Portfolio Construction with Single Stage Optimization

- Zoe Zhao, Utkarshbhanu Andurkar,
Jinwei Zhu

Portfolio Construction Problem

Traditional 2 Step Approach

Predict ->

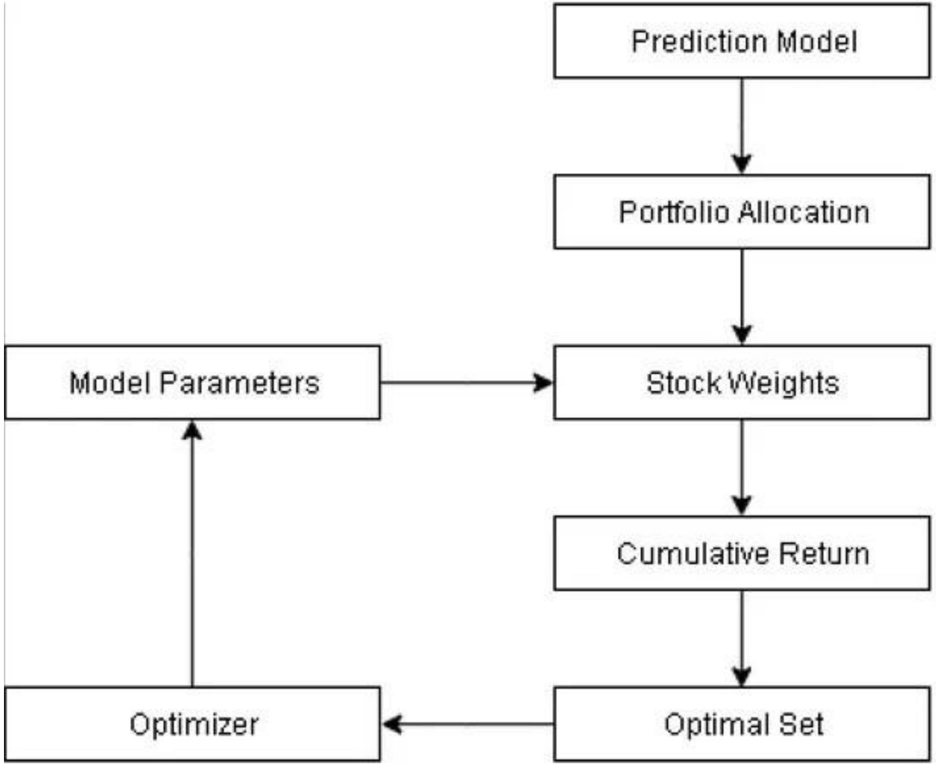
$$\mu, \Sigma$$

Minimize ->

$$w^T \Sigma w$$

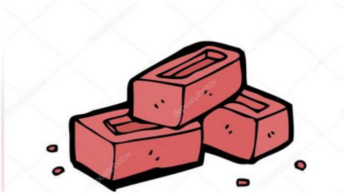
Given,

$$w^T \mu = E$$

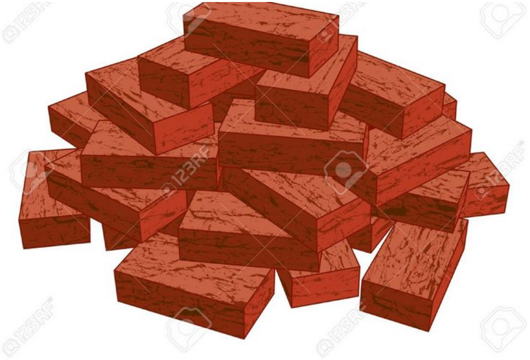

Prediction Model
Portfolio Allocation
Model Parameters
Stock Weights
Cumulative Return
Optimizer
Optimal Set

Drawbacks

Error in step 1



Accumulated error in step 2



Single Stage Portfolio Optimization With Deep Learning

Model Free Approach with feed forward layer

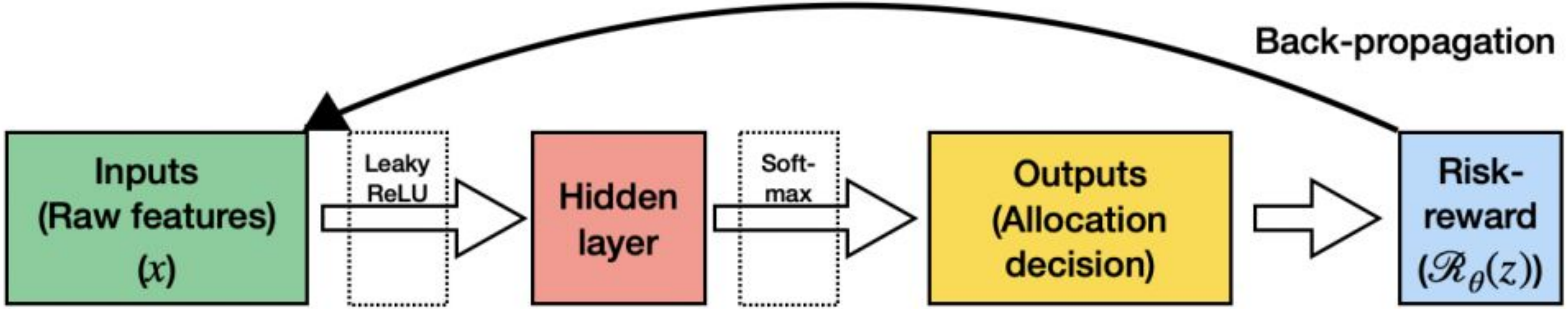


Figure 1: Computational graph of model-free approach.

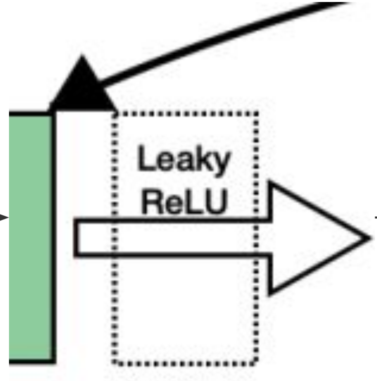
Forward Propagation

$$X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}$$

$$E = \begin{bmatrix} E_{11} \\ E_{21} \end{bmatrix}$$

Forward Propagation

$$X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}$$



$$Y = \text{LeakyReLU}(X)$$

$$Y = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix}$$

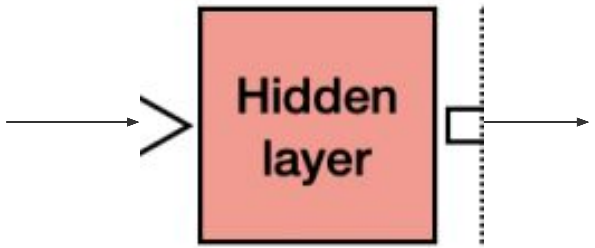
where,

$$Y_{11} = \max(x_{11}, 0) + \min(x_{11}, 0) \times 0.01$$
$$Y_{12} = \max(x_{12}, 0) + \min(x_{12}, 0) \times 0.01$$
$$Y_{21} = \max(x_{21}, 0) + \min(x_{21}, 0) \times 0.01$$
$$Y_{22} = \max(x_{22}, 0) + \min(x_{22}, 0) \times 0.01$$

Forward Propagation

$$Z = Y.W + B$$

$$Y = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix}$$



$$\begin{bmatrix} Z_{11} \\ Z_{21} \end{bmatrix} = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} \begin{bmatrix} W_{11} \\ W_{21} \end{bmatrix} + \begin{bmatrix} b \\ b \end{bmatrix}$$

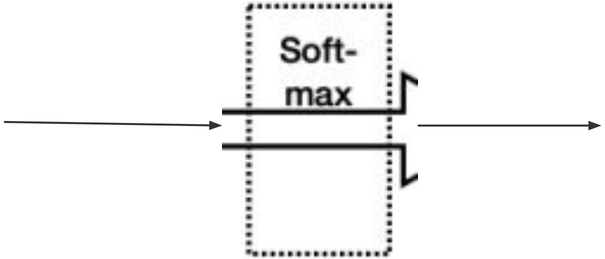
$$Z_{11} = W_{11}Y_{11} + W_{21}Y_{12} + b$$

$$Z_{21} = W_{11}Y_{21} + W_{21}Y_{22} + b$$

Forward Propagation

$$S = \text{Softmax}(Z)$$

$$\begin{bmatrix} Z_{11} \\ Z_{21} \end{bmatrix}$$

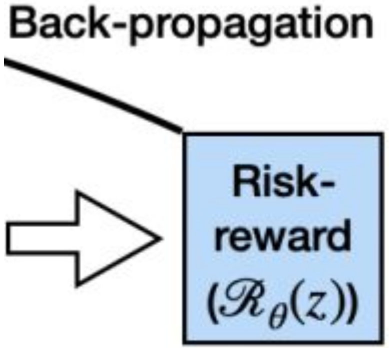


$$S_{11} = \frac{e^{Z_{11}}}{e^{Z_{11}} + e^{Z_{21}}}, \quad S_{21} = \frac{e^{Z_{21}}}{e^{Z_{11}} + e^{Z_{21}}}$$

$$\begin{bmatrix} S_{11} \\ S_{21} \end{bmatrix} = \frac{1}{e^{Z_{11}} + e^{Z_{21}}} \begin{bmatrix} e^{Z_{11}} \\ e^{Z_{21}} \end{bmatrix}$$

Forward Propagation

$$\begin{bmatrix} S_{11} \\ S_{21} \end{bmatrix}$$



$$-\begin{bmatrix} S_{11} & S_{21} \end{bmatrix} \begin{bmatrix} E_{11} \\ E_{21} \end{bmatrix}$$

$$L = -(S_{11}E_{11} + S_{21}E_{21})$$

$$L = -S^{\top}.E$$

Backpropagation

We have 3 parameters to tune: W_{11}, W_{21}, b.

Using Gradient descent with learning rate α, we will have following set of equations for the backpropagation step.

$$W_{11} = W_{11} - \alpha \left(\frac{\partial L}{\partial W_{11}} \right)$$
$$W_{21} = W_{21} - \alpha \left(\frac{\partial L}{\partial W_{21}} \right)$$
$$b = b - \alpha \left(\frac{\partial L}{\partial b} \right)$$

We need to calculate the derivatives to update the weights.

Formulae for Backpropagation

$$W_{11} = W_{11} - \alpha S_{11} S_{21} \left(Y_{21} - Y_{11}\right)\left(E_{11} - E_{21}\right)$$

$$W_{21} = W_{21} - \alpha S_{11} S_{21} \left(Y_{22} - Y_{12}\right)\left(E_{11} - E_{21}\right)$$

$$b = b$$

Toy Example

Let us use a concrete example to solidify the concepts we learned in the previous section. The matrix X contains daily returns for 2 assets for yesterday and day before yesterday, while E vector contains daily returns for those two assets today.

$$X = \begin{pmatrix} 2 & 1 \\ 4 & -2 \end{pmatrix}, \quad E = \begin{pmatrix} 4 \\ 6 \end{pmatrix}, \quad W = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}, \quad b = 0.5$$

Toy Example - Continued

1. **LeakyReLU**
   As $\text{LeakyReLU}(x) = \max(x, 0) + \min(x, 0) \times 0.01$

$$\text{LeakyReLU}(X) = Y = \begin{pmatrix} 2 & 1 \\ 4 & -0.02 \end{pmatrix}$$

NYU

Toy Example - Continued

2. **Linear Layer**
   $Z = Y \cdot W + b$

$$Z = \begin{pmatrix} 2 & 1 \\ 4 & -0.02 \end{pmatrix} \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix} + \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$$

$$Z = \begin{pmatrix} 2 \\ 2.49 \end{pmatrix}$$

NYU

Toy Example - Continued

3. **Softmax**

   $S = \text{Softmax}(Z)$

$$S = \begin{pmatrix} \frac{e^2}{e^2+e^{2.49}} \\ \\ \frac{e^{2.49}}{e^2+e^{2.49}} \end{pmatrix}$$

$$S = \begin{pmatrix} \frac{7.38906}{19.45033} \\ \\ \frac{12.06128}{19.45033} \end{pmatrix}$$

$$\text{Therefore, } S = \begin{pmatrix} 0.3799 \\ 0.6201 \end{pmatrix}$$

NYU

Toy Example - Continued

$$
\begin{aligned}
s_{11} &= 0.3799, \quad s_{21} = 0.6201 \\
y_{11} &= 2, \quad y_{12} = 1, \\
y_{21} &= 4, \quad y_{22} = -0.02 \\
E_{11} &= 4, \quad E_{21} = 6 \\
W_{11} &= 0.5, \quad W_{21} = 0.5, \alpha = 0.01
\end{aligned}
$$

Toy Example - Continued

$$W_{11} = W_{11} - \alpha s_{11} s_{21} (y_{21} - y_{11})(E_{11} - E_{21})$$

$$= 0.5 - 0.01 \times 0.3799 \times 0.6201 \times 2 \times -2$$
$$= 0.5 + 9.4230396 \times 10^{-3}$$
$$W_{11} = 0.5094$$

NYU

Toy Example - Continued

$$
\begin{aligned}
W_{21} &= W_{21} - \alpha s_{11} s_{21} (y_{22} - y_{12})(E_{11} - E_{21}) \\
&= 0.5 - 0.01 \times 0.3799 \times 0.6201 \times (-1.02) \times -2 \\
&= 0.5 - 4.8058 \times 10^{-3} \\
W_{21} &= 0.4952
\end{aligned}
$$

$$b = b = 0.5$$

NYU

Toy Example - Continued

Inference:

Let us predict what should be our allocation tomorrow with the weight we learned from the toy data. Now the X matrix will have daily returns for 2 assets for today and yesterday.

$$X = \begin{pmatrix} 4 & 2 \\ 6 & 4 \end{pmatrix}, \quad W = \begin{pmatrix} 0.5094 \\ 0.4952 \end{pmatrix}, \quad b = 0.5$$

NYU

Toy Example - Continued

1. **LeakyReLU**

$$\text{LeakyReLU}(X) \Rightarrow Y = \begin{pmatrix} 4 & 2 \\ 6 & 4 \end{pmatrix}$$

Toy Example - Continued

2. **Linear Layer**

$$Z = Y\,W + b$$

$$Z = \begin{pmatrix} 4 & 2 \\ 6 & 4 \end{pmatrix} \begin{pmatrix} 0.5094 \\ 0.4952 \end{pmatrix} + \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$$

$$Z = \begin{pmatrix} 3.0281 \\ 5.0373 \end{pmatrix} + \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$$

$$Z = \begin{pmatrix} 3.5281 \\ 5.5373 \end{pmatrix}$$

Toy Example - Continued

3. **Softmax**

$$S = \text{Softmax}(Z)$$

$$S = \begin{pmatrix} \frac{e^{3.5281}}{e^{3.5281}+e^{5.5373}} \\ \\ \frac{e^{5.5373}}{e^{3.5281}+e^{5.5373}} \end{pmatrix}$$

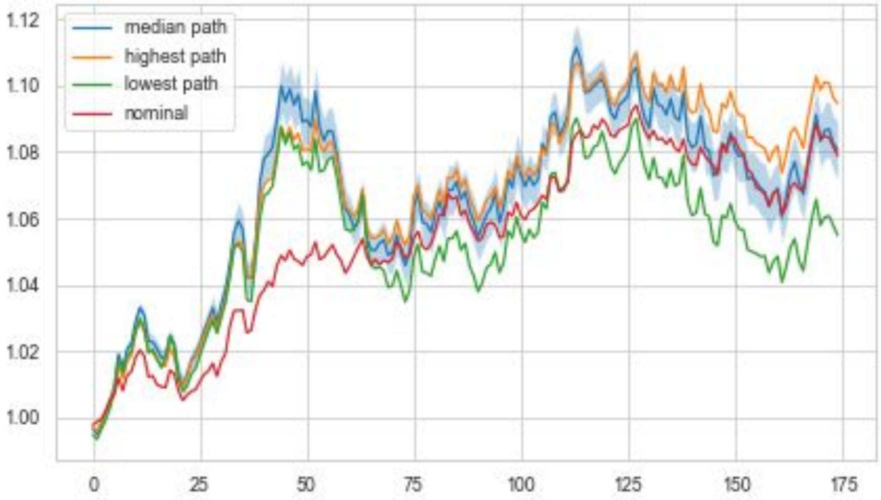
$$S = \begin{pmatrix} 0.1182 \\ 0.8818 \end{pmatrix}$$

This means we should allocate 11.82% of our wealth to asset 1 and 88.18% of our wealth to asset 2 tomorrow.

NYU

Results and Performance Analysis

Results on Simulated Data



NYU

Risk Parity

Category	Description
Asset Types	Stocks and Bonds
Annual Volatility	Stocks: 15%, Bonds: 5%
Asset Allocation	Stocks: 25%, Bonds: 75%
Risk Contribution	Equal risk contribution from both stocks and bonds

Results on Real Market Data

Portfolio	Return	Volatility	Sharpe Ratio	Return/Ave.DD
	2017-2021/06			
Nominal RP	0.0609	0.0604	0.7906	3.2988
Model-free (cumulative return)	0.0040	0.1812	-0.0463	-0.0633
Model-free (Sharpe ratio)	0.0508	0.1194	0.3168	0.6190

Table 1: Portfolio Performance [*Uysal, A.S (2021)*]

NYU

Hypothesis Testing

$$H_0 : \hat{R}_{end-to-end} \geq R_{parity}$$

$$H_\alpha : \hat{R}_{end-to-end} < R_{parity}$$

The corresponding p value is 0.00001.
Which means we can reject the null hypothesis.

Conclusion and Improvement

- Model free approach fails to beat risk parity benchmark
- Improvement in term of Model based approach

Model Based Approach Structure

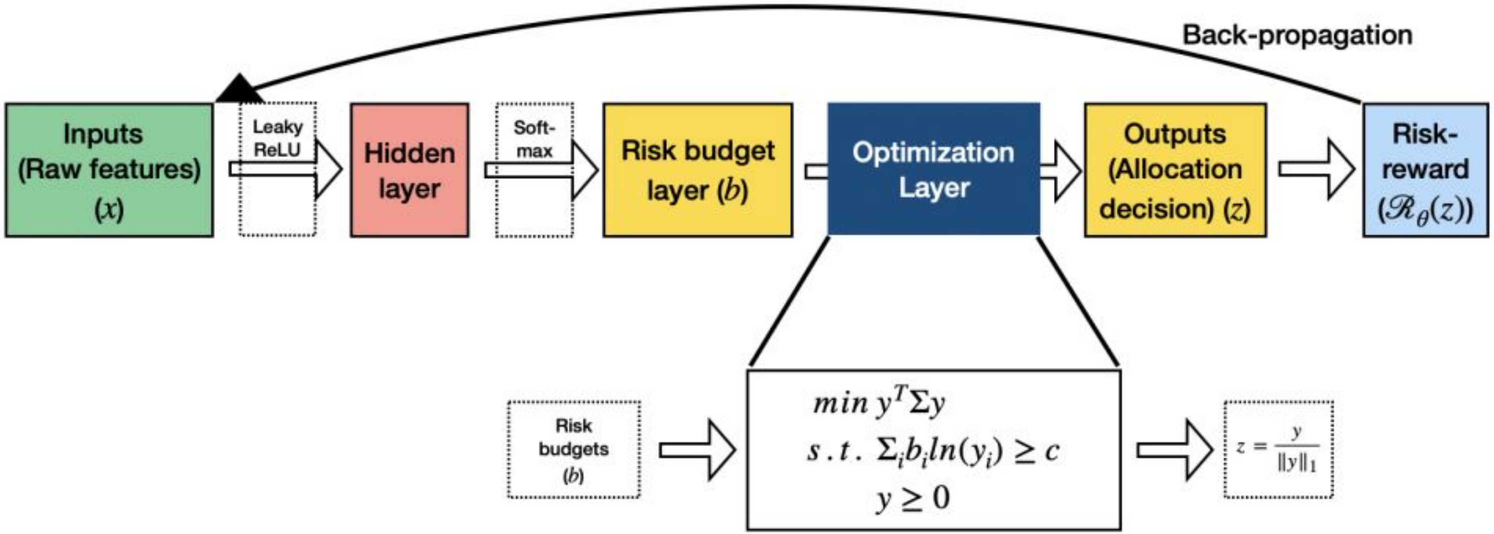


Figure 2: Computational graph of model-based approach.

NYU

Industry Challenges

- Computationally intensive than simple regression models
- Slower than simple regression models
- Less Interpretability

NYU

1. What is the objective of Markowitz portfolio optimization?

01. Maximize Expected Returns

02. Minimize Expected Returns

03. Maximize Variance

04. Minimize Variance for a given level of Expected Return

NYU

2. What is the issue with two stage portfolio optimization process, which makes us prefer end to end approach?

01. It is fast

02. It is slow

03. It is computationally expensive

04. It can lead to error accumulation and subsequently - suboptimal asset allocation decision

NYU

3. What is the function of Softmax layer in the structure of model free neural network?

01. Make the neural network fast

02. Make the neural network slow

03. Make it possible to treat outputs of the neural network as allocation decisions

04. It has no role

NYU

4. Which of the following is benefit of model free end to end approach?

0 1 . It follows single stage optimization

0 2 . It is unable to outperform risk parity benchmark

0 3 . It is computationally expensive than simple linear regression

0 4 . It is less interpretable

NYU

5. Which is not a major risk associated with using deep learning models for financial domain?

01. High computational cost

02. Low Interpretability

03. Overfitting

04. Better Performance

NYU

References

- Uysal, A.S., Li, X. & Mulvey, J.M. ”End-to-end risk budgeting portfolio optimization with neural networks.” Ann Oper Res (2023). (https://arxiv.org/abs/2107.04636)

- Bahri Sales, J., Pakmaram, A., & Valizadeh, M. (2018). ”Selection and Portfolio Optimization by Mean–Variance Markowitz Model and Using the Different Algorithms.” Financial Knowledge of Securities Analysis, 11(37), 43-53.

- Toy Code for model free approach https://colab.research.google.com/drive/1bMgKXGBvse9cgXVviR19YC8NV-bLekjP?usp=sharing