University Of Waterloo

STAT 906

Computer Intensive Method in Finance

Technical Report
Multi-period mean–variance portfolio optimization based on Monte-Carlo simulation

Name:

Chao Qian

12 August, 2020

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. This optimization problem is solved by introducing the Lagrange multiplier to associate the above-mentioned two criteria. However, due to the presence of the conditional variance term in the objective, the problem is not in a standard form for dynamic programming. It is shown that in (3; 6), 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, I would briefly introduce a typical way to solve the constraint mean-variance optimization problem - the finite difference method suggested in (4) to solve the resulting nonlinear Hamilton-Jacobi-Bellman (HJB) PDE. The numerical solutions of the problem are presented in the report in the form of an efficient frontier comparison under certain constraints on the portfolio wealth and optimal control. Later in this report, I will walk through another technique to deal with constraint dynamic mean-variance optimization problem - a simulated-based approach. I discuss two simulation methods presented in the paper: 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 comparison under constraint on the portfolio wealth and control with detailed analysis. The main reference paper for this project is (4) and (1).

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 Markowitz's pioneering work (**?**), the single-period mean-variance portfolio optimization model has been widely used and innovated by scholar in academia and industrial practitioners. As an extension to Markowitz's work, multi-period mean-variance asset allocation problem have also been well studied over the years (3**?** ; 2; 1). Many financial applications are using the continuous-mean variance model including pension asset allocation (2) (later revealed in the report), insurance (5), and for hedging derivatives (**?**). For the sake of simplicity, 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 occurrence of non-linear conditional variance term. To resolve this problem, according to (4), 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 (4), one can implement the finite difference method to solve the Hamilton-Jacobi-Bellman (HJB) PDE iteratively and the analytic solution also exists under specific conditions.

Using a finite difference method to solve for HJB equations will give us accurate and stable results. However, it is rather computational expensive to implement the algorithm in only two assets realm, not to mention if we are trying to extend to mulit-dimensional case. the authors in (1) proposed a better method of applying the Monte-Carlo Simulation under the same transformed problem using the stochastic LQ method, but using a forward fashion multi-stage strategy and an improved backward recursive strategy based from the former policy via Monte-Carlo simulation. Throughout this report, I 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.

3 Problem Setting

This section provides the problem settings for the dynamic portfolio optimization problem in a defined pension allocation plan for both (**?**) and (1). Let,

- Ω be the set of all possible realizations in the financial market within the time horizon [0, T].
- $\mathcal{F}$ be the sigma algebra of events at time T i.e. $\mathcal{F} = \mathcal{F}_T$
- $\mathcal{P}$ be the probability measure define on $\mathbb{F}$
- $\mathbb{D}$:= the set of all admissible wealth w(t), for 0$\leq$ t $\leq$ T, and $\mathbb{P}$:= the set of all admissible controls p(t,w), for 0$\leq$ t $\leq$ T and w $\in \mathbb{D}$

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.

The original objective function is the following, where λ (> 0) represents the level of risk aversion,

$$V_t(W_t) = \max_{\{p_s\}_{s=t}^{T-\Delta t}} \{\underbrace{\mathrm{E}^{t=0}[W_T|W_t]}_{Reward} - \lambda \underbrace{\mathrm{Var}^{t=0}[W_T|W_t]}_{Risk}\}, \text{ subject to } W_t \tag{1}$$

subject to the wealth restriction:

$$W_{s+\Delta t} = W_s \cdot (p_s R_s^e + R_f) + \pi \cdot \Delta t, \quad s = t, \cdots, T - \Delta t. \tag{2}$$

Where, p_s denotes the asset allocation of the investor's wealth in the risky asset in the period $[s, s+\Delta t]$. we assume that $p_t \in F_t$. R_f is the risk-free return in one time step, which is assumed to be constant for simplicity, and R_s^e is the excess return of the risky asset during the same investment period. We assume that the excess returns $\{p_s\}_{s=t}^{T-\Delta t}$ are statistically independent. $C \cdot \Delta t$ stands for a contribution of the investor in the portfolio during this period.λ is the risk reversion parameter. A very important concept throughout this report is the value function, $v_t(W_t)$ which depends on the portfolio wealth W_t. This value function measures the investor's investment opportunities at time t with wealth W_t.

Due to non-linearity of conditional variance term in (1), dynamic programming valuation approach is not applicable. Namely, $Var[Var[W_T|F_t]|F_s] \neq Var[W_T|F_s], \forall s \leq t$. However, due to following theorem by Zhou and Li(2000), this mean-variance problem can be transformed into another one, a target-based optimization problem, leaving the optimal control unchanged.

Theorem 3.1. *if $p^*(t,w) \in$ the optimal control of problem (5), then $p^*(t,w)$ is also an optimal control of*

$$\min_{\{p_s\}_{s=t}^{T-\Delta t}} E^{t=0}\left[(W_T - \frac{\gamma}{2})^2 \Big| W_t = w\right] \tag{3}$$

Where,

$$\gamma = \frac{1}{\lambda} + 2E_{p^*}^{t=0}[W_T|W_t] \tag{4}$$

Proof. see Appendix A.3 □

This LQ formulation also called pre-commitment optimization problem and for numerical computation, this pre-commitment problem is usually formulated as a HJB PDE, with appropriate boundary condition, just as did in (4).

We have the general setting of the problem defined for both papers. I will briefly go through the results in the (4), for more numerical details, please refer to my previous technical report.

4 Wang and Forsyth (2010)

The Numerical HJB PDE approach follows the following steps: problem formulation, localization, and discretization.

4.1 Mean variance efficient wealth case

Based on the theoretical results in (6), we embedded the improved objective into a stochastic LQ problem by defining

$$\begin{aligned} V(w,t) &= \inf_{p\in\mathbb{P}} E^{t=0}\left[(W_T-\frac{\gamma}{2})^2 \middle| W_t = w\right] && (5)\\ &= \inf_{p\in\mathbb{P}} J(t,w,p) && (6)\end{aligned}$$

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,

$$\begin{aligned} dV &= V_t + V_w dW_t + \frac{1}{2}V_{ww}d[W,W]_t && (7)\\ &= V_t dt + V_w\left[(r+p\xi_1\sigma_1)w+\pi\right]dt + pw\sigma_1 dZ_t + \frac{1}{2}(pw\sigma_1)^2 V_{ww}dt && (8)\\ &= \left(V_t + V_w \underbrace{\left[(r+p\xi_1\sigma_1)w+\pi\right]}_{\mu_w^p} + \frac{1}{2}\underbrace{(pw\sigma_1)^2}_{(\sigma_w^p)^2} V_{ww}\right)dt + \underbrace{pw\sigma_1 dZ_t}_{\text{local martingale}} && (9)\end{aligned}$$

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

$$\begin{aligned} &V_t + \inf_{p\in\mathbb{P}} \underbrace{\left(\mu_w^p V_w + \frac{1}{2}(\sigma_w^p)^2 V_{ww}\right)}_{L^p V} = 0 && (10)\\ \Rightarrow &V_t + \inf_{p\in\mathbb{P}} L^p V = 0 && (11)\end{aligned}$$

with terminal condition,

$$V(w,t=T) = (w-\frac{\gamma}{2})^2 \tag{12}$$

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

$$U_t = -\left(\mu_w^p U_w + \frac{1}{2}(\sigma_w^p)^2 U_{ww}\right)_{W(t)=w,p(t,w)=p^*(t,w)} \tag{13}$$

with terminal condition,

$$U(w, t = T) = w \tag{14}$$

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

$$V(\hat{w}_0, t = 0) = \mathrm{E}_{p^*}^{t=0}[W_T] - \gamma \mathrm{E}_{p^*}^{t=0}[W_T] + \frac{\gamma^2}{4} \tag{15}$$

$$U(\hat{w}_0, t = 0) = \mathrm{E}_{p^*}^{t=0}[W_T] \tag{16}$$

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,

$$\gamma = \frac{1}{\lambda} + 2\mathrm{E}_{p^*}^{t=0}[W_T] \tag{17}$$

$$\Rightarrow \frac{1}{2\lambda} = \frac{\gamma}{2} - \mathrm{E}_{p^*}^{t=0}[W_T] > 0 \tag{18}$$

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 and Boundary constraints

From the practitioner point of view, 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 paper, to meets such practical circumstances,the authors discussed three cases where restrictions on asset allocations and portfolio wealth are imposed by boundary conditions.

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. 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)$\geq$0) and the investor cannot short sell the risky assets (i.e. p(t,w)$\geq$0) i.e. $\mathbb{D} = [0, +\infty)$ and $\mathbb{P} = [0, +\infty)$.

1. Numerically, we use $\hat{\mathbb{D}} = [0, w_{max}]$ to approximate $\mathbb{D} = [0, +\infty)$.
2. The author make assumption that $p^*(t, w_{max}) \approx 0$, as investor has large enough wealth, he/she prefers the risk-less asset
3. 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

$$
\begin{aligned}
V_\tau(w_{min} = 0, \tau) &= \pi V_w \\
U_\tau(w_{min} = 0, \tau) &= \pi U_w
\end{aligned}
\tag{19}
$$

No Bankruptcy, bounded control In this case, we assume that bankruptcy is prohibited (i.e. W(t)$\geq$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	$[0, w_{\max}]$	$[0, p_{\max}]$

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 (5). Firstly, we need to discretize both the wealth and time step with uniform grid, second, use Policy-iteration algorithm, the grid searching algorithm to find the solution for U and V mentioned above, and finally we assemble the U and V to get the mean variance pair $(Std_{p^*}^{t=0}[Z_T], E_{p^*}^{t=0}[Z_T])$ for this γ. For more technical details please look at Chao.

4.3 Numerical Results

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

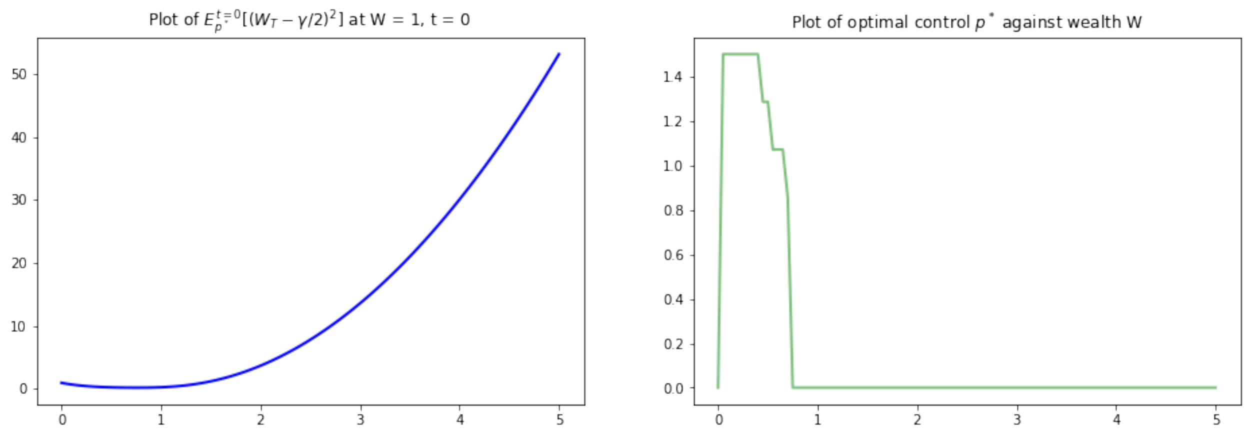

Figure 1: Plot of V, p^* versus wealth ($\gamma = 9.125$)

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 (**?**) where people would invest all money in the non-risky asset when the wealth is relatively large.

4.3.1 Allowing bankruptcy

On top of the fist trial, we implement two algorithms in the paper (**?**) 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 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

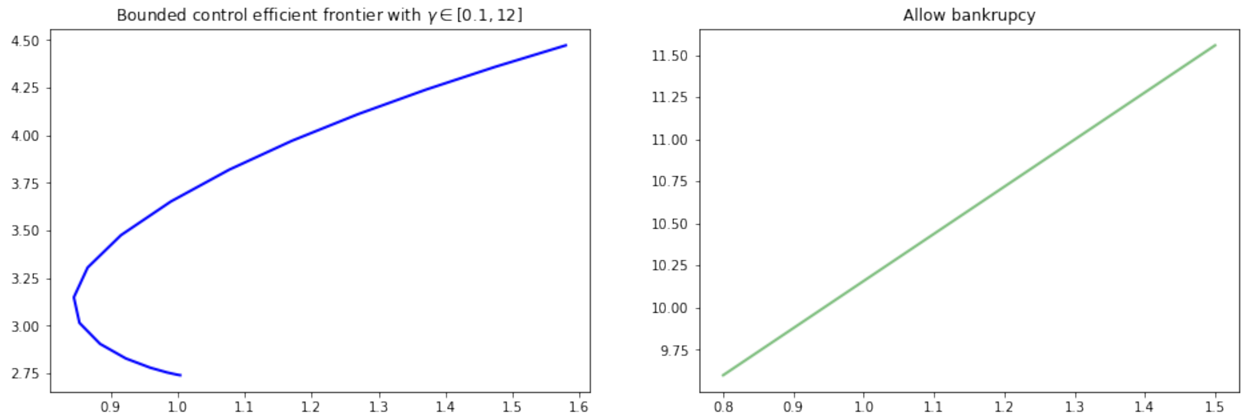

Figure 2: Plot of V, p^* versus wealth

net wealth)

Table 3,4 are reproduced in the context of $\gamma = 9.125$, and $|Wmin| = |Wmax| = 5925$ to

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

approximate the infinity wealth boundary. For this part of the convergence study, we do not see that the $\text{Std}_{p^*}^{t=0}[W_T]$ and $\text{E}_{p^*}^{t=0}[W_T]$ converge monotonically to the analytic solution, but rather in a bumpy fashion. The possible explanation for this mismatch with the results in (**?**) 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	$\mathrm{Std}_{p^*}^{t=0}[W_T]$ [$\star$]	$\mathrm{E}_{p^*}^{t=0}[W_T]$ [$\bigcirc$]	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

$\mathrm{Std}_{p^*}^{t=0}[W_T]$ and $\mathrm{E}_{p^*}^{t=0}[W_T]$ are evaluated at (W = 1, t = 0). Analytic solution is ($\mathrm{Std}_{p^*}^{t=0}[W_T]$, $\mathrm{E}_{p^*}^{t=0}[W_T]$) = (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 a
use constant time steps

Table 6: Convergence Study: Std and Mean

Nodes (W)	Timesteps	Nonlinear iterations	$\mathrm{Std}_{p^*}^{t=0}[W_T]$ [$\star$]	$\mathrm{E}_{p^*}^{t=0}[W_T]$ [$\bigcirc$]	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

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

4.3.2 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 (**?**): 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”.

5 Cong and Cornelis (2016)

5.1 Problem Reformulation

In this supporting paper (1), the author used the same setting as paper (**?**) including portfolio construction, the dynamic mean-variance problem embedded into the stochastic LQ problem (9). The major difference between this paper and the paper in the previous section is that in this case we consider the simulation-based approach. 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

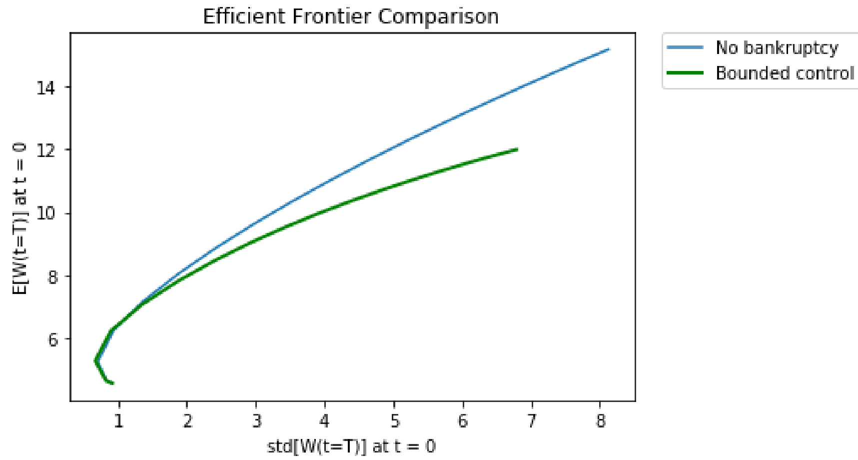

Figure 3: Comparison between Bounded control and No bankruptcy case

and to minimized the investment risk. Theoretically, the investor's investment problem is the same as in (1).

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

$$V_t(W_t) = J_t(W_t) = \min_{\{p_s\}_{s=t}^{T-\Delta t}} \{\mathrm{E}\left[(W_T - \frac{\gamma}{2})^2 \Big| W_t\right]\}, \text{ where } \gamma = \frac{1}{\lambda} + 2\mathrm{E}_{p^*}[W_T|W_t] \tag{20}$$

Adding constaints will remove the smoothness of the value function, which in vein the derivative the derivative-based approach(solving the first order condition). So one has to use a grid searching algorithm similar to the Policy-iteration algorithm proposed by (4). However, casting constraints on the control in the HJB PDE approach as in (4), is not trivial. Specifically, each constraint on the portfolio wealth or the control give rise to a different boundary condition, in other words, a new problem to solve. This is of course not ideal. On the other hand, From the error aspect, using the value function J to transfer the information between two recursive steps will introduce accumulated numerical error as the iteration proceeds.

Based on the above in convenience and concerns, (1) propose two better approaches than the numerical approach proposed by (4), that is, to solve the continuous mean-variance optimization problem via Monte-Carlo simulation. The authors introduced two strategies in general : a multi-stage strategy in a forward fashion and an improved backward recursive dynamic programming calculation based upon 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. In particular, we treat the mean-variance problem from a different angel. Rewrite the original mean-variance problem (1) into the LQ/pre-commitment problem (3). Then the objective becomes to minimize the difference between the final wealth and a predetermined target. The optimization problem at the time t is:

$$J_t^{*pc} := \min_{\{p_s\}_{s=t}^{T-\Delta t}} \{\mathrm{E}\left[(W_T - \frac{\gamma}{2})^2) \Big| W_t\right]\} \tag{21}$$

$$= \min_{\{p_s\}_{s=t}^{T-\Delta t}} \{\mathrm{E}\left[J_{t+\Delta t}^{*pc}(W_{t+\Delta t}) \big| W_t\right]\}, \text{ with } J_t^{*pc}(W_T) = (W_T - \frac{\gamma}{2})^2 \tag{22}$$

where at state (t, W_t), the value function J_t^{*pc} depends on optimal control p at each successive time paths. Clearly observe that equation (22) is the one we discuss in the previous paper (4). Solving this dynamical programming problem backward in time recursively suffers from two problems. First, it is computational inefficient at each time step and back-propagate the value function accumulate the error gradually, so it is numerically inaccurate.

Therefore, reflections on these two issues leads the author in the paper (1) 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,

$$p_t^{*ms} := \arg\min_{p_t}\{\mathrm{E}\left[W_t \cdot (p_t R_t^e + R_f) + \pi \cdot \Delta t - \delta_{t+\Delta t} | \, W_t\right]\} \tag{23}$$

where

$$\delta_t = \frac{\frac{\gamma}{2} - \pi \cdot \Delta t \cdot \frac{1-(R_f)^{\frac{T-t}{\Delta t}}}{1-R_f}}{(R_f)^{\frac{T-t}{\Delta t}}} \tag{24}$$

So in this case, instead of looking at the optimality in the future, the author perform a single stage optimization problem at each discrete 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 (1), 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,

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

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

$$p_t^{*pc}(W_t) = \arg\min_{p_t}\{\mathrm{E}\left[\left(W_t \cdot (p_t R_t^e + R_f) \cdot \prod_{s=t+\Delta t}^{T-\Delta t} (p_s^{*pc} R_s^e + R_f) - \frac{\gamma}{2}\right)^2 \middle| W_t\right]\} \tag{25}$$

where $\{p_s^{*pc}\}_{s=t}^{T-\Delta t}$ denote the optimal allocations at times $s = t + \Delta t, \cdots, T - \Delta t$. The multi-stage optimization problem is,

$$p_t^{*ms}(W_t) = \arg\min_{p_t}\{\mathrm{E}\left[\left(W_t \cdot (p_t R_t^e + R_f) \cdot \prod_{s=t+\Delta t}^{T-\Delta t} (R_f) - \frac{\gamma}{2}\right)^2 \middle| W_t\right]\} \tag{26}$$

In this sub-optimal strategy, clearly, we set $\{p_s^{*pc}\}_{s=t+\Delta t}^{T-\Delta t}$ equal to zero. I.e., we only consider the current moment, regardless the optimality in the future. Thus, this 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: since we changed the objective of the problem, the optimal control obtained from this sub-optimal problem is different from the pre-commitment one, in general. 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 (1) 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 (1) for further interests.

Note One crucial restriction(assumption) in this multi-stage strategy is that there must exist a risk-free asset in the market.

5.3 Stochastic Grid Bundling Method

When the multi-period stochastic optimization problem needs to be solved, one has to deal with the conditional expectations. To accurately and efficiently calculate the conditional expectations, the author adopt the Monte Carlo simulation adn cross-path least squares regression. This is a recently developed method by Cong and Oosterlee, the Stochastic Grid Bundling Method(SGBM). Regression based approach is not a new idea, however, Cong uses a different way to set up the regression step - instead of choosing the regressors and regressands from different time steps, this so called "Regress later" approach use the regressors and regressands from the same time step. The detailed introduction is introduced in the following section.

5.3.1 Regress-Later Technique

Assume we would like to compute the value function $v_t(W_t)$ as in equation(1). Using the Regress-Later approach means that we first approximate the value function $v_{t+\Delta t}(W_{t+\Delta t})$ with basis function formed by wealth $W_{t+\Delta t}$. That is,

$$v_{t+\Delta t}(W_{t+\Delta t}) \approx \sum_{k=0}^{K} \alpha_k \cdot \phi_k(W_{t+\Delta t}) \tag{27}$$

Then substitute into

$$v_t(W_t) = E(v_{t+\Delta t}(W_{t+\Delta t})|W_t, p_t^*) \tag{28}$$

we get,

$$\begin{aligned} v_t(W_t) &= E(\sum_{k=0}^{K} \alpha_k \cdot \phi_k(W_{t+\Delta t})|W_t, p_t^*) && (29) \\ &= \sum_{k=0}^{K} \alpha_k \cdot E(\phi_k(W_{t+\Delta t})|W_t, p_t^*) && (30) \end{aligned}$$

From equation (30), we have to know the conditional expectations of the basis functions, so as to implement the Regress-Later approach.

5.3.2 Bundling Approach

"Bundling" approach proposed by the author provides additional information in the regression step. We again consider the problem of computing the conditional expectations shown in Equation (28). Assume that $X_{t+\Delta t}\Phi_{t+\Delta t}$, where $\Phi_{t+\Delta t}$ denotes a sub-domain of $\mathbb{R}$. Based on the information about the sub-domains, using the Regress-Later technique, we can approximate $v_{t+\Delta t}(W_{t+\Delta t})$ by,

$$v_{t+\Delta t}(W_{t+\Delta t}) \approx \sum_{k=0}^{K} \hat{\alpha}_k \cdot \phi_k(W_{t+\Delta t}) \tag{31}$$

Here we still consider the basis functions $\{\phi_k(W_{t+\Delta t})\}_{k=0}^K$ as used in Equation (27). Since we have the information that $W_{t+\Delta t}$ belongs to a specified sub-domain, the coefficients $\{\hat{\alpha}_k\}_{k=0}^K$ may be different from $\{\alpha_k\}_{k=0}^K$. Approximating value functions in a sub-domain usually requires fewer basis functions for achieving satisfactory accuracy.

5.4 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 (**?**), the domain $D_{t+\Delta t}$ can be approximated as

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

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 $f_{t+\Delta t}^b(\cdot)$ using up to second order basis functions to regress the continuation values $J_{t+\Delta t}^b(i)_{i=1}^{N_b}$ on the wealth values $W_{t+\Delta t}^b(i)_{i=1}^{N_b}$ at time $t + \Delta t$.

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

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

3. Calculate the old continuation value $\tilde{J}_t^B(i)$. Update the allocation $x_t^b(i)v_{i=1}^{N_b}$ based on the condition if the new continuation value,$\hat{J}_t^b(i)$, is less than the original one, $\tilde{J}_t^b(i)$.

4. Once the updated allocations are obtained, again by regression we can calculate the "updated" continuation value $J_t^b(i)_{i=1}^{N_B}$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.

Note about the implementation: Assume that an optimal strategy for W_T is p_T^*. Then by taking into account these two values, one can estimate the projected wealth level $W_(T+1)$. Assume that we already know the value function, V/(J), at time T+1, we can then get $V_(T+1)$ for each individual path. Here, we can apply the "Regress-Later" idea by regressing $V_(T+1)$ on the basis functions formed by $W_(T+1)$. Say, we get a formula like: $V_(T+1) = W_(T+1)+W_(^2T+1)$. By further taking into account that p_T^* can be written as an analytic formula of W_T, and $E[W_(T+1)]$ should be a function of W_T and p_T^*, we can estimate $E[V_(T+1)|W_T]$ in the format of "$a+b*W_T+c*W_T^2$" (some math should be done carefully here). In other words, this gives the continuation value the function for any paths whose wealth levels are located in a certain bundle. – 2(c).

5.5 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.

5.5.1 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)

$$0 \le p_t \le 1 + \frac{\pi \Delta t}{W_t \cdot R_f} \tag{32}$$

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.

5.5.2 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 (1), 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_t^e are $R_t^{e,\alpha}$ and $R_t^{e,1-\alpha}$ correspondingly. It follows by paper (1), with certainty $1-2\alpha\%$ and the constraint from below, The bound for the control p_t can be computed as,

$$\frac{-\pi\Delta t - W_t \cdot R_f}{W_t \cdot R_t^{e,1-\alpha}} \leq p_t \leq \frac{-\pi\Delta t - W_t \cdot R_f}{W_t \cdot R_t^{e,\alpha}} \tag{33}$$

5.5.3 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.

5.6 Numerical Results

5.6.1 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.

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		

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

5.6.2 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 (1), as we find out that the simple five thousands 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 (1). We also tried to increase the number of paths in the simulation experiment, this results in smoother frontier indeed.

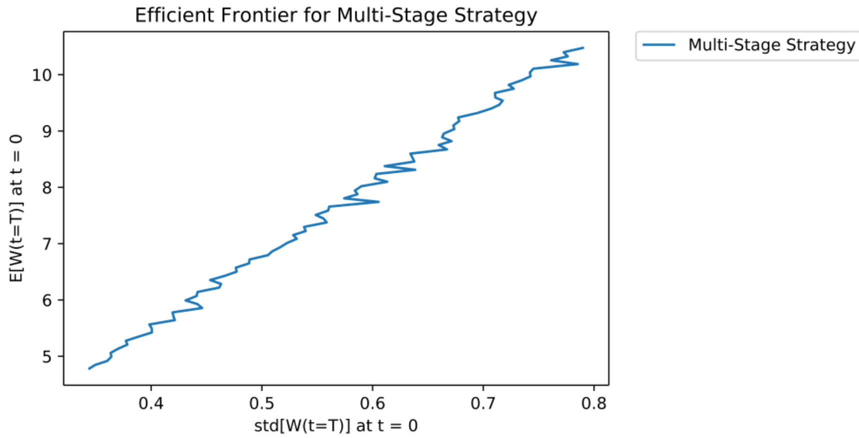

Figure 4: Efficient Frontier for Multi-stage Strategy

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 (1).

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.

6 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.

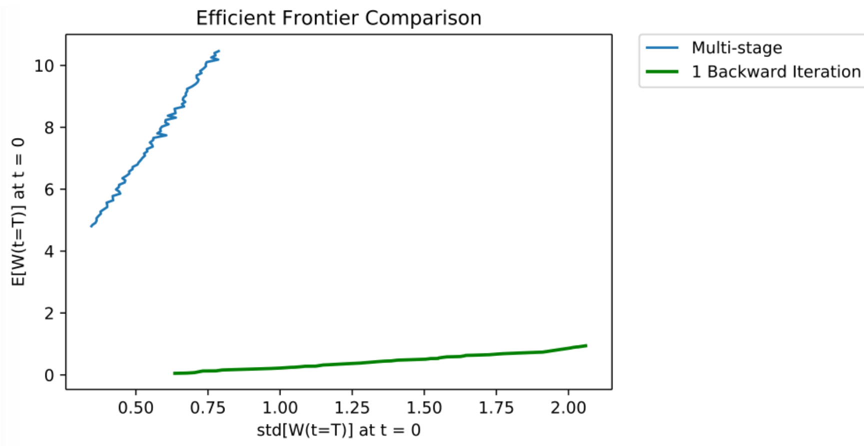

Figure 5: Comparison between Multi-Stage and Backward recursion approach

7 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 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 (**?**), the author made the assumption here for large w, the optimal control p ≈ 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

$$V_t + (rw + \pi)V_w = 0 \tag{33}$$

with terminal condition (15),

$$V(w, t = T) = (w - \frac{\gamma}{2})^2 \tag{34}$$

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

$$V(w, t = T - \tau) = H_1(\tau)w^2 + H_2(\tau)w + H_3(\tau) \tag{35}$$
$$\Rightarrow V_t = \frac{\partial V}{\partial \tau} \cdot \frac{\partial \tau}{\partial t} = -\frac{\partial V}{\partial \tau} = -H_1^{'}(\tau)w^2 - H_2^{'}(\tau)w - H_3^{'}(\tau) \tag{36}$$
$$\Rightarrow V_w = 2H_1(\tau)w + H_2(\tau) \tag{37}$$

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

$$V_t(\tau = T - t = 0) = w^2 - \gamma w + \frac{\gamma^2}{4} = H_1(0)w^2 + H_2(0)w + H_3(0) \tag{38}$$

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 $W(p, t = T - \tau = 0) = \hat{w}_0$)

$$\begin{cases} H_1^{'}(\tau) = 2rH_1(\tau) \\ H_2^{'}(\tau) = 2\pi H_1(\tau) + rH_2(\tau) \\ H_3^{'}(\tau) = \pi H_2(\tau) \\ H_1(0) = \hat{w}_0^2 \\ H_2(0) = -\hat{w}_0\gamma \\ H_3(0) = \frac{\gamma^2}{4} \end{cases} \tag{39}$$

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,

$$V(w, t = T - \tau) = H_1(\tau)w^2 + H_2(\tau)w + H_3(\tau) \tag{40}$$

where $\tau = T - t$ and,

$$\begin{cases} H_1(\tau) = \mathrm{e}^{2r\tau} \\ H_2(\tau) = -w_0\gamma\mathrm{e}^{r\tau} + \frac{2\pi}{r}\mathrm{e}^{r\tau}(\mathrm{e}^{r\tau} - 1) \\ H_3(\tau) = \frac{\gamma^2}{4} - \frac{w_0\pi\gamma(\mathrm{e}^{r\tau}-1)}{r} + \frac{\pi^2(\mathrm{e}^{r\tau}-1)^2}{r^2} \end{cases} \tag{41}$$

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 $p*$,

$$p^*(t, w) = -\frac{\xi_1}{\sigma_1 w}[w - \frac{w_0\gamma}{2}\mathrm{e}^{-r(T-t)} + \frac{\pi}{r}\mathrm{e}^{-r(T-t)}(\mathrm{e}^{-r(T-t)} - 1)], \forall t \in [0, T] \tag{42}$$

For γ part Observed that the only unknown part for $p*$'s formula is the γ part for large w. According to (2), 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,

$$dW_t = [(r + p^* \xi_1 \sigma_1)W_t + \pi]\, dt + p^* W_t \sigma_1 dZ_t \tag{43}$$

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 $\frac{\partial f}{\partial t} = 0$, $\frac{\partial f}{\partial W} = 2W_t, \frac{\partial^2 f}{\partial W^2} = 2$)

$$dW_t^2 = \left(2[\mu_w^{p*} W_t] + (\sigma_w^{p*})^2\right) dt + 2\sigma_w^{p*} W_t dZ_t \tag{44}$$

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 (2), 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,

$$\mathrm{E}_{p^*}[W_T] = (w_0 + \frac{\pi}{r})\mathrm{e}^{-(\xi_1^2 - r)T} + \frac{\gamma}{2}(1 - \mathrm{e}^{-\xi_1^2 T}) - \frac{\pi}{r}\mathrm{e}^{-\xi_1^2 T} \tag{45}$$

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

$$\mathrm{E}_{p^*}[W_T^2] = (w_0 + \frac{\pi}{r})^2\mathrm{e}^{-(\xi_1^2 - 2r)T} + \frac{\gamma}{2}(1 - \mathrm{e}^{-\xi_1^2 T}) - \frac{2\pi}{r}(w_0 + \frac{\pi}{r})\mathrm{e}^{-(\xi_1^2 - r)T} + \frac{\pi^2}{r^2}\mathrm{e}^{-\xi_1^2 T} \tag{46}$$

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

$$\begin{aligned}\gamma &= \frac{1}{\lambda} + 2((w_0 + \frac{\pi}{r})\mathrm{e}^{-(\xi_1^2 - r)T} + \frac{\gamma}{2}(1 - \mathrm{e}^{-\xi_1^2 T}) - \frac{\pi}{r}\mathrm{e}^{-\xi_1^2 T}) \\ \Rightarrow \gamma &= \mathrm{e}^{\xi_1^2 T}\frac{1}{\lambda} + 2(w_0 + \frac{\pi}{r})\mathrm{e}^{rT} - \frac{2\pi}{r}\end{aligned} \tag{47}$$

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

$$p^*(t,w) = -\frac{\xi_1}{\sigma_1 w}[w - (\hat{w}_0 e^{rt} + \frac{\pi}{r}(e^{rt} - 1)) - \frac{e^{-r(T-t) + \xi_1^2 T}}{2\lambda}], \forall t \in [0,T] \tag{48}$$

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

$$\mathrm{E}_{p^*}^{t=0}[W_T] = w_0\mathrm{e}^{rT} + \pi\frac{\mathrm{e}^{rT} - 1}{r} + \frac{\mathrm{e}^{\xi_1^2 T} - 1}{2\lambda} \tag{49}$$

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,

$$\mathrm{Var}_{p^*}^{t=0}[W_T] = \frac{e^{\xi_1^2 T} - 1}{4\lambda^2} \tag{50}$$

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,

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

$$\mathrm{E}_{p^*}^{t=0}[W_T] = w_0 \mathrm{e}^{rT} + \pi \frac{\mathrm{e}^{rT} - 1}{r} + \frac{\mathrm{e}^{\xi_1^2 T} - 1}{2\lambda} = w_0 \mathrm{e}^{rT} + \pi \frac{\mathrm{e}^{rT} - 1}{r} + \sqrt{\mathrm{e}^{\xi_1^2 T} - 1}\,\mathrm{Std}(W_T) \tag{51}$$

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

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

- **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 $\{W_t(i)\}_{i=1}^{N}, t = 0, \ldots, T$. 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 $t = T - \Delta t, \ldots, \Delta t, 0$.
- **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 $\{\hat{x}_t^b(i)\}_{i=1}^{N_B}$.

(1),

  - 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] F. Cong and C. W. Oosterlee. Multi-period mean–variance portfolio optimization based on monte-carlo simulation. *Journal of Economic Dynamics and Control*, 64:23–38, 2016.

[2] B. Højgaard, E. Vigna, et al. Mean-variance portfolio selection and efficient frontier for defined contribution pension schemes. 2007.

[3] D. Li and W.-L. Ng. Optimal dynamic portfolio selection: Multiperiod mean-variance formulation. *Mathematical finance*, 10(3):387–406, 2000.

[4] J. Wang and P. A. Forsyth. Numerical solution of the hamilton–jacobi–bellman formulation for continuous time mean variance asset allocation. *Journal of Economic Dynamics and control*, 34(2):207–230, 2010.

[5] Z. Wang, J. Xia, and L. Zhang. Optimal investment for an insurer: The martingale approach. *Insurance: Mathematics and Economics*, 40(2):322–334, 2007.

[6] X. Y. Zhou and D. Li. Continuous-time mean-variance portfolio selection: A stochastic lq framework. *Applied Mathematics and Optimization*, 42(1):19–33, 2000.