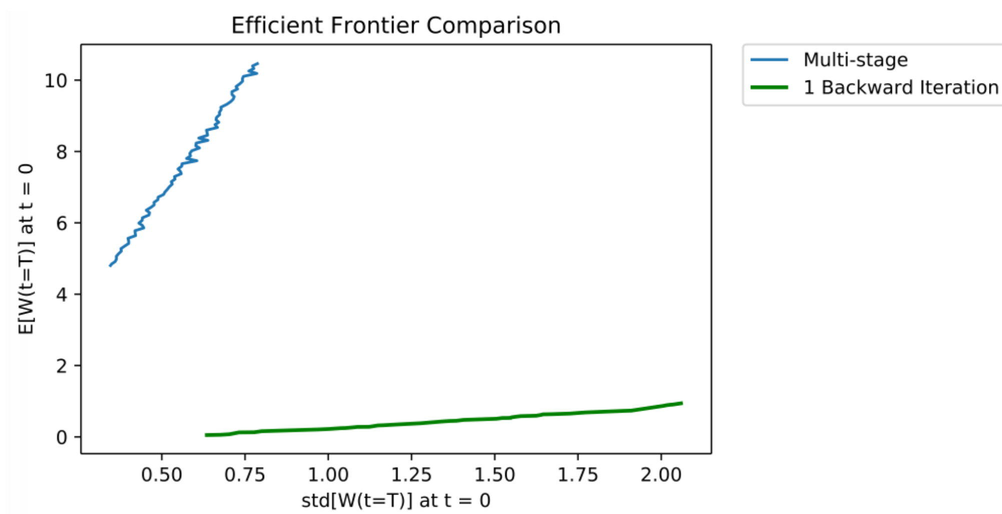

Figure 7: Comparison between Multi-Stage and Backward recursion approach

7 Conclusion and Enlightenment

Based on our understanding of the numerical results of the papers, we can make the following conclusions. For the pension plan problem with one risky asset and one risk-free asset, one should still use the numerical PDE method, for the sake of stability and computationally convenience. On the other hand, when one deals with high-dimensional problems with several risky assets, the multi-stage and backward recursive strategies via Monte-Carlo Simulations are recommended in terms of lower numerical error. The main deliverable for the paper is the GitHub code provides in Appendix C. with some useful utility function for one's convenience.

8 Acknowledgements

We would like to express our appreciation to all those who provided us with the advice and comments to finish this report, including but not limited to suggestions to provide further detailed derivations in the Appendices. Special gratitude to Professor Ben that provided us with encouragement and suggestions during the research. Furthermore, we would also like to appreciate the guidance given by all the authors from reference papers that render us a chance to study in the field of continuous mean-variance portfolio asset allocation problem.

Appendix A Prerequisite Knowledge and Proofs

A.1 Prerequisite Knowledge

HJB For a continuous-time dynamic programming, time $t \in R_+$, state $x \in \mathbb{X}$, and control $p \in \mathbb{P}$, $dx = u(p(t,x))dt + \sigma(p(t,x))dW(t)$ with a cost function V(x,t), the equation

$$0 = \inf_{p \in \mathbb{P}} \{V_t(x,p) + \mu_x V_x + \frac{1}{2}(\sigma_x)^2 V_{xx}\}$$

is called the Hamilton-Jacobo-Bellman equation

Ito's Formula Assume $dX = udt + \sigma dW$, let $Z(t) = f(t, X(t))$ be a $C^{1,2}$ function, then Z has a stochastic differential given by

$$df(t,X(t)) = \left(\frac{\partial f}{\partial t} + \mu \frac{\partial f}{\partial x} + \frac{\sigma^2}{2}\frac{\partial^2 f}{\partial x^2}\right) dt + \sigma \frac{\partial f}{\partial x} dW(t)$$

Proposition for Differentiable Convex Functions The differentiable function f of n variables defined on a convex set S is convex on S if and only if

$$f(x) - f(x^*) \geq \sum_{i=1}^{n} f_i'(x^*) \cdot (x_i - x_i^*), \forall x \in S, x^* \in S$$

A.2 Equivalence of (4) & (7)

For a fixed γ and $\lambda > 0$ we consider rearranging (4),

$$\begin{aligned}
&\max_{p(t,w)\in\mathbb{P}} \{\mathrm{E}^{t=0}[\mu W_T - \lambda W_T^2]\} \\
&= \max_{p(t,w)\in\mathbb{P}} \lambda\{\mathrm{E}^{t=0}[\frac{\mu}{\lambda} W_T - W_T^2]\} \\
&= \max_{p(t,w)\in\mathbb{P}} \{\mathrm{E}^{t=0}[\gamma W_T - W_T^2]\} \\
&= \inf_{p(t,w)\in\mathbb{P}} \{\mathrm{E}^{t=0}[W_T^2] - \gamma W_T\} \\
&= \inf_{p(t,w)\in\mathbb{P}} \{\mathrm{E}^{t=0}[W_T^2 - \gamma W_T + \frac{\gamma}{4}]\} \\
&= \inf_{p(t,w)\in\mathbb{P}} \{\mathrm{E}^{t=0}[(W_T - \frac{\gamma}{2})^2]\} \\
&= (7)
\end{aligned} \tag{46}$$

A.3 Proof of Theorem 3.1

Proof. Firstly, we denote the problem (3) as Q(p)

$$\Pi_{Q(p)} = \{p(\cdot) \mid p(\cdot) \text{ is an optimal control of Q(p)}\}$$

Secondly, we denote problem (4) as A(p,u),

$$\Pi_{A(p,u)} = \{p(\cdot) \mid p(\cdot) \text{ is an optimal control of A(p,u)}\}$$

where u $\in (-\infty, +\infty)$. Let $p^* \in \Pi_{Q(p)}$ and suppose $p^* \notin \Pi_{A(p,u)}$, as p^* is not an optimal of A(p,u), A(p,u) > A(p^*,u). It follows that there exist $p(\cdot)$ with corresponding $w(\cdot)$ s.t.

$$u(\mathrm{E}_p^{t=0}[W_T] - \mathrm{E}_{p^*}^{t=0}[W_T]) - \lambda(\mathrm{E}_p^{t=0}[W_T^2] - \mathrm{E}_{p^*}^{t=0}[W_T^2]) > 0 \tag{47}$$

Consider setting a function

$$\pi(x,y) = -\lambda x + \lambda y^2 + y \tag{48}$$

One can check that the hessian matrix of π has non-negative entries everywhere, then $\pi(x,y)$ is a convex function in (x,y) and

$$\pi(\mathrm{E}_p^{t=0}[W_T^2], \mathrm{E}_p^{t=0}[W_T]) = \mathrm{E}_p^{t=0}[W_T] - \lambda \mathrm{Var}_p^{t=0}[W_T]$$

which is exactly the objective function Q(p). Now, by the proposition in Appendix A.1 and the convexity of π (noting that $\pi_x(x,y)$ = -λ, $\pi_y(x,y) = 2\lambda y + 1$), we have

$$\begin{aligned}\pi(\mathrm{E}_p^{t=0}[W_T^2], \mathrm{E}_p^{t=0}[W_T]) \geq\ & \pi(\mathrm{E}_{p^*}^{t=0}[W_T^2], \mathrm{E}_{p^*}^{t=0}[W_T]) - \lambda(\mathrm{E}_p^{t=0}[W_T^2] - \mathrm{E}_{p^*}^{t=0}[W_T^2]) + \\ & \underbrace{(1 + 2\mathrm{E}_{p^*}^{t=0}[W_T])\lambda)}_{\mu}(\mathrm{E}_p^{t=0}[W_T] - \mathrm{E}_{p^*}^{t=0}[W_T]) \\ \geq\ & \pi(\mathrm{E}_{p^*}^{t=0}[W_T^2], \mathrm{E}_{p^*}^{t=0}[W_T]) + (-\lambda(\mathrm{E}_p^{t=0}[W_T^2] - \mathrm{E}_{p^*}^{t=0}[W_T^2]) - \\ & \mu(\mathrm{E}_p^{t=0}[W_T] - \mathrm{E}_{p^*}^{t=0}[W_T]))\end{aligned}$$

From inequality (47), it follows that

$$\begin{aligned}\pi(\mathrm{E}_p^{t=0}[W_T^2], \mathrm{E}_p^{t=0}[W_T]) &> \pi(\mathrm{E}_{p^*}^{t=0}[W_T^2], \mathrm{E}_{p^*}^{t=0}[W_T]) + 0 \\ &= \pi(\mathrm{E}_{p^*}^{t=0}[W_T^2], \mathrm{E}_{p^*}^{t=0}[W_T])\end{aligned} \tag{49}$$

By (49), p^* is not optimal for problem Q(p), but by assumption $p^* \in \Pi_{Q(p)}$, leading to a contradiction, so $p^* \in \Pi_{A(p,u))}$ □

A.4 General solution of U and V for large w

Let $\tau = T - t$ and perform change of variable for equations (14) & (16), we have

$$V_\tau = \mu_w^p V_w + \frac{1}{2}(\sigma_w^p)^2 V_{ww}$$
$$U_\tau = \mu_w^p U_w + \frac{1}{2}(\sigma_w^p)^2 U_{ww}$$

For U part, use $U_w(\tau) = J_1(\tau)$ and $U_{ww} = 0$ based from (24), it follows that

$$U_\tau = k_1 w U_w + \frac{1}{2}k_2 U_{ww} + \pi U_w = k_1 J_1(\tau) w + \pi J_1(\tau) \approx J_1'(\tau) w + J_2'(\tau) \tag{50}$$

where we denote $k_1 = r + p\sigma_1\xi_1$ & $k_2 = (p\sigma_1)^2$ and take the initial condition $U(\tau = 0) = w$ into account, then

$$U(\tau = 0) \approx J_1(0) w + J_2(0) = w \tag{51}$$

By (50)&(51), We have

$$\begin{cases} J_1'(\tau) = k_1 J_1(\tau) \\ J_1(0) = 1 \end{cases} \tag{52}$$

$$\begin{cases} J_2'(\tau) = \pi J_1(\tau) \\ J_2(0) = 0 \end{cases} \tag{53}$$

One can solve the above systems of linear ODEs and obtain $J_1(\tau) = e^{k_1\tau}$ and substitute it into (53), which gives $J_2(\tau) = \pi(e^{k_1\tau} - 1)$. Thus, $U(\tau) = e^{k_1\tau} w + \pi(e^{k_1\tau} - 1)$, and as w $\to \pm\infty$, $\pi(e^{k_1\tau} - 1)$ term can be omitted, so

$$U(\tau) \approx e^{k_1\tau} w \tag{54}$$

For V part, use $V_w(\tau) = 2H_1(\tau)w + H_2(\tau)$ and $V_{ww} = 2H_1(\tau)$ based from (23), it follows that

$$\begin{aligned}
V_\tau &\approx k_1 V_w w + \frac{1}{2}k_2 V_{ww} + \pi V_w = k_1[2H_1(\tau)w + H_2(\tau)]w + \frac{1}{2}[2H_1(\tau)k_2]w^2 + \pi[2H_1(\tau)w + H_2(\tau)] \\
&= \underbrace{[2k_1 H_1(\tau) + H_1(\tau)k_2]}_{w^2-term} w^2 + \underbrace{[k_1 H_2(\tau) + \pi 2H_1(\tau)]}_{w-term} w + \pi H_2(\tau) \\
&\approx H_1^{'}(\tau)w^2 + H_2^{'}(\tau)w + H_3^{'}(\tau)
\end{aligned} \tag{55}$$

where $k_1 = r + p\sigma_1\xi_1$ and $k_2 = (p\sigma_1)^2$. Take the initial condition $V(\tau = 0) = (w - \frac{\gamma}{2})^2$ into account, then

$$V(\tau = 0) \approx H_1(0)w^2 + H_2(0)w + H_3(0) = w^2 - \gamma w + \frac{\gamma^2}{4} \tag{56}$$

By (55)&(56), as w $\to \pm\infty$, w^2 will grow faster than w term and the constant term, so we only need to solve

$$\begin{cases} H_1^{'}(\tau) = (2k_1 + k_2)H_1(\tau) \\ H_1(0) = 1 \end{cases} \tag{57}$$

One can solve the above system and get $H_1(\tau) = e^{(2k_1+k_2)\tau}$. Thus, as w $\to \pm\infty$, we have

$$V(\tau) \approx e^{(2k_1+k_2)\tau} w^2 \tag{58}$$

A.5 Analytic Solution for Unconstrained Case

For p* part According to the RHS of equation (14), i.e. the inf$\{\cdot\}$ part,

$$\begin{aligned}
&\inf_{p\in\mathbb{P}}\{(\pi + w(r + p\sigma_1\xi_1))V_w + \frac{1}{2}(p\sigma_1 w)^2 V_{ww}\} \\
\Rightarrow &= \inf_{p\in\mathbb{P}}\{\underbrace{\frac{1}{2}\sigma_1^2 w^2 V_{ww}}_{A>0} p^2 + \underbrace{w\sigma_1\xi_1 V_w}_{B} p + \underbrace{wrV_w + \pi V_w}_{C}\} \\
\Rightarrow &= \inf_{p\in\mathbb{P}}\{Ap^2 + Bp + C\} \\
\Rightarrow &= \inf_{p\in\mathbb{P}}\{A(p + \frac{B}{2A})^2 - \frac{B^2}{4A} + C\} \Rightarrow \inf_{p\in\mathbb{P}}\{A(p + \frac{B}{2A})^2\}, A > 0
\end{aligned}$$

In this case, we are minimizing a quadratic equation here with positive constant A. The only case where the quadratic equation can be minimized is when $p* = \frac{B}{-2A} = -\frac{w\sigma_1\xi_1 V_w}{\sigma_1^2 w^2 V_{ww}} = -\frac{\xi_1}{\sigma_1 w} \cdot \frac{V_w}{V_{ww}}$ $(\star)$. Observe that in order to obtain the analytic solution of $p*$, we still need to solve for the exact solution of V, and that is the next part of our derivation.

For exact V part From [8], the author made the assumption here for large w, the optimal control p $\approx$ 0. This is reasonable in the sense that if the investor have large enough of wealth, then the marginal gain investing in the risky securities is relative small comparing to invest in a risk-free asset. Under this assumption, when w is large, the HJB PDE (14) can be reduce to

$$V_t + (rw + \pi)V_w = 0 \tag{59}$$

with terminal condition (15),

$$V(w, t = T) = (w - \frac{\gamma}{2})^2 \tag{60}$$

According to (23), We know the solution of V is a function of w that take the form of a quadratic equation,

$$V(w, t = T - \tau) = H_1(\tau)w^2 + H_2(\tau)w + H_3(\tau) \tag{61}$$

$$\Rightarrow V_t = \frac{\partial V}{\partial \tau} \cdot \frac{\partial \tau}{\partial t} = -\frac{\partial V}{\partial \tau} = -H_1^{'}(\tau)w^2 - H_2^{'}(\tau)w - H_3^{'}(\tau) \tag{62}$$

$$\Rightarrow V_w = 2H_1(\tau)w + H_2(\tau) \tag{63}$$

Now, we plug in the initial condition (63) into (64)

$$V_t(\tau = T - t = 0) = w^2 - \gamma w + \frac{\gamma^2}{4} = H_1(0)w^2 + H_2(0)w + H_3(0) \tag{64}$$

Plug (62) & (63) into (59) and match the coefficient of w^2, w and constant term accordingly, we have the following linear system of first-order Ordinary differential equations (ODE) with initial conditions, (assume $W(p, t = T - \tau = 0) = \hat{w}_0$)

$$\begin{cases} H_1'(\tau) = 2rH_1(\tau) \\ H_2'(\tau) = 2\pi H_1(\tau) + rH_2(\tau) \\ H_3'(\tau) = \pi H_2(\tau) \\ H_1(0) = \hat{w}_0^2 \\ H_2(0) = -\hat{w}_0\gamma \\ H_3(0) = \frac{\gamma^2}{4} \end{cases} \tag{65}$$

One can easily solve for coefficients $H_1(\tau)$, $H_2(\tau)$ and $H_3(\tau)$ of the exact solution V(t, w) to PDE (61), when w is large, and the exact solution of V(t, w) for large w is as below,

$$V(w, t = T - \tau) = H_1(\tau)w^2 + H_2(\tau)w + H_3(\tau) \tag{66}$$

where $\tau = T - t$ and,

$$\begin{cases} H_1(\tau) = \mathrm{e}^{2r\tau} \\ H_2(\tau) = -w_0\gamma\mathrm{e}^{r\tau} + \frac{2\pi}{r}\mathrm{e}^{r\tau}(\mathrm{e}^{r\tau} - 1) \\ H_3(\tau) = \frac{\gamma^2}{4} - \frac{w_0\pi\gamma(\mathrm{e}^{r\tau}-1)}{r} + \frac{\pi^2(\mathrm{e}^{r\tau}-1)^2}{r^2} \end{cases} \tag{67}$$

Therefore, we can use the exact solution of V to calculate V_w and V_{ww} as the following,

$$\begin{cases} V_w = 2H_1(\tau)w + H_2(\tau) = 2\mathrm{e}^{2r\tau}w + \left(-w_0\gamma\mathrm{e}^{r\tau} + \frac{2\pi}{r}\mathrm{e}^{r\tau}(\mathrm{e}^{r\tau} - 1)\right) \\ V_{ww} = 2H_1(\tau) = 2\mathrm{e}^{2r\tau} \end{cases}$$

then we substitute V_w and V_{ww} into ($\star$) to obtain $p*$,

$$p^*(t, w) = -\frac{\xi_1}{\sigma_1 w}[w - \frac{w_0\gamma}{2}\mathrm{e}^{-r(T-t)} + \frac{\pi}{r}\mathrm{e}^{-r(T-t)}(\mathrm{e}^{-r(T-t)} - 1)], \forall t \in [0, T] \tag{68}$$

For γ part Observed that the only unknown part for $p*$'s formula is the γ part for large w. According to [4], we can consider taking the expectation of dW an dW^2 and solve the resultant system of linear ODEs to obtain the value for γ.

According to (2), we can obtain the dynamic of wealth under optimal control (68) as,

$$dW_t = [(r + p^*\xi_1\sigma_1)W_t + \pi]\, dt + p^*W_t\sigma_1 dZ_t \tag{69}$$

then by application of Ito's lemma from Appendix A.1 to dW_t, we obtain the following SDE that gives the dynamic of $f(w, t) = W_t^2$ under optimal control p^* (68), (as $\frac{\partial f}{\partial t} = 0$, $\frac{\partial f}{\partial W} = 2W_t, \frac{\partial^2 f}{\partial W^2} = 2$)

$$dW_t^2 = \left(2[\mu_w^{p^*}W_t] + (\sigma_w^{p^*})^2\right) dt + 2\sigma_w^{p^*}W_t dZ_t \tag{70}$$

where μ_w^p is previously defined as $(r + p\xi_1\sigma_1)W_t + \pi$ and σ_w^p is previously defined as $pW_t\sigma_1$.

if we consider taking the expectation for both equation (69)& (70), the dW term would vanish and leave them as two system of linear ODEs as below,

$$d\mathrm{E}_{p^*}[W_t] = [(r + p^*\xi_1\sigma_1)E[W_t] + \pi]dt, \text{ with } \mathrm{E}[W_0] = w_0$$
$$d\mathrm{E}_{p^*}[W_t^2] = 2dt\left[r + p^*\xi_1\sigma_1)E[W_t^2] + \pi E[W_t]\right] + (p^*W_t\sigma_1)^2 dt, \text{ with } \mathrm{E}[W_0^2] = w_0^2$$

According to [4], by solving the system of ODEs, we can find the expected value of the portfolio wealth at terminal time T under p^* is as the following,

$$\mathrm{E}_{p^*}[W_T] = (w_0 + \frac{\pi}{r})\mathrm{e}^{-(\xi_1^2 - r)T} + \frac{\gamma}{2}(1 - \mathrm{e}^{-\xi_1^2 T}) - \frac{\pi}{r}\mathrm{e}^{-\xi_1^2 T} \tag{71}$$

Similarly, the expected value of the square portfolio wealth at time T under under p^* is as the following,

$$\mathrm{E}_{p^*}[W_T^2] = (w_0 + \frac{\pi}{r})^2\mathrm{e}^{-(\xi_1^2 - 2r)T} + \frac{\gamma}{2}(1 - \mathrm{e}^{-\xi_1^2 T}) - \frac{2\pi}{r}(w_0 + \frac{\pi}{r})\mathrm{e}^{-(\xi_1^2 - r)T} + \frac{\pi^2}{r^2}\mathrm{e}^{-\xi_1^2 T} \tag{72}$$

Recalled that $\gamma = \frac{1}{\lambda} + 2\mathrm{E}_{p^*}[W_T]$ (6), and based on (71), we can obtain the following equation,

$$\begin{aligned} \gamma &= \frac{1}{\lambda} + 2((w_0 + \frac{\pi}{r})\mathrm{e}^{-(\xi_1^2 - r)T} + \frac{\gamma}{2}(1 - \mathrm{e}^{-\xi_1^2 T}) - \frac{\pi}{r}\mathrm{e}^{-\xi_1^2 T}) \\ \Rightarrow \gamma &= \mathrm{e}^{\xi_1^2 T}\frac{1}{\lambda} + 2(w_0 + \frac{\pi}{r})\mathrm{e}^{rT} - \frac{2\pi}{r} \end{aligned} \tag{73}$$

Now, we have the value of γ from (73), we can obtain the exact value for the optimal control p^*,

$$p^*(t, w) = -\frac{\xi_1}{\sigma_1 w}[w - (\hat{w_0}e^{rt} + \frac{\pi}{r}(e^{rt} - 1)) - \frac{e^{-r(T-t) + \xi_1^2 T}}{2\lambda}], \forall t \in [0, T] \tag{74}$$

Furthermore, since we have the value of γ from (73), the expected value of terminal wealth under optimal control can be written as below,

$$\mathrm{E}_{p^*}^{t=0}[W_T] = w_0\mathrm{e}^{rT} + \pi\frac{\mathrm{e}^{rT} - 1}{r} + \frac{\mathrm{e}^{\xi_1^2 T} - 1}{2\lambda} \tag{75}$$

Apart from that, we can also determine the variance of the terminal wealth under optimal control based on (72)&(73) using variance formula as below,

$$\mathrm{Var}_{p^*}^{t=0}[W_T] = \frac{e^{\xi_1^2 T} - 1}{4\lambda^2} \tag{76}$$

We can further simplified the analytic solution of the expected value of terminal wealth and variance of the terminal wealth under optimal control for large w as,

$$\begin{aligned} \mathrm{Var}_{p^*}^{t=0}[W_T] &= \frac{e^{\xi_1^2 T} - 1}{4\lambda^2} \\ \mathrm{E}_{p^*}^{t=0}[W_T] = w_0\mathrm{e}^{rT} + \pi\frac{\mathrm{e}^{rT} - 1}{r} + \frac{\mathrm{e}^{\xi_1^2 T} - 1}{2\lambda} &= w_0\mathrm{e}^{rT} + \pi\frac{\mathrm{e}^{rT} - 1}{r} + \sqrt{\mathrm{e}^{\xi_1^2 T} - 1}\,\mathrm{Std}(W_T) \end{aligned} \tag{77}$$

Eventually, we have the analytic solution for unconstrianed control case derived as (74) & (77).

A.6 Derivative Discretization

This is a general review of using Taylor expansion to estimate first and second derivatives of a function f(x). Consider the function of one variable f(x) shown in Figure 8., by using the forward and backward scheme of the Taylor series around point x, the point we would like to estimate for f(x)

$$\textbf{Forward scheme: } f(x+h) = f(x) + hf^{'}(x) + \frac{h^2}{2}f^{''}(x) + \frac{h^3}{3}f^{'''}(x) + \mathcal{O}(h^4) \tag{78}$$

$$\textbf{Backward scheme: } f(x-h) = f(x) - hf^{'}(x) + \frac{h^2}{2}f^{''}(x) - \frac{h^3}{3}f^{'''}(x) + \mathcal{O}(h^4) \tag{79}$$

Based on (78) & (79), the following approximations can be derived with the order of accuracy:

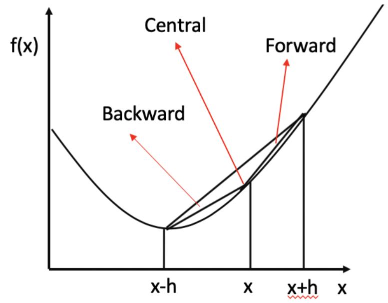

Figure 8: Illustration of PDE Discretisation

- Forward approximation of 1^{st} derivative of f(x)

$$f'(x) = \frac{\mathrm{f}(x+h) - \mathrm{f}(x)}{\mathrm{h}} + \mathcal{O}(h)$$

- Backward approximation of 1^{st} derivative of f(x)

$$f'(x) = \frac{\mathrm{f}(x) - \mathrm{f}(x-h)}{\mathrm{h}} + \mathcal{O}(h)$$

- Central approximation of 1^{st} derivative of f(x)

$$f'(x) = \frac{\mathrm{f}(x+h) - \mathrm{f}(x-h)}{2h} + \mathcal{O}(h^2)$$

- Standard approximation of 2^{nd} derivative of f(x)

$$f''(x) = \frac{\mathrm{f}(x+h) - 2f(x) + \mathrm{f}(x-h)}{\mathrm{h}^2} + \mathcal{O}(h^2)$$

A.7 Discrete equation coefficients

A.7.1 Central Scheme

Based on Appendix A.6, we applied forward difference in time formula of $\frac{\partial V}{\partial t}$, central difference approximation of $\frac{\partial V}{\partial w}$ and standard approximation of $\frac{\partial^2 V}{\partial w^2}$ to the operator (note: $w_i = i\Delta w$) $(L_p^h V^n)_i$ (or $(L_p^h U^n)_i$),

$$(L_p^h V^n)_i = \frac{1}{2}\underbrace{(pi\Delta W\sigma_1)^2}_{(\sigma_{w_i^p})^2}\frac{V_{i+1}^n - 2V_i^n + V_{i-1}^n}{(\Delta W)^2} + \underbrace{(\pi + (r + p\xi_1)i\Delta W)}_{(\mu_{w_i^p})}\frac{V_{i+1}^n - V_{i-1}^n}{2\Delta W}$$

we then use the above approximation to the operator $(L_p^h V^n)_i$ (or $(L_p^h U^n)_i$) and substitute it into (14) (or (16)). After careful re-arrangement, the HJB equation becomes,

$$\frac{V_i^{n+1} - V_i^n}{\Delta t} + \inf_{p\in P}\Bigg[\underbrace{\left(\frac{1}{2}\sigma_1^2 p^2 i^2 - \frac{\pi + i\Delta W(r + p\sigma_1\xi_1)}{2\Delta W}\right)}_{\alpha_{i,\text{central}}^n} V_{i-1}^n$$

$$+ \underbrace{\left(\frac{1}{2}\sigma_1^2 p^2 i^2 + \frac{\pi + i\Delta W(r + p\sigma_1\xi_1)}{2\Delta W}\right)}_{\beta_{i,\text{central}}} V_{i+1}^n - \underbrace{(\sigma_1^2 p^2 i^2)}_{(\alpha_{i,\text{central}}^n + \beta_{i,\text{central}}^n)} V_i^n\Bigg] = 0$$

with coefficients for the general formula (30) as below,

$$\alpha_{i,\text{central}}^n = \frac{1}{2}\sigma_1^2 p^2 i^2 - \frac{\pi + i\Delta W(r + p\sigma_1\xi_1)}{2\Delta W} \tag{80}$$

$$\beta_{i,\text{central}}^n = \frac{1}{2}\sigma_1^2 p^2 i^2 + \frac{\pi + i\Delta W(r + p\sigma_1\xi_1)}{2\Delta W} \tag{81}$$

A.7.2 Forward Scheme

Based on Appendix A.6, we applied forward difference in time formula of $\frac{\partial V}{\partial t}$, forward difference approximation of $\frac{\partial V}{\partial w}$ and standard approximation of $\frac{\partial^2 V}{\partial w^2}$ to the operator (note: $w_i = i\Delta w$) $(L_p^h V^n)_i$ (or $(L_p^h U^n)_i$),

$$(L_p^h V^n)_i = \frac{1}{2}\underbrace{(pi\Delta W\sigma_1)^2}_{(\sigma_{w_i}^p)^2}\frac{V_{i+1}^n - 2V_i^n + V_{i-1}^n}{(\Delta W)^2} + \underbrace{(\pi + (r + p\xi_1)i\Delta W)}_{(\mu_{w_i}^p)}\frac{V_{i+1}^n - V_i^n}{\Delta W}$$

we then use the above approximation to the operator $(L_p^h V^n)_i$ (or $(L_p^h U^n)_i$) and substitute it into (14) (or (16)). After careful re-arrangement, the HJB equation becomes,

$$\frac{V_i^{n+1} - V_i^n}{\Delta t} + \inf_{p\in P}\left[\underbrace{\frac{1}{2}\sigma_1^2 p^2 i^2}_{\alpha_{i,\text{forward}}^n} V_{i-1}^n + \underbrace{\left(\frac{1}{2}\sigma_1^2 p^2 i^2 + \frac{\pi + i\Delta W(r + p\sigma_1\xi_1)}{\Delta W}\right)}_{\beta_{i,\text{forward}}^n} V_{i+1}^n \right.$$
$$\left. - \underbrace{\left(\sigma_1^2 p^2 i^2 + \frac{\pi + i\Delta W(r + p\sigma_1\xi_1)}{\Delta W}\right)}_{(\alpha_{i,\text{forward}}^n + \beta_{i,\text{forward}}^n)} V_i^n\right] = 0$$

with coefficients for the general formula (30) as below,

$$\alpha_{i,\text{forward}}^n = \frac{1}{2}\sigma_1^2 p^2 i^2 \tag{82}$$

$$\beta_{i,\text{forward}}^n = \frac{1}{2}\sigma_1^2 p^2 i^2 + \frac{\pi + i\Delta W(r + p\sigma_1\xi_1)}{\Delta W} \tag{83}$$

A.7.3 Backward Scheme

Based on Appendix A.6, we applied forward difference in time formula of $\frac{\partial V}{\partial t}$, backward difference approximation of $\frac{\partial V}{\partial w}$ and standard approximation of $\frac{\partial^2 V}{\partial w^2}$ to the operator (note: $w_i = i\Delta w$) $(L_p^h V^n)_i$ (or $(L_p^h U^n)_i$),

$$(L_p^h V^n)_i = \frac{1}{2}\underbrace{(pi\Delta W\sigma_1)^2}_{(\sigma_{w_i}^p)^2}\frac{V_{i+1}^n - 2V_i^n + V_{i-1}^n}{(\Delta W)^2} + \underbrace{(\pi + (r + p\xi_1)i\Delta W)}_{(\mu_{w_i}^p)}\frac{V_i^n - V_{i-1}^n}{\Delta W}$$

we then use the above approximation to the operator $(L_p^h V^n)_i$ (or $(L_p^h U^n)_i$) and substitute it into (14) (or (16)). After careful re-arrangement, the HJB equation becomes,

$$\frac{V_i^{n+1} - V_i^n}{\Delta t} + \inf_{p\in P}\left[\underbrace{\left(\frac{1}{2}\sigma_1^2 p^2 i^2 - \frac{\pi + i\Delta W(r + p\sigma_1\xi_1)}{\Delta W}\right)}_{\alpha_{i,\text{backward}}^n} V_{i-1}^n + \underbrace{\frac{1}{2}\sigma_1^2 p^2 i^2}_{\beta_{i,\text{backward}}^n} V_{i+1}^n \right.$$
$$- \underbrace{\left(\sigma_1^2 p^2 i^2 - \frac{\pi + i\Delta W(r + p\sigma_1\xi_1)}{\Delta W}\right)}_{(\alpha_{i,\text{backward}}^n + \beta_{i,\text{backward}}^n)} V_i^n] = 0$$

with coefficients for the general formula (30) as below,

$$\alpha_{i,\text{backward}}^n = \frac{1}{2}\sigma_1^2 p^2 i^2 - \frac{\pi + i\Delta W(r + p\sigma_1\xi_1)}{\Delta W} \tag{84}$$

$$\beta_{i,\text{backward}}^n = \frac{1}{2}\sigma_1^2 p^2 i^2 \tag{85}$$

A.8 Transferring (31)&(32) into the Matrix form

Follow the instructions from [8], we can transfer equations (31) & (32) into matrix form. Firstly, we write the differential operator $(L_p^h V^n)_i$ (30) as,

$$
\begin{aligned}
(L_p^h V^n)_i &= \alpha_i^n V_{i-1}^n + \beta_i^n V_{i+1}^n - (\alpha_i^n + \beta_i^n) V_i^n \\
&= (A^n(P^n)V^n)_i
\end{aligned}
$$

where $P^n = [p_0^n, \ldots, p_k^n]$ is the row vector of local optimal control according to the local wealth $W = [W_0 = W_{min}, \ldots, W_k = W_{max}]$.

Now, Define a boundary condition vector $G^n = [G_0^n, \ldots, G_k^n]$ and the last entry G_k^n are calculated using the general solution of boundary condition (25) for large W, (G_0^n is trivial to show for W_{min})

$$
\begin{aligned}
V(w_{max}, \tau = T - t) &= \mathrm{e}^{(2(r + p\sigma_1\xi_1) + (p\sigma_1)^2)\tau} W_{max}^2 \\
\Rightarrow G_p^n = V(w_{max}, T - n\Delta t) &= e^{(2(r + p\sigma_1\xi_1) + (p\sigma_1)^2)(T - n\Delta t)} W_{max}^2
\end{aligned}
$$

We now can transfer equation (31) into matrix form,

$$
\begin{aligned}
&\frac{V_i^{n+1} - V_i^n}{\Delta\tau} - \inf_{p^n \in \mathbb{P}_h} \{(L_{p^n}^h V^n)_i\} = 0 \\
\Rightarrow &\frac{V_i^{n+1} - V_i^n}{\Delta\tau} - A^n(P^n)V^n = 0 \\
\Rightarrow &V_i^{n+1} - V_i^n - \Delta\tau A^n(P^n)V^n = 0 \\
\Rightarrow &\left[I - \Delta\tau A^n(P^n)\right] V_i^n = V_i^{n+1} + \left(G^{n+1} - G^n\right), \text{ with } p_i^n = \inf_{p \in P_h} \{(L_p^h V^n)_i\} \qquad (86)
\end{aligned}
$$

In this case, we need to modify the first and last row of A^n under different boundary condition in separate cases discuss in section 4.1.2. The Dirichlet boundary condition is enforced by using term $(G^{n+1} - G^n)$.

In addition, the same rationale can apply to U. As a result, we have the matrix form for equation (32),

$$
\left[I - \Delta\tau A^n(P^n)\right] U_i^n = U_i^{n+1} + \left(H^{n+1} - H^n\right), \text{ with each } p_i^n \text{ from (86)} \qquad (87)
$$

A.9 Grid for allowing bankruptcy scenario

For the allowing bankruptcy case, we have $D = (-\infty, \infty)$ and $P = (-\infty, \infty)$. In this case, according to [8], the W grid becomes $[w_{min}, \ldots, w_{-2}, w_{-1}, w_1, w_2, \ldots, w_{max}]$. Observe that, the new W grid does not contain w = 0. This is due to the fact that no information can be passed between the negative nodes and positive nodes if $w = 0$. As in the algorithm, we are moving in a iterative scheme if one of the nodes does not contain the information, then the iteration process will stop at that node. For that reason, we have to make sure $w = 0$ will not in the grid by adding two new nodes $w_{-1}^{new} = w_{-1}/2$ between node w_{-1} and w = 0, and $w_1^{new} = w_1/2$ between node w = 0 and w_1 as the following figure.

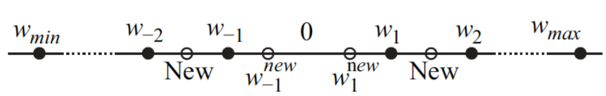

Figure 9: Node in W grid for allowing bankruptcy case

A.10 Derivation for no bankruptcy case constraint

Firstly, to make sure that the allocation at time t will not lead to bankruptcy in the successive time $t + \Delta t$ we need,

$$
\begin{aligned}
&W_t \cdot (p_t R_t^e + R_f) + \pi \cdot \Delta t \geq 0 \\
\Rightarrow &p_t R_t^e \geq -R_f - \frac{\pi \Delta t}{W_t}, \text{as } W_t > 0 \text{ for no bankruptcy case,}
\end{aligned}
$$

Since we know $R_t^e = \mathrm{e}^{r_t^e} - R_f$, r^e a random variable, we require,

$$-R_f \geq -\frac{R_f}{p_t} - \frac{\pi\Delta t}{W_t \cdot p_t}, \text{ and } p_t \geq 0$$

we then re-arrange the constraint from above to ensure a valid point $p_t \geq 0$. Eventually, after careful re-arrangement, we have

$$\begin{aligned} &- R_f p_t \geq -R_f - \frac{\pi\Delta t}{W_t} \ \&\ p_t \geq 0 \\ \Rightarrow & p_t \leq 1 + \frac{\pi\Delta t}{W_t \cdot R_f} \ \&\ p_t \geq 0, \text{ as } R_f \text{ always positive} \\ \Rightarrow & 0 \leq p_t \leq 1 + \frac{\pi\Delta t}{W_t \cdot R_f} \end{aligned} \tag{88}$$

Appendix B Algorithms

B.1 Multi-stage strategy in a forward-fashion

The policy for Multi-stage strategy with a forward-fashion is as follow,

- Generate the intermediate target values at each re-balancing time.
- Compute the optimal allocation step by step starting from the initial state.

B.2 Backward recursive programming

In this case, we consider letting $x_t = p_t$ for the consistency between the report and the original paper [2],

- **Step 1: Initiation**:
  Generate an initial guess of optimal asset allocations $\{\tilde{x}_t\}_{t=0}^{T-\Delta t}$ and simulate the paths of optimal wealth values $\{W_t(i)\}_{i=1}^{N}, t = 0, \dots, T$. At the terminal time T, we have the determined value function $J_T(W_T)$.
  The following three steps are subsequently performed, recursively, backward in time, at $t = T - \Delta t, \dots, \Delta t, 0$.
- **Step 2: Solving**
  Bundle paths into B partitions, where each bundle contains a similar number of paths and the paths inside a bundle have similar values at time t. Denote the wealth values associated to the paths in the bundle by $\{W_t^b(i)\}_{i=1}^{N_B}$, where N_B is the number of paths in the bundle. Within each bundle, we perform the following procedure.
  - For paths in the bundle, we have the corresponding wealth values $\{W_{t+\Delta t}^b(i)\}_{i=1}^{N_B}$ and the continuation values $\{J_{t+\Delta t}^b(i)\}_{i=1}^{N_B}$ at time $t+\Delta t$. So, a function $f_{t+\Delta t}^b(\cdot)$, which satisfies $J_{t+\Delta t}^b = f_{t+\Delta t}^b(W_{t+\Delta t}^b)$ on the local domain, can be determined by *regression*.[5]
  - For all paths in the bundle, since the value function $f_{t+\Delta t}^b(W_{t+\Delta t}^b)$ has been approximated, we solve the optimization problem by calculating the first-order conditions. In this way, we get new asset allocations $\{\hat{x}_t^b(i)\}_{i=1}^{N_B}$.
  - Since the wealth values $\{W_t^b(i)\}_{i=1}^{N_B}$ and the allocations $\{\hat{x}_t^b(i)\}_{i=1}^{N_B}$ are known, by *regression* we can also compute the new continuation values $\{\hat{J}_t^b(i)\}_{i=1}^{N_B}$. Here $\hat{J}_t^b(i)$ is the expectation of $J_{t+\Delta t}(W_{t+\Delta t})$ conditional on $W_t^b(i)$ and $\hat{x}_t^b(i)$, that is,

    $$\hat{J}_t^b(i) = \mathbb{E}[J_{t+\Delta t}(W_{t+\Delta t}) | W_t = W_t^b(i), x_t = \hat{x}_t^b(i)].$$

- **Step 3: Updating**
  For the paths in a bundle, since we have an old guess $\{\tilde{x}_t^b\}_{i=1}^{N_B}$ for the asset allocations, by *regression* we can also calculate the old continuation values $\{\tilde{J}_t^b(i)\}_{i=1}^{N_B}$. For the i-th path, if $\tilde{J}_t^b(i) > \hat{J}_t^b(i)$, we choose $\hat{x}_t^b(i)$ as the updated allocation. Otherwise we retain the initial allocation. We denote the updated allocations by $\{x_t^b(i)\}v_{i=1}^{N_B}$.
- **Step 4: Evolving**
  Once the updated allocations $\{x_t^b(i)\}_{i=1}^{N_B}$ are obtained, again by *regression* we can calculate the "updated" continuation values $\{J_t^b(i)\}_{i=1}^{N_B}$ and proceed with the backward recursion.

Appendix C Python Code

GitHub URL: https://github.com/c3qian/Continuous-time-Portfolio-Optimization

References

[1] Barles, G., Souganidis, P.,1991. Convergence of approximation schemes for fully nonlinear equations. Asymptotic Analysis 4,271–283.

[2] Cong, Fei, and Cornelis W. Oosterlee."Multi-period mean–variance portfolio optimization based on Monte-Carlo simulation." Journal of Economic Dynamics and Control 64 (2016): 23-38.

[3] Duffie, D., Richardson, H.,1991.Mean-variance hedging in continuous time. Annals of Applied Probability 1,1–15.

[4] Højgaard, B., & Vigna, E. 2007. Mean-variance portfolio selection and efficient frontier for defined contribution pension schemes. Department of Mathematical Sciences, Aalborg University. Research Report Series, No. R-2007-13

[5] Jain, Shashi, Oosterlee, Cornelis W., 2015. The stochastic grid bundling method: efficient pricing of Bermudan options and their Greeks. Appl. Math. Comput. 269(1),412–431.

[6] Li, D., Ng, W.-L., 2000. Optimal dynamic portfolio selection: multi-period mean variance formulation. Mathematical Finance 10, 387–406.

[7] Markowitz H (1952) Portfolio selection. J Finance 7: 77–91

[8] Wang, J. & Forsyth, P.A., 2010. "Numerical solution of the Hamilton-Jacobi-Bellman formulation for continuous-time mean-variance asset allocation," Journal of Economic Dynamics and Control, Elsevier, vol. 34(2), pages 207-230, February.

[9] Wang, J., Forsyth, P.,2008.Maximal use of central differencing for Hamilton–Jacobi–Bellman PDEs in finance.SIAM Journal on Numerical Analysis 46, 1580–1601.

[10] Wang, Z., Xia, J., Zhang, L., 2007. Optimal investment for the insurer: the martingale approach. Insurance: Mathematics and Economics 40, 322–334.

[11] Zhou,X.,Li,D.,2000.Continuous time mean variance portfolio selection:a stochastic LQ framework. Applied Mathematics and Optimization 42, 19–33.