# Summary of the Master's thesis

Mohammad Joshaghani*

August 2022

## 1 Introduction

This thesis discusses a data-driven decision making problem, in which investors are seeking to create and protect their wealth by investing in financial institutions. Due to various factors such as economic and political conditions, financial institutions exhibit stochastic behavior, requiring advanced mathematical models for effectively managing risks and returns of assets.

Deep reinforcement learning is a framework that can uncover hidden patterns in the data distribution and aid decision-making problems. The objective of this study is to enhance the performance of WaveCorr (Marzban et al. (2021)), a specific deep reinforcement learning model, with a distributionally robust end-to-end learning approach for portfolio optimization. Specifically, the distributionally robust end-to-end learning approach takes into account the risk of the model and the risk of return predictions, both of which are not considered by WaveCorr alone.

In this approach, the decision-making layer optimizes the portfolio by solving a minimax problem, assuming that the distribution of asset returns belongs to an ambiguity set centered around a nominal distribution. The model parameters are updated using implicit differentiable optimization layers.

To evaluate these models, data from the 20 highest large capital companies listed on the Tehran Stock Exchange between 2009 and 2022 is collected. The results indicate that WaveCorr with the distributionally robust end-to-end learning approach improves the performance of portfolio optimization.

## 2 Methodology

Embedded optimization within an end-to-end learning approach is recently proposed for various tasks, such as portfolio optimization (Costa & Iyengar (2023);Butler & Kwon (2023);Uysal et al. (2023)). We follow (Costa & Iyengar (2023)) to define the distributionally robust end-to-end learning approach. As figure 1 shows, it consists of three layers, i.e., prediction layer, decision layer, and task loss. Prediction layer forecasts assets' return given features, decision layer assigns portfolio weights given assets' return and forecast errors, task loss part provides loss which, with stochastic optimization improve the performance of the two previous layers.

*MSc of Business Administration-Finance

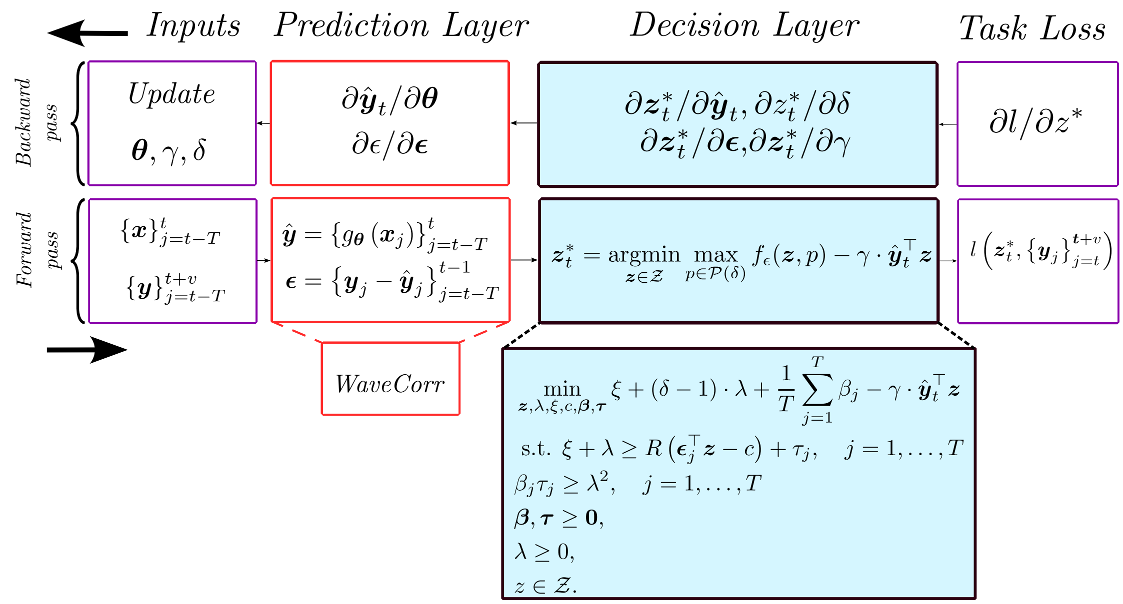


Figure 1: General framework.

lets assume $\boldsymbol{x}_t \in \mathbb{R}^m$ be $m$ predictive features at time $t$, and a goal is to predict $n$ assets returns $\hat{\boldsymbol{y}}_t \in \mathbb{R}^n$. Let $\{\boldsymbol{x}_j \in \mathbb{R}^m : j = t-T, ..., t-1\}$ be the time series of predictive features, and $\{\boldsymbol{y}_j \in \mathbb{R}^n : j = t-T, ..., t-1\}$ the time series of assets returns for $T$ time steps. Let $g_\theta : \mathbb{R}^m \to \mathbb{R}^n$ be a predictive model that maps $\boldsymbol{x}_t$ to $\mathbb{E}[\hat{\boldsymbol{y}}_{\boldsymbol{t}}]$. We consider $g_\theta$ be differentiable on $\theta$ and $\hat{\boldsymbol{y}}_t \triangleq g_\theta(\boldsymbol{x}_t)$. We can consider WaveCorr as $g_\theta(.)$, which its neural network parameters, $\theta$, are trained using gradient policy and Bellman optimality equation. We refer readers to (Marzban et al. (2021)) for more detail about WaveCorr architecture and the reinforcement learning formulation.

let $\bar{\boldsymbol{\epsilon}}_t \triangleq \bar{\boldsymbol{y}}_t - \hat{\boldsymbol{y}}_t = \bar{\boldsymbol{y}}_t - g_\theta(\boldsymbol{x}_t) \in \mathbb{R}^n$ denotes prediction error. We can define portfolio as:

$$\mathbb{Z} \triangleq \{\boldsymbol{z} \in \mathbb{R}^n : \boldsymbol{z} \geq \mathbf{0}, \mathbf{1}^T \boldsymbol{z} = 1\} \tag{1}$$

Let $\boldsymbol{\epsilon} = \{\boldsymbol{\epsilon}_j \in \mathbb{R}^n : j = 1, ..., T\}$ be a finite set of observable prediction errors; and $\mathbf{z} \in \mathbb{Z} \subseteq \mathbb{R}^n$ represents the selected portfolio.

Also, consider a closed convex function $\mathbf{R} : \mathbb{R} \to \mathbb{R}_+ \cup \{+\infty\}$ such that $\mathbf{R}(0) = 0$ and $\mathbf{R}(X) = \mathbf{R}(-X)$.

Let $\mathbf{p}$ be a probability mass function on the probability simplex $\mathcal{Q}$, defined as:

$$\mathcal{Q} \triangleq \{\mathbf{p} \in \mathbb{R} : \mathbf{p} \geq \mathbf{0}, \mathbf{1}^T \boldsymbol{p} = 1\} \tag{2}$$

We can define the deviation risk associated with the existing errors $\boldsymbol{\epsilon}$, the selected portfolio $\mathbf{z}$, and the probability density function $\mathbf{p}$:

$$f_\epsilon(\mathbf{z}, \mathbf{p}) \triangleq \min_{c} \sum_{j=1}^{T} p_j \cdot \mathbf{R}(\boldsymbol{\epsilon}_j^T \mathbf{z} - c) \tag{3}$$

Here, $c$ plays the role of a central parameter.

When $R(X) = X^2$, we can define the deviation risk as:

$$f_\epsilon(\mathbf{z}, \mathbf{p}) = \sum_{j=1}^{T} p_j \cdot (\boldsymbol{\epsilon}_j^T \mathbf{z} - \sum_{k=1}^{T} p_k \cdot \boldsymbol{\epsilon}_k^T \mathbf{z})^2 \tag{4}$$

where $f_\epsilon(\mathbf{z}, \mathbf{p})$ represents the variance. Furthermore, we define the worst-case deviation risk by taking the maximum over $\mathbf{p} \in \mathcal{P}(\delta)$.

Assume that the observation of each component $\epsilon = \epsilon_j : j = t - T, ..., t - 1$ has equal probabilities. In other words, suppose that $q \in Q$ represents a uniform distribution with $q_j = 1/T$ for all $j$. Then, distributional robustness can be defined as the ability of $p$ to vary within a certain range of $q$. The ambiguity set $\mathcal{P}(\delta)$ for the distribution $\mathbf{q}$ can be defined as follows:

$$\mathcal{P}(\delta) \triangleq \left\{ \boldsymbol{p} \in \mathbb{R}^T : \boldsymbol{p} \geq \mathbf{0}, \mathbf{1}^\top \boldsymbol{p} = 1, I_\phi(\boldsymbol{p}, \boldsymbol{q}) \leq \delta \right\} \tag{5}$$

where $\delta$ represents the maximum possible value that the distribution $\mathbf{p}$ can differ from the uniform distribution $\mathbf{q}$.

The optimal distributionally robust layer selects the optimal portfolio $\mathbf{z}_t^*$ by considering the worst-case behavior of $\mathbf{p} \in \mathcal{P}(\delta)$. It solves the following optimization problem where $\gamma \in \mathbb{R}_+$ is the risk-aversion parameter and $f_\epsilon$ is the risk deviation measurement as defined in formula:

$$\boldsymbol{z}_t^* = \underset{\boldsymbol{z} \in \mathcal{Z}}{\operatorname{argmin}} \max_{\boldsymbol{p} \in \mathcal{P}(\delta)} f_{\boldsymbol{\epsilon}}(\boldsymbol{z}, \boldsymbol{p}) - \gamma \cdot \hat{\boldsymbol{y}}_t^\top \boldsymbol{z} \tag{6}$$

Usually, solving the above optimization problem requires complex and time-consuming mathematical operations, which may lead to stopping points in local optimal solutions.

Based on formula 6, it can be observed that the objective function is convex with respect to $\mathbf{z}$ . However, the convexity of the objective function with respect to the risk deviation measurement $f_\epsilon(\mathbf{z}, \mathbf{p})$ needs to be examined more carefully. Therefore, using convex duality, the objective function in equation 6 can be rewritten as a convex function with respect to the parameters and variables.

By applying minimax theorem for convex duality, the objective function in equation 6 can be written as follows:

$$\begin{aligned} \max_{\boldsymbol{p} \in \mathcal{P}(\delta)} f_{\boldsymbol{\epsilon}}(\boldsymbol{z}, \boldsymbol{p}) &= \max_{\boldsymbol{p} \in \mathcal{P}(\delta)} \min_{c \in \mathbb{R}} \sum_{j=1}^{T} p_j \cdot R\left(\boldsymbol{\epsilon}_j^\top \boldsymbol{z} - c\right) \\ &= \min_{c \in \mathbb{R}} \max_{\boldsymbol{p} \in \mathcal{P}(\delta)} \sum_{j=1}^{T} p_j \cdot R\left(\boldsymbol{\epsilon}_j^\top \boldsymbol{z} - c\right) \end{aligned} \tag{7}$$

Then, maximizing with respect to $\mathcal{P}(\delta)$ in equation 5 can be rewritten as follows:

$$
\begin{aligned}
\max_{\boldsymbol{p}} & \sum_{j=1}^{T} p_j \cdot R\left(\boldsymbol{\epsilon}_j^\top \boldsymbol{z} - c\right) \\
\text{s.t.} & \sum_{j=1}^{T} p_j = 1 && \text{(8a)} \\
& I_\phi(\boldsymbol{p}, \boldsymbol{q}) = \sum_{j=1}^{T} q_j \cdot \phi\left(p_j / q_j\right) \leq \delta && \text{(8b)} \\
& \boldsymbol{p} \geq \mathbf{0}
\end{aligned}
$$

By assigning a dual variable $\xi$ to the constraint in 8a and a dual variable$\lambda \geq 0$ to the constraint in 8b, the Lagrangian dual function $f_{\boldsymbol{\epsilon}}^{\delta}(\boldsymbol{z}, c, \lambda, \xi)$ can be defined as follows, where $\phi(.)^*$ is the convex conjugate function for $\phi(.)$. Additionally, $\lambda \geq 0, \quad \xi \in \mathbb{R}$ are auxiliary variables added to the model by constructing the dual function:

$$
\begin{aligned}
& f_{\boldsymbol{\epsilon}}^{\delta}(\boldsymbol{z}, c, \lambda, \xi) \\
& \quad \triangleq \max_{\boldsymbol{p} \geq 0} \sum_{j=1}^{T} p_j \cdot R\left(\boldsymbol{\epsilon}_j^\top \boldsymbol{z} - c\right) + \xi \cdot \left(1 - \mathbf{1}^\top \boldsymbol{p}\right) \\
& \qquad + \lambda \cdot \left(\delta - \sum_{j=1}^{T} q_j \cdot \phi\left(p_j / q_j\right)\right) \\
& \quad = \xi + \delta \cdot \lambda + \max_{\boldsymbol{p} \geq \mathbf{0}} \sum_{j=1}^{T} \left(p_j \cdot \left(R\left(\boldsymbol{\epsilon}_j^\top \boldsymbol{z} - c\right) - \xi\right)\right. \\
& \qquad \left. - \lambda \cdot q_j \cdot \phi\left(p_j / q_j\right)\right) \\
& \quad = \xi + \delta \cdot \lambda + \sum_{j=1}^{T} \max_{p_j \geq 0} \left(p_j \cdot \left(R\left(\boldsymbol{\epsilon}_j^\top \boldsymbol{z} - c\right) - \xi\right)\right. \\
& \qquad \left. - \lambda \cdot q_j \cdot \phi\left(p_j / q_j\right)\right) \\
& \quad = \xi + \delta \cdot \lambda + \sum_{j=1}^{T} q_j \cdot \max_{s \geq 0} \left(s \cdot \left(R\left(\boldsymbol{\epsilon}_j^\top \boldsymbol{z} - c\right) - \xi\right)\right. \\
& \qquad \left. - \lambda \cdot \phi(s)\right) \\
& \quad = \xi + \delta \cdot \lambda + \sum_{j=1}^{T} q_j \cdot (\lambda \phi)^* \left(R\left(\boldsymbol{\epsilon}_j^\top \boldsymbol{z} - c\right) - \xi\right) \\
& \quad = \xi + \delta \cdot \lambda + \frac{\lambda}{T} \sum_{j=1}^{T} \phi^* \left(\frac{R\left(\boldsymbol{\epsilon}_j^\top \boldsymbol{z} - c\right) - \xi}{\lambda}\right). && \text{(9a)}
\end{aligned}
$$

In the following, it will be shown that in formula 9, if $\phi$ be the Helinger distance, it can be rewritten as a second-order cone problem in formula 15.

The Helinger distance is defined as:

$$I_{\phi}^{H}(\boldsymbol{p}, \boldsymbol{q}) \triangleq \sum_{j=1}^{T} q_j \cdot \phi_H\left(p_j / q_j\right) = \sum_{j=1}^{T}\left(\sqrt{p_j} - \sqrt{q_j}\right)^2 \tag{10}$$

and

$$\phi_H(w) \triangleq (\sqrt{w} - 1)^2, \quad w \geq 0 \ \phi_H^*(s) \quad \triangleq \frac{s}{1-s}, \quad s < 1. \tag{11}$$

Let's assume:

$$h(s) \triangleq \frac{1}{1-s} = \phi_H^*(s) + 1 \quad h^{-1}(s) \triangleq 1 - \frac{1}{s} \tag{12}$$

Then we can write 9a as:

$$\xi + (\delta - 1) \cdot \lambda + \frac{1}{T} \sum_{j=1}^{T} \lambda \cdot h\left(\frac{R\left(\boldsymbol{\epsilon}_j^{\top} \boldsymbol{z} - c\right) - \xi}{\lambda}\right) - \gamma \cdot \hat{\boldsymbol{y}}_t^{\top} \boldsymbol{z} \tag{13}$$

Then using the auxiliary variable $\beta \in \mathbb{R}^T$, we can write:

$$\begin{aligned} \beta_j \geq \lambda \cdot h\left(\frac{R\left(\boldsymbol{\epsilon}_j^{\top} \boldsymbol{z} - c\right) - \xi}{\lambda}\right) & \iff h^{-1}\left(\frac{\beta_j}{\lambda}\right) \geq \frac{R\left(\boldsymbol{\epsilon}_j^{\top} \boldsymbol{z} - c\right) - \xi}{\lambda} \\ & \iff \lambda - \frac{\lambda^2}{\beta_j} \geq R\left(\boldsymbol{\epsilon}_j^{\top} \boldsymbol{z} - c\right) - \xi \end{aligned} \tag{14}$$

Since $\phi_H^*(s)$ requires $s < 1$, we have $\boldsymbol{\beta} \geq \mathbf{0}$. And we can write with the help of the auxiliary variable $\boldsymbol{\tau} \in \mathbb{R}_+^T$:

$$\begin{array}{lll} \min\limits_{\boldsymbol{z}, \lambda, \xi, c, \boldsymbol{\beta}, \boldsymbol{\tau}} & \xi + (\delta - 1) \cdot \lambda + \frac{1}{T} \sum_{j=1}^{T} \beta_j - \gamma \cdot \hat{\boldsymbol{y}}_t^{\top} \boldsymbol{z} & \\ \text{s.t.} & \xi + \lambda \geq R\left(\boldsymbol{\epsilon}_j^{\top} \boldsymbol{z} - c\right) + \tau_j, & j = 1, \ldots, T \\ & \beta_j \tau_j \geq \lambda^2, & j = 1, \ldots, T \\ & \boldsymbol{\beta}, \boldsymbol{\tau} \geq \mathbf{0}, & \\ & \lambda \geq 0, & \\ & \boldsymbol{z} \in \mathcal{Z} & \end{array} \tag{15}$$

## 2.1 Loss function and metrics

We used MSE as a return prediction loss function:

$$l_{\text{mse}}\left(\hat{\boldsymbol{y}}_t, \boldsymbol{y}_t\right) \triangleq \frac{1}{n}\left\|\boldsymbol{y}_t - \hat{\boldsymbol{y}}_t\right\|_2^2 \tag{16}$$

Moreover, the task loss function ,$l_{\text{task}}$ , is composed of two parts: the MSE loss ,$l_{\text{mse}}$, and the portfolio value loss $l_{\text{PR}}$, where $l_{\text{PR}}$ exhibits portfolio final value when started with 1$:

$$
\begin{aligned}
l_{\text{task}}\ \left(\boldsymbol{z}_t^*, \hat{\boldsymbol{y}}_t, \left\{\boldsymbol{y}_j\right\}_{j=t}^{t+v}\right) =& 0.5 \cdot l_{\text{mse}}\left(\hat{\boldsymbol{y}}_t, \boldsymbol{y}_t\right) + l_{\text{PR}}\left(\boldsymbol{z}_t^*, \left\{\boldsymbol{y}_j\right\}_{j=t}^{t+v}\right), \\
l_{\text{PR}}\left(\boldsymbol{z}_t^*, \left\{\boldsymbol{y}_j\right\}_{j=t}^{t+v}\right) =&\ \text{-}\mathit{cumprod}(1 + \left\{\boldsymbol{y}_j\right\})^{\frac{1}{v}}
\end{aligned}
\tag{17}
$$

# 3 Numerical experiment

We used CvxpyLayer package (Agrawal et al. (2019)) to implement implicit differentiable optimization layer. Also, for hyper-parameters optimization we consider different weight decays ($WD$) and epochs : $WD \in \{0.009, 0.007, 0.005\}$ and $\mathit{epoch} \in \{1, 2, 3, 4\}$; and select one that yields a lower validation loss, i.e., $WD = 0.007, \mathit{epoch} = 3$ for $WaveCorr\_DRO$ and $WD = 0.007, \mathit{epoch} = 1$ for $WaveCorr\_Casual$.

Figure 2 shows each model final portfolio value, where $WaveCorr\_Casual$ is the benchmark WaveCorr, $WaveCorr\_DRO$ is the distributionally robust approach, and *EW* represents a method which each asset gets equal weight. As it shows, final value in the test set for $WaveCorr\_DRO$, $WaveCorr\_Casual$, and *EW* are 1.7, 1.4, and 1.2 respectively. We also performed *Independent Samples t Test*; and it illustrates distributionally robust end-to-end learning approach enhances WaveCorr performance.

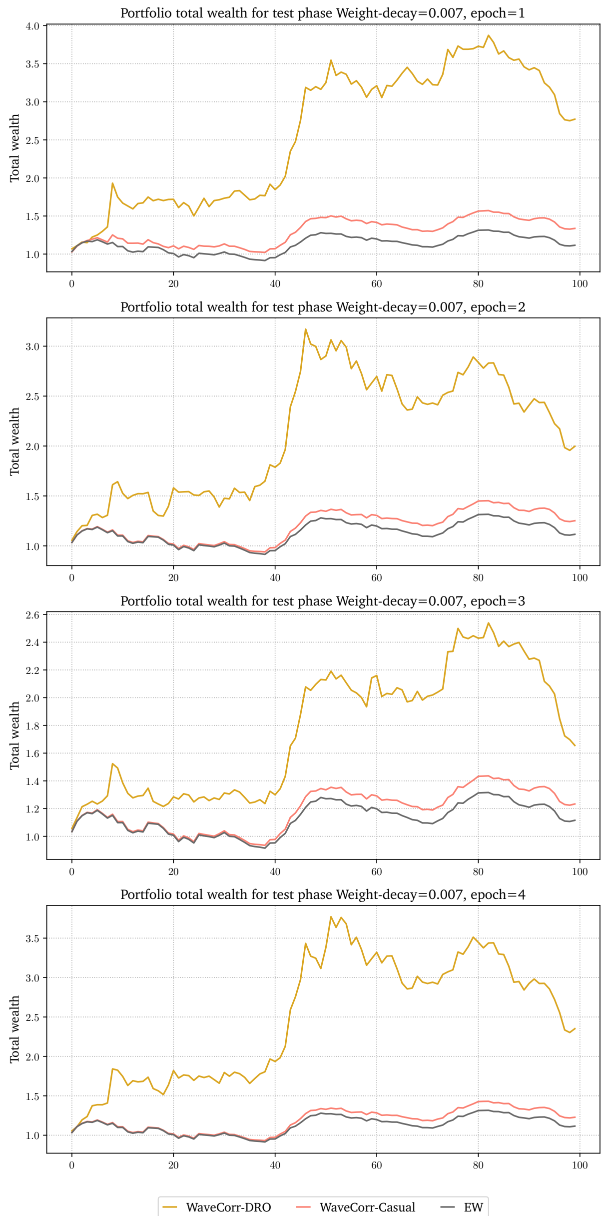


Figure 2: Portfolio performance in the test phase.

# References


Agrawal, A., Amos, B., Barratt, S., Boyd, S., Diamond, S., & Kolter, J. Z. (2019). Differentiable convex optimization layers. *Advances in neural information processing systems*, *32*.

Butler, A., & Kwon, R. H. (2023). Integrating prediction in mean-variance portfolio optimization. *Quantitative Finance*, *23*, 429–452.

Costa, G., & Iyengar, G. N. (2023). Distributionally robust end-to-end portfolio construction. *Quantitative Finance*, *23*, 1465–1482.

Marzban, S., Delage, E., Li, J. Y., Desgagne-Bouchard, J., & Dussault, C. (2021). Wavecorr: Correlation-savvy deep reinforcement learning for portfolio management. `arXiv:2109.07005`.

Uysal, A. S., Li, X., & Mulvey, J. M. (2023). End-to-end risk budgeting portfolio optimization with neural networks. *Annals of Operations Research*, (pp. 1–30).