金融机器学习课程报告

基于稳健风险感知强化学习框架的投资组合优化问题求解

学院 中国人民大学统计学院

小组 第五小组

姓名及学号

姜颖昕 2024103987

韩艳栋 2023103283

郭镇涛 2024103985

杨哲 2024103999

吴晶晶 2023101730

2025 年 1 月 20 日

摘　要

本报告基于 Jaimungal 等人提出的基于稳健优化的风险感知强化学习（RL）方法 [3]，允许代理人考虑各种风险-回报特征，并应用秩相依期望效用（RDEU）来评估策略的价值。为了在模型不确定性下稳健地优化策略，我们不是直接根据原始策略来对其评估，而是讨论它在 Wasserstein 球内的最坏可能分布，寻找最坏情况下的最优策略。因此，我们的问题表述可以被视为一个代理 (Agent) 选择策略（外部问题），而对手 (Adversary) 则试图扭曲该策略的性能（内部问题）。报告中为内部和外部问题提出了明确的策略梯度公式，并给出了对应伪代码。

为验证该框架的适用性，我们进一步提出了稳健投资组合优化问题，希望利用该框架解决现实中的投资组合优化问题。为此，我们引入近十年（2015 年 1 月 1 日至 2025 年 1 月 1 日）10 个涵盖股票、债券、货币、期货等不同类型资产进行实证分析，采用 PCA 方法分解资产回报因子，然后通过上述稳健强化学习框架进行策略学习，以 RDEU 为评价指标，并考虑 Wasserstein 距离大小带来的影响。结果表明当 Wasserstein 距离较大时，CVaR 正向增大，UTE 减小，投资者倾向于稳健投资策略，反之亦然。

目录

1 引言 **1**

2 稳健随机优化问题框架 **1**

2.1 秩相依效用（Rank Dependent Expected Utility） 1

2.2 Wasserstein Distance 2

2.3 稳健随机优化问题框架 2

3 基于强化学习策略梯度算法的问题求解 **2**

3.1 内部问题求解 3

3.2 外部问题求解 4

3.3 算法设计 4

4 稳健投资组合优化问题 **5**

4.1 基于稳健随机优化框架下的问题描述 5

4.2 问题具体设定 5

4.3 实证数据的收集与处理方法 6

4.4 风险测度设定 8

5 实证结果 **8**

6 结论与讨论 **11**

6.1 结论 11

6.2 讨论 11

7 附录：团队成员信息及分工 **13**

1 引言

在金融数学、经济学和工程的诸多领域，经常需要解决包含不确定性的优化问题。这些问题通常可以通过随机优化的方法来形式化，其中决策者（或代理）必须在给定的概率模型下选择最优控制策略。

然而，现实世界是复杂多变的，任何模型都只能作为现实的近似。因此，理解并量化模型不确定性对决策的影响，是制定有效策略的关键。在金融市场中，投资者面临的一个主要挑战是如何在追求收益的同时管理风险。传统的期望效用理论提供了一种在风险和回报之间权衡的方法，但它并不适用于所有情况 [4],[2]。例如，投资者可能对概率的主观看法与客观概率不同，或者他们可能更关心避免损失而非仅仅最大化期望收益。因此，引入了如条件风险价值（CVaR）等更为复杂的风险度量方法。

稳健优化（Robust Optimization）提供了一种框架，允许决策者在不确定性下选择策略，即使在最坏情况下也能保持其性能 [1]。在稳健优化问题中，代理不仅要考虑其策略在预期模型下的表现，还要考虑在一系列可能的模型中的表现。这种方法特别适用于那些对模型不确定性特别敏感的领域，如金融市场的资产配置 [5]。

本文将基于 Jaimungal 等人提出的稳健强化学习框架 [3]，进一步结合现实中的收益率数据得到最终的实证结果。在第二部分，我们将介绍稳健随机优化框架，包括核心的 RDEU 概念、Wasserstein 距离，并提出我们解决的目标问题；在第三部分，我们将给出强化学习的策略梯度算法及其伪代码；在第四部分，我们将介绍利用该框架求解的稳健随机优化问题，并说明该问题的具体设定及实际数据；第五部分，我们将给出实际资产数据得到的结果；最后在第六部分，我们将得到最后的结论与讨论。

2 稳健随机优化问题框架

本部分将对稳健随机优化问题进行概述，分别介绍度量随机变量风险的秩相依效用、描述分布不确定性的 Wasserstein 距离，并基于两者构建稳健随即优化问题的框架。

2.1 秩相依效用（Rank Dependent Expected Utility）

秩相依效用框架由 Yaari 提出 [7], 其定义如下

定义 2.1. *随机变量的秩相依期望效用可以通过 Choquet 积分定义为：*

$$\mathcal{R}_g^U[Y] := \int_{-\infty}^{0} (1 - g(P(U(Y) > y)))\, dy - \int_{0}^{+\infty} g(P(U(Y) > y))\, dy \tag{1}$$

其中扭曲函数（distortion function）$g : [0,1] \to [0,1]$ 是一个递增函数，满足 $g(0) = 0$ 和 $g(1) = 1$，而 U 是一个非递减的凹效用函数。我们假设 U 在几乎所有地方都是可微的。

上述定义中，假设正结果对应于收益，而负结果对应于损失。RDEU 框架包含了扭曲风险度量类，例如当 $U(x) = x$ 时，上述结果可包括条件风险价值（CVaR）。此外，当

$g(x) = x$ 时，它包括期望效用框架，此时 $\mathcal{R}_g^U[Y] = -\mathbb{E}[U(Y)]$。在本文中，我们通常将随机变量的 RDEU 称为随机变量的风险。

2.2 Wasserstein Distance

定义 2.2. 随机变量 X 与 Y 间的 p 阶 *Wasserstein* 距离被定义为：

$$d_p[X,Y] := \inf_{\chi \in \Pi(F_X,F_Y)} \left(\int_{\mathbb{R}^2} |x-y|^p \chi(dx,dy) \right)^{\frac{1}{p}} \tag{2}$$

其中 $\Pi(F_X,F_Y)$ 是所有边际分布为 F_X, F_Y 的联合概率测度。

p 阶 Wasserstein 距离在概率测度空间上定义了一种度量，可以用来表示随机变量 X、Y 的分布的相近程度。

2.3 稳健随机优化问题框架

基于上述秩相依效用与 Wasserstein 距离的表述，可以得到一种稳健随机优化问题的框架。问题分析如下：

当代理人采取行动 $\phi \in \Phi$ 时将产生一个随机变量 X^ϕ，代理人希望关于该随机变量的风险最小化，即 $R_g[X^\phi]$ 最小。然而由于模型和分布的不确定性，代理人将会选择使得 X^ϕ 最坏情况风险最小的行动，θ 属于与代理人行动 ϕ 有关的不确定集 ν_ϕ，ν_ϕ 可通过 Wasserstein 距离来描述 X^ϕ 与 X^θ 之间的相似性。

根据上述分析，可以得到如下稳健随机优化问题框架：

$$\inf_{\phi \in \Phi} \sup_{\theta \in \nu_\phi} R_g[X^\theta],$$
$$\text{where } \nu_\phi := \{\theta \in \Theta : d_p[X^\theta, X^\phi] \le \varepsilon\}$$

3 基于强化学习策略梯度算法的问题求解

在强化学习里的策略梯度方法 [6] 提供了一种通过在值函数梯度的方向上采取步骤来逐步改进策略/动作的序列的方法，其中梯度是相对于策略参数求得的。

在稳健随机优化框架中，我们的优化目标为

$$\inf_{\phi \in \varphi} \sup_{\theta \in \nu_\phi} R_g^U[X^\theta]$$
$$\nu_\phi := \{\theta \in \nu : d_p[X^\theta, X^\phi] \le \varepsilon\}$$

而该目标可以拆解为内部问题 $\sup_{\theta \in \nu_\phi} R_g^U[X^\theta]$ 和外部问题 $\inf_{\phi \in \varphi} \sup_{\theta \in \nu_\phi} R_g^U[X^\theta]$，在本节中，我们将基于内部问题和外部问题推导出相应公式，用于稳健随机优化问题的策略梯度更新规则，从而实现对 ϕ 和 θ 的优化。

3.1 内部问题求解

在内部问题中，可以将 X^ϕ 视为一个固定的随机变量，进而使用增广拉格朗日法，把不确定集中关于 Wasserstein 距离的限制纳入公式，则增广拉格朗日函数可以表示成

$$L[\theta,\phi] = R_g^U[X^\theta] + \lambda c[X^\theta, X^\phi] + \frac{\mu}{2}(c[X^\theta, X^\phi])^2 \tag{3}$$

其中 $c[X^\theta, X^\phi] := \left((d_p[X^\theta, X^\phi])^p - \varepsilon^p\right)_+$，$\lambda$ 为拉格朗日乘子，μ 为惩罚项。在增广拉格朗日算法中，固定 λ 和 μ，利用随机梯度下降使得 $L[\theta,\phi]$ 最小，每经过 N 轮循环后增大 λ 和 μ，即可实现该问题的优化。

下面推导该内部问题的梯度公式。令 X_c^ϕ 为 X^ϕ 秩排列后的结果，当扭曲函数 g 左可导时，可以得到内部问题的梯度公式。

命题 3.1. 内部梯度公式

$$\nabla_\theta L[\theta,\phi] = E\left[\left\{U'(X^\theta)\gamma(F_\theta(X^\theta)) - p\Lambda|X^\theta - X_c^\phi|^{p-1} sgn(X^\theta - X_c^\phi)\right\}\frac{\nabla_\theta F_\theta(x)|_{x=X^\theta}}{f_\theta(X^\theta)}\right]$$

其中 $\gamma:(0,1)\to\mathbb{R}$ 为 $\gamma(u) := \partial_- g(x)|_{x=1-u}$，常数 $\Lambda := (\lambda+\mu c[X^\theta, X^\phi]) I_{d_p[X^\theta,X^\phi]>\varepsilon}$.

该公式中的 $\nabla_\theta F_\theta(x)$ 无法直接得出，因此需要核密度估计法。利用 (X^θ, X_c^ϕ) 的同单调小批量样本 $\{(x_\phi^{(1)}, x_\theta^{(1)}), \ldots, (x_\phi^{(N)}, x_\theta^{(N)})\}$，可以得到 $F_\theta(x)$ 的核密度估计 $\hat{F}_\theta(x)$，表示为

$$\hat{F}_\theta(x) = \frac{1}{N}\sum_{i=1}^{N}\Phi(x - x_\theta^{(i)})$$

其中 Φ 为可能的分布核函数，一般使用高斯分布。根据上式可以得到：

$$\nabla_\theta F_\theta(x) = -\frac{1}{N}\sum_{i=1}^{N}\Phi'(x - x_\theta^{(i)})\nabla_\theta x_\theta^{(i)}$$

其中 Φ' 为选择的核函数对应的密度函数。将该估计代入内部问题的梯度公式，可以得到用于优化算法的最终内部问题梯度公式：

$$\begin{aligned}\nabla_\theta L[\theta,\phi] \approx -\frac{1}{N}\sum_{i=1}^{N}\Bigg[\Big(&U'(x_\theta^{(i)})\gamma(\hat{F}_\theta(x_\theta^{(i)}))\\ &-p\Lambda|x_\theta^{(i)} - x_{\phi,c}^{(i)}|^{p-1}\mathrm{sgn}(x_\theta^{(i)} - x_{\phi,c}^{(i)})\Big)\frac{\sum_{j=1}^{N}\Phi'(x_\theta^{(i)} - x_\theta^{(j)})\nabla_\theta x_\theta^{(j)}}{\sum_{k=1}^{N}\Phi'(x_\theta^{(i)} - x_\theta^{(k)})}\Bigg]\end{aligned}$$

其中 Wasserstein 距离的公式通过小批量样本估计量得到

$$\left(\frac{1}{N}\sum_{i=1}^{N}|x_\theta^{(i)} - x_{\phi,c}^{(i)}|\right)^{\frac{1}{p}}.$$

3.2 外部问题求解

基于外部问题，同样使用增广拉格朗日法得到目标函数，进而对 θ 求导。由于求导要求，需要假设 X^θ 的生成方式，我们假设 $X^\theta = H(X^\phi, Y)$，其中 Y 是产生 X^θ 的另一个随机影响。

根据与内部问题相同的思路，可以得到外部问题的梯度公式。

命题 3.2. 外部梯度公式

$$\begin{aligned}\nabla_\phi L[\theta,\phi] =& \mathbb{E}\left[U'(X^\theta)\gamma(F_\theta(X^\theta))\frac{\nabla_\phi F_\theta(x)|_{x=X^\theta}}{f_\theta(X^\theta)}\right. \\ &\left. -p\Lambda|X^\theta - X_c^\phi|^{p-1}sgn(X^\theta - X_c^\phi)\left(\frac{\nabla_\phi F_\theta(x)|_{x=X^\theta}}{f_\theta(X^\theta)} + \frac{\nabla_\phi G_\phi(x)|_{x=X^\phi}}{g_\phi(X^\phi)}\right)\right]\end{aligned}$$

再次利用核密度估计法，得到用于优化的最终外部问题公式为：

$$\begin{aligned}\nabla_\phi L[\theta,\phi] \approx& -\frac{1}{N}\sum_{i=1}^{N}\left[U'(x_\theta^{(i)})\gamma(\hat{F}_\theta(x_\theta^{(i)}))\frac{\sum_{j=1}^N \Phi'(x_\theta^{(i)} - x_\theta^{(j)})\nabla_\phi x_\theta^{(j)}}{\sum_{k=1}^N \Phi'(x_\theta^{(i)} - x_\theta^{(k)})}\right. \\ &\left. -p\Lambda|x_\theta^{(i)} - x_{\phi,c}^{(i)}|^{p-1}\mathrm{sgn}(x_\theta^{(i)} - x_{\phi,c}^{(i)})\left(\frac{\sum_{j=1}^N \Phi'(x_\theta^{(i)} - x_\theta^{(j)})\nabla_\phi x_\theta^{(j)}}{\sum_{k=1}^N \Phi'(x_\theta^{(i)} - x_{\phi,c}^{(k)})}\right)\right]\end{aligned}$$

为计算简便，F_θ 与 G_ϕ 的核密度估计核函数相同，$\nabla_\phi x_\theta, \nabla_\phi x_{\phi,c}$ 可通过 $X^\theta = H(X^\phi, Y)$ 和神经网络的反向传播算法得到。

3.3 算法设计

通过上述分析，利用人工神经网络方法，可以得到求解一般稳健随机优化问题的算法。

Algorithm 1 强化学习策略梯度算法下稳健随机优化类问题的算法

```
初始化神经网络 θ, ϕ;
初始化拉格朗日乘子 λ = 1, μ = 10, α = 1.5;
for i = 1 to M_outer do
    模拟生成 X^ϕ 小批次样本;
    for j = 1 to M_inner do
        模拟生成 X^θ 的小批次样本（基于外层循环中固定的 X^ϕ）；
        利用内部问题公式计算 ∇_θ L[θ, ϕ];
        用 ADAM 算法更新神经网络 θ 的参数;
        if (j + 1)%N_Lagrange = 0 then
            更新拉格朗日乘子 λ ← λ + μc(θ*), μ ← αμ;
        end if
    end for
    重复内部循环至 d_p[X^θ, X^ϕ] ≤ ε 且 R_γ^U[X^θ] 在过去的 100 次循环中没有继续增加;
    基于 X^ϕ 与更新后的内部网络 θ 生成 X^θ 的小批次样本;
    利用外部问题公式计算 ∇_ϕ L[θ, ϕ];
    用 ADAM 算法更新神经网络 ϕ 的参数;
    重复外部循环至 R_γ^U[X^θ] 在过去的 100 次循环中没有继续增加;
end for
```

上述算法中，M_{outer} 和 M_{inner} 分别为内部循环和外部循环的轮数，N_{Lagrange} 为设定的拉格朗日乘子更新的间隔轮数。

4 稳健投资组合优化问题

4.1 基于稳健随机优化框架下的问题描述

假设 φ 是 d 维概率单纯形，对于 $\varphi \in \varphi$，令 $X^\varphi = \varphi^T X$，其中 X 为 d 维向量，代表 d 个交易资产的回报，即 $X = (X_1, \ldots, X_d)$。因此 X^φ 为投资组合的总回报。而 $X^\varphi = H_\theta(X^\varphi)$，其中 $H_\theta(\cdot)$ 为以 θ 为参数的人工神经网络。

基于上述描述，内部问题的目标是在 Wasserstein 球内寻找所有可能的分布函数，以稳健评估 X^φ（即投资组合）的风险。外部问题的目标则是寻找一个最优稳健投资组合。

4.2 问题具体设定

假设市场中有 d 种资产，它们的回报由系统性因子 ζ 和特质性因子 $Z_i, i \in D := \{1, \ldots, d\}$ 组成。资产的对应回报为 $R_i = \zeta + Z_i$，总投资回报为 $X^\varphi = \varphi^T R$。外部策略

π^{φ} 通过一个输入为 θ，输出为 φ 的神经网络得到（激活函数选取 softmax 函数以使得 $\sum_{i\in D}\varphi_i = 1$），选取 Wasserstein 距离的阶数为 1 阶。本报告采用上述假设，选取了资产的实证数据进行资产回报的模拟。

4.3 实证数据的收集与处理方法

针对强化学习在稳健投资组合优化问题上的应用，原文章提出的系统性因子和特质性因子都给出了假定的分布，如 $\zeta \sim N(0, 0.02^2), Z_i \sim N(0.03i, (0.025i)^2), i = 1, 2, ..., d$，本文力求结合真实的资产数据给出更真实的因子模型模拟资产路径。

为保证数据的真实性和时效性，实证部分选取了 10 个涵盖股票、债券、货币、期货等不同类型的资产，资产数据的时间跨度为近十年即 2015 年 1 月 1 日至 2025 年 1 月 1 日，数据频率为日度，数据来源为 Yahoo Finance。选取的 10 种资产分别为：股票 AAPL、MSFT、GOOGL、AMZN、NVDA，债券 TLT，货币 GLD，期货 CL=F，以及加密货币 BTC-USD，选取的资产回报使用最常用的对数收益率的形式，即

$$r_t = \log\left(\frac{P_t}{P_{t-1}}\right),$$

其中 P_t 是时间 t 的资产价格（选取收盘价），r_t 是时间 t 的对数收益率即资产回报。为了提取资产回报中的系统性因子和特质性因子，本文采用主成分分析（PCA）方法。

对每日对数收益率矩阵 $\mathbf{R}$ 进行主成分分析，设 $\mathbf{R}$ 为 $n \times d$ 的矩阵，其中 n 为时间序列长度，d 为资产数量。PCA 的目标是将 $\mathbf{R}$ 分解为系统性因子和特质性因子。由于

$$\mathbf{R} = \mathbf{U}\Sigma\mathbf{V}^T$$

其中，$\mathbf{U}$ 是左奇异矩阵，Σ 是对角矩阵，$\mathbf{V}$ 是右奇异矩阵。$\mathbf{V}$ 的列向量即为主成分。我们将第一个主成分作为系统性因子。设 $\mathbf{v}_1$ 为第一个主成分向量，则系统性因子 ζ 可以表示为：

$$\zeta = \mathbf{R}\mathbf{v}_1,$$

这样解释真实数据的系统性因子在本文看来是合理的，因为 PCA 的第一主成分 $\mathbf{v}_1$ 是数据中方差最大的方向。根据 PCA 的原理，第一个主成分是整个数据集（在此为每日对数收益率）的最大变异来源。也就是说，$\mathbf{v}_1$ 是数据中最大程度上解释了收益率变化的方向，反映了影响所有资产共同变化的主要因素。

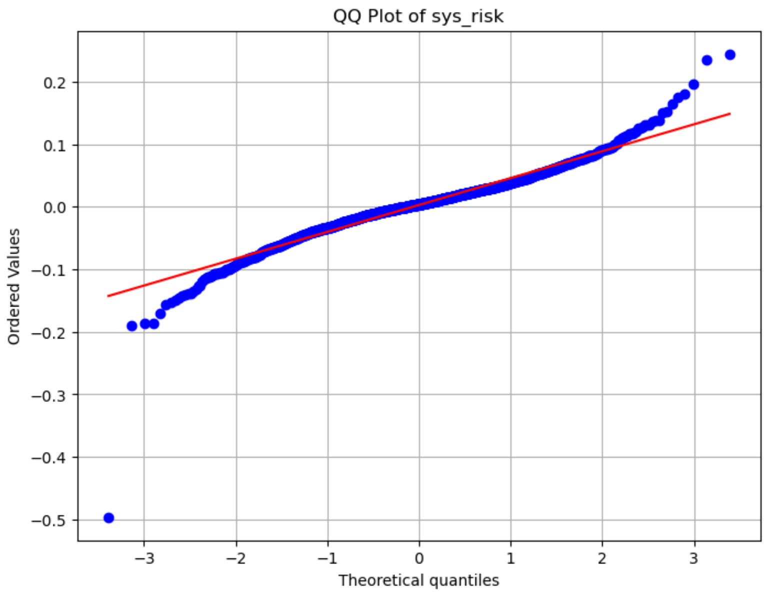
在实际应用中，第一主成分通常与市场整体的变动（例如市场指数的变动）高度相关，因此它经常被解读为系统性风险因子。它代表了影响大多数资产（或所有资产）的共同因素，反映了市场整体的趋势。

而特质性因子是每日对数收益率减去系统性因子的部分。设 $\mathbf{Z}$ 为特质性因子矩阵，则有：

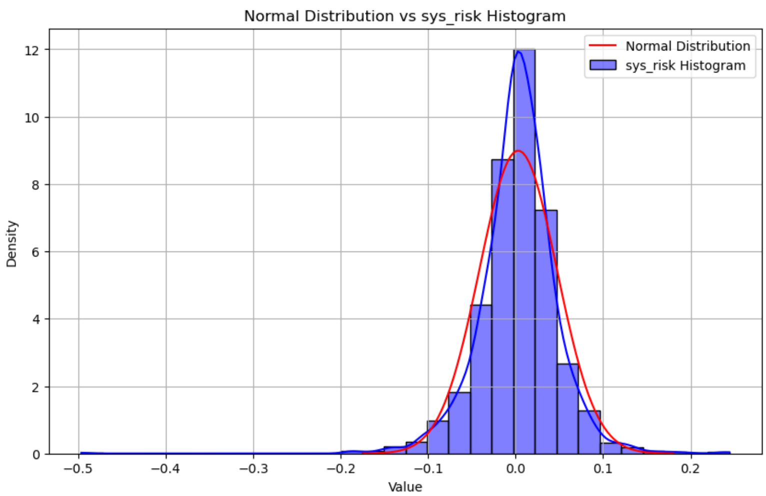
$$\mathbf{Z} = \mathbf{R} - \zeta\mathbf{v}_1^T.$$

通过上述方法我们成功将资产回报分解为系统性因子和特质性因子，为后续的稳健投资组合优化问题提供了数据基础，同时本文的处理方法相比参考文章更加真实可靠符合市场，对强化学习在稳健投资组合问题上的应用更具有实际价值。

由于强化学习的对抗过程需要对投资组合不停进行迭代，运用得到的系统性因子和特质性因子模拟资产路径是十分必要的。本文首先考虑了如果以系统性因子和特质性因子分别服从正态分布和多元正态分布从中抽取样本进行资产路径模拟的方法，但在系统性因子 Q-Q 图1(a)以及系统性因子分布与正态分布对比的图1(b)中发现，系统性因子并不服从正态分布，因此我们采用了基于核密度估计分布抽取样本作为模拟资产路径的方法。



(a) 系统性因子 Q-Q 图



(b) 系统性因子分布与正态分布对比

4.4 风险测度设定

如前部分所说，我们使用 RDEU 框架度量风险，在该问题中，我们选择 $\alpha-\beta$ 风险测度，即

$$\gamma(u)=\frac{1}{\eta}\left(p\mathbb{I}_{\{u\leq\alpha\}}+(1-p)\mathbb{I}_{\{u>\beta\}}\right),$$

其中 $\eta=p\mathbb{I}_{\{u\leq\alpha\}}+(1-p)\mathbb{I}_{\{u>\beta\}}$，$0<\alpha\leq\beta<1$，而 $p\in[0,1]$。

这样设置的函数为 U 型，并且兼顾了投资者的风险厌恶和风险偏好。

- 当 $p=1$ 时，风险测度即为 α 下的 CVaR；
- 当 $p=0$ 时，风险测度即为 β 下的上尾期望（Upper Tail Expectation）。
- 当 $p>0.5$ 时，风险测度更强调损失的度量。
- 当 $p<0.5$ 时，风险测度更强调收益的度量。

5 实证结果

实证方面，我们采用上文提到的 $\alpha-\beta$ 风险测度，选取 $\alpha=0.1,\beta=0.9$，并且 $p=0.75$，即更强调损失的度量。对于强化学习的框架，分别设定内层训练迭代 500 次，外层训练迭代 81 次，共训练 40500 次，外层训练每迭代 40 次生成参数、效用函数以及资产组合权重的图像。分别对 Wasserstein 距离从 $\epsilon=10^{-3},10^{-2},10^{-1},1$ 扭曲投资组合。

下面给出部分训练结果，完整的训练结果见附录。

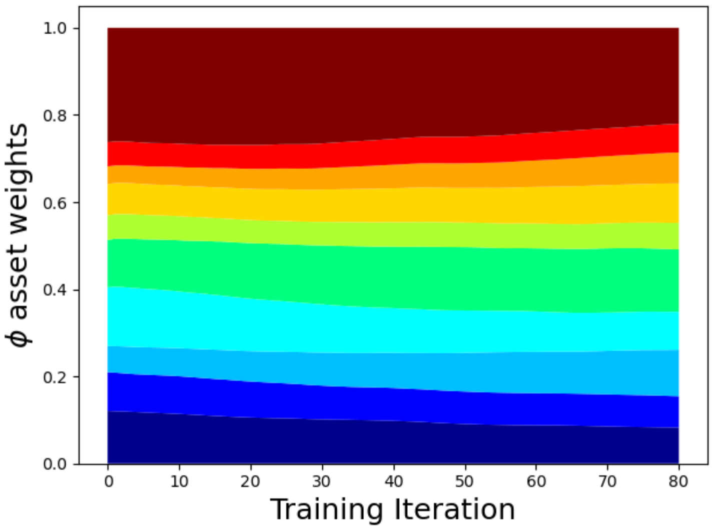
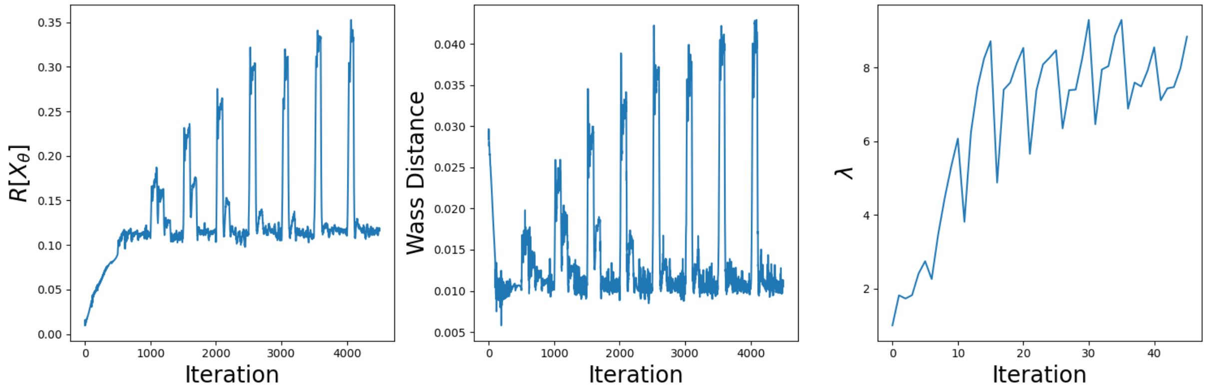
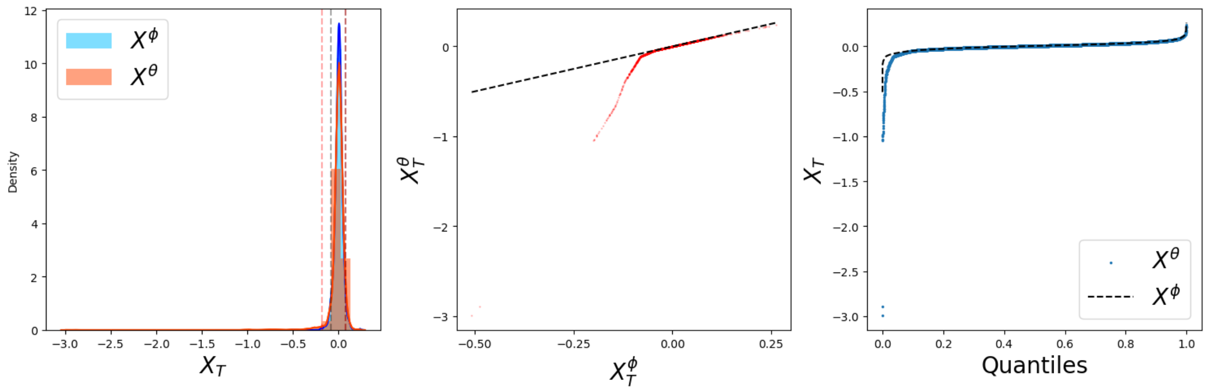


图 1 Wasserstein 距离 $\epsilon=10^{-1}$ 时的外层训练迭代 80 次后的最优资产组合



(a) 训练过程中效用函数、Wasserstein 距离 ϵ、Lagrange 乘子 λ 迭代过程



(b) 最优投资组合与扭曲投资组合分布及关系图

稳健投资组合优化问题的目标自然是在 Wasserstein 距离变化时，即最优投资组合在扭曲时，强化学习的 Agent 仍然能选择最有利于投资者的投资组合。下图是 Wasserstein 距离从 $\epsilon = 10^{-3}$ 到 $\epsilon = 1$ 时最优投资组合中各资产的权重变化。

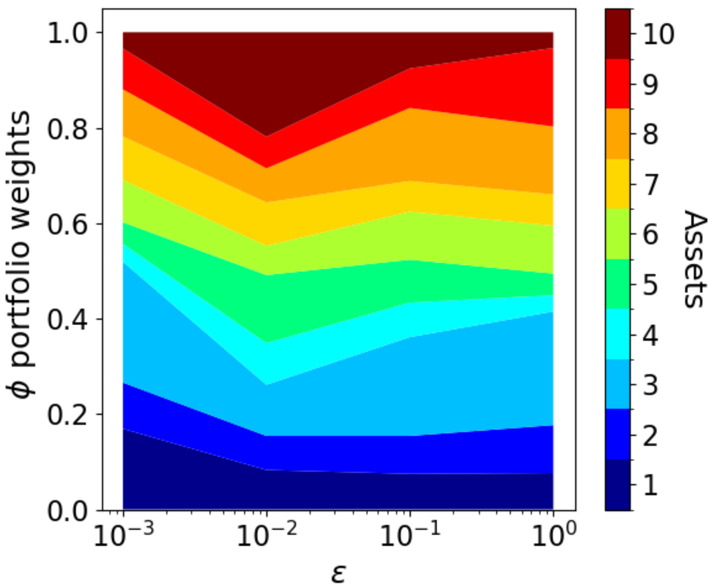


图 2 最优投资组合资产权重随 ϵ 变化

下图是随着 Wasserstein 距离 $\epsilon = 10^{-3}, 10^{-2}, 1$ 变化时终端时刻投资组合的价值分布图：

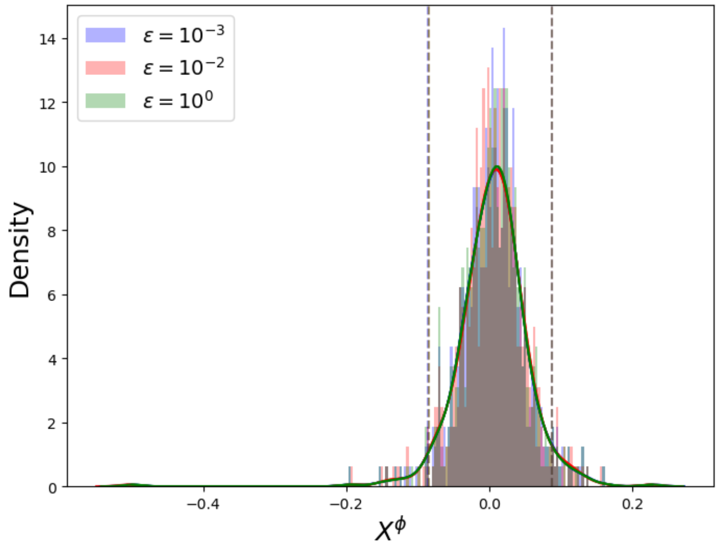


图 3 最优投资组合终端时刻财富的分布

图上的竖直线分别是水平为 $\alpha = 0.1$ 的条件风险价值 (Conditional Value at Risk, CVaR_α) 和终期财富效用 (Utility of Terminal Wealthy, $\mathrm{UTE}_{1-\alpha}$)，分别是 $\alpha = 0.1$ 的下分位数和上分位数。计算结果如表一所示 (CVaR_α 取负值相当于亏损)

表 1 不同 ϵ 终端时刻最优投资组合的 $\mathrm{CVaR}_\alpha, \mathrm{UTE}_{1-\alpha}$

ϵ	CVaR_α	$\mathrm{UTE}_{1-\alpha}$
10^{-3}	-0.08524	0.08730
10^{-2}	-0.08516	0.08716
10^{0}	-0.08471	0.08701

由表中数据可以发现，当 Wasserstein 距离增加时，即扭曲最优投资组合的程度增大时，CVaR_α 在往正向增大，即投资者希望此时的风险减少，寻求可能损失更少的投资组合，而 $\mathrm{UTE}_{1-\alpha}$ 也在减小，由于我们的风险测度 $\gamma(u)$ 取 $p = 0.75$，即更强调损失的度量，这说明 ϵ 增大即对抗者生成更坏的扭曲时投资者宁愿减少收益也要减少损失，符合我们选择风险测度的要求。而对于最优投资组合的不同资产变化图，可以看出当 ϵ 较小时投资者愿意选择风险更大收益更大的股票类而不是债券和货币类产品，随着 ϵ 增大，投资者更倾向于选择债券类收益稳定损失风险小的产品，这也符合稳健投资策略。

6 结论与讨论

6.1 结论

本研究通过引入稳健随机优化框架，结合强化学习策略梯度算法，成功解决了金融投资组合优化问题。主要结论如下。

第一基于秩相依效用（RDEU）和 Wasserstein 距离构建的稳健随机优化框架，能够有效处理金融市场的不确定性，为金融投资组合优化提供了一种新的方法。该框架不仅考虑了期望效用，还通过扭曲函数和 Wasserstein 距离，量化了模型不确定性和分布不确定性对决策的影响。

第二本文提出的基于强化学习策略梯度算法的求解方法，通过内部问题和外部问题的交替优化，成功实现了对投资组合的稳健优化。算法在不同 Wasserstein 距离下的表现表明，该方法能够有效应对不同程度的模型扭曲，选择出最优的投资组合。

第三通过实证数据的收集与处理，结合主成分分析（PCA）方法，我们成功提取了系统性因子和特质性因子，为投资组合优化提供了真实可靠的数据基础。实证结果表明，随着 Wasserstein 距离的增加，投资者更倾向于选择风险较低的资产，选择稳健投资策略。具体表现为：

- 当 Wasserstein 距离较小时，投资者愿意选择风险更大、收益更高的股票类资产；
- 随着 Wasserstein 距离增大，投资者更倾向于选择债券类收益稳定、损失风险小的产品；
- 条件风险价值（CVaR）在 Wasserstein 距离增加时往正向增大，表明投资者希望减少风险，寻求可能损失更少的投资组合；
- 终期财富效用（UTE）在 Wasserstein 距离增加时减小，表明投资者宁愿减少收益也要减少损失，符合风险测度的要求。

第四本文选择的 $\alpha-\beta$ 风险测度在实证中表现出良好的适用性。该风险测度兼顾了投资者的风险厌恶和风险偏好，特别是在 $p=0.75$ 时，更强调损失的度量，符合金融市场的实际情况。

6.2 讨论

尽管本研究取得了一定的成果，但仍有一些值得进一步探讨和改进的地方。

首先，在模型假设部分，虽然本研究结合了真实的资产数据，但在某些假设上仍存在简化。例如，假设系统性因子和特质性因子的分布形式，以及神经网络的结构和参数选择，这些假设可能在不同的市场条件下需要进一步调整和验证。

其次，强化学习策略梯度算法在处理大规模数据和复杂模型时，计算复杂性较高。未来可以探索更高效的算法和计算方法，以提高算法的运行效率和收敛速度。

第三，风险测度的多样性：虽然 $\alpha-\beta$ 风险测度在本研究中表现出良好的适用性，但金融市场的风险特征是多样的。未来可以考虑引入更多的风险测度方法，如熵风险测度、多目标优化等，以更全面地反映投资者的风险偏好。

第四，在实际应用中，市场条件和投资者的风险偏好是动态变化的。未来可以研究动态调整机制，使投资组合能够根据市场变化和投资者的风险偏好进行实时调整，提高投资组合的适应性和稳健性。

最后尽管本研究选取较多种类资产如股票、债券、货币、期货等，但金融市场的资产种类繁多。未来可以扩展到更多的资产类别，如房地产、艺术品等，以构建更全面的投资组合。

综上所述，本研究为金融投资组合优化提供了一种新的稳健优化方法，通过实证数据验证了其有效性和适用性。未来的研究可以在此基础上，进一步完善模型假设，提高算法效率，引入更多风险测度方法，以及探索动态调整机制，以更好地应对金融市场的复杂性和不确定性。

参考文献

[1] Carole Bernard, Silvana M Pesenti, and Steven Vanduffel. Robust distortion risk measures. *Mathematical Finance*, 34(3):774–818, 2024.

[2] Yingjie Fei, Zhuoran Yang, Yudong Chen, Zhaoran Wang, and Qiaomin Xie. Risk-sensitive reinforcement learning: Near-optimal risk-sample tradeoff in regret. *Advances in Neural Information Processing Systems*, 33:22384–22395, 2020.

[3] Sebastian Jaimungal, Silvana M Pesenti, Ye Sheng Wang, and Hariom Tatsat. Robust risk-aware reinforcement learning. *SIAM Journal on Financial Mathematics*, 13(1):213–226, 2022.

[4] Oliver Mihatsch and Ralph Neuneier. Risk-sensitive reinforcement learning. *Machine learning*, 49:267–290, 2002.

[5] Georg Pflug and David Wozabal. Ambiguity in portfolio selection. *Quantitative Finance*, 7(4):435–442, 2007.

[6] Aviv Tamar, Yinlam Chow, Mohammad Ghavamzadeh, and Shie Mannor. Policy gradient for coherent risk measures. *Advances in neural information processing systems*, 28, 2015.

[7] Menahem E Yaari. The dual theory of choice under risk. *Econometrica: Journal of the Econometric Society*, pages 95–115, 1987.

7　附录：团队成员信息及分工

序号	成员姓名	学号	期末报告分工	课堂展示分工
1	姜颖昕	2024103987	引言、随机优化理论框架、强化学习策略梯度算法部分（报告 1 至 3 部分）	论文理论和算法部分汇报
2	韩艳栋	2023103283	论文修改与完善	论文实验部分汇报
3	郭镇涛	2024103985	稳健投资组合优化问题与实证结果，数据处理与实证代码的整理	代码展示，复现论文结果
4	杨哲	2024103999	论文修改与完善	代码展示
5	吴晶晶	2023101730	摘要、结论与讨论、参考文献、格式修正	个人展示：基于券商分析师评级报告的投资决策分析