Forecasting Portfolios

Author

Autumn 2024

1 Goal

Find optimal portfolio weights based on MV approach considering that the error of expected return prediction between assets can be correlated.

2 Definitions

The weights of the tangency portfolio with k assets are defined by:

$$\mathbf{w}_t = \frac{\Sigma^{-1}\Lambda}{\mathbf{1}\Sigma^{-1}\Lambda_t}$$

Where $\Lambda_{t+1}^{\top} = [\lambda_{1,t+1}, \lambda_{2,t+1}, \lambda_{3,t+1}, \ldots, \lambda_{k,t+1}]$ and

$$\lambda_{i,t+1} := \mathbb{E}[\tilde{r}_{i,t+1}], \quad \text{for } i = 1, \ldots, k$$

$\mathbb{E}[\tilde{r}_{i,t+1}]$ is the expected excess return for asset i, which can be calculated by an unknown function $f_i(X)$:

$$\mathbb{E}[\tilde{r}_{i,t+1}] = f_i(X_t)$$

In sample:

$$\hat{\lambda}_{i,t+1} = \hat{f}_i(X_t)$$

$\hat{\lambda}_{i,t+1}$ can be estimated by any machine learning model.

For instance, in a deep neural network with no hidden layer, squared error loss and no activation function (linear regression) aiming to estimate the $\hat{\lambda}_{i,t+1}$ for $i = 1, \ldots, k$ with the same features, the general loss function is defined by:

$$L = \sum_{i=1}^{k} \|\lambda_i - \hat{\lambda}_i\|^2$$

Where λ_i is a vector of the time-series excess returns of asset i and $\hat{\lambda}_i$ is the estimated λ_i.

$$\lambda_i^\top := [\lambda_{i,1}, \lambda_{i,2}, \ldots, \lambda_{i,T}]$$

$$\hat{\lambda}_i^\top := \left[\hat{\lambda}_{i,1}, \hat{\lambda}_{i,2}, \ldots, \hat{\lambda}_{i,T}\right]$$

As mentioned, the predictions for any t considering the example of the linear regression:

$$\hat{\lambda}_{i,t} = \hat{f}_i(X_{t-1}) = \hat{\theta}^\top X_{t-1}$$

Where $\hat{\theta}$ is the vector of estimated parameters considering any loss function (in our current example, considering squared error loss).

For an example, lets consider that we have two assets in our portfolio:

$$L = \|\lambda_1 - \hat{\lambda}_1\|^2 + \|\lambda_2 - \hat{\lambda}_2\|^2$$

For a deep neural network with no hidden layers, the predictions considering X are independent, since the parameters used to estimate $\hat{\lambda}_1$ are different from the ones used to estimate $\hat{\lambda}_2$

$$L = \|\lambda_1 - \theta_1^\top \mathbf{X}\|^2 + \|\lambda_2 - \theta_2^\top \mathbf{X}\|^2$$

Where θ_1 are the parameters used for the estimation of asset 1 and $\hat{\theta}_2$ are the parameters used for the estimation of asset 2.

$$\theta := \{\theta_1, \theta_2\}$$

$$\hat{\theta}_1, \hat{\theta}_2 = \text{argmin}_{\theta_1,\theta_2} L$$

3 Our Suggestion

- A prediction of returns is more important if we have more weight for it.
- We want the error of the predictions to compensate each other at any given time: for instance, if in time t, you have an error that is positive for asset 1, you want a negative error in your prediction for asset 2 in the same time. Nonetheless, if asset 1 is more important in the portfolio than asset 2, you should weight the error of asset 1 more than the one of asset two.

The new loss for 2 assets portfolio:

$$\begin{aligned}
L &= \alpha\left(\|\lambda_1-\hat{\lambda}_1\|^2+\|\lambda_2-\hat{\lambda}_2\|^2\right)+(1-\alpha)\sum_{i=1}^{T}\left[w_1(\lambda_{1,i}-\hat{\lambda}_{1,i})+w_2(\lambda_{1,i}-\hat{\lambda}_{1,i})\right]\\
&= \alpha\left(\|\lambda_1-\hat{\lambda}_1\|^2+\|\lambda_2-\hat{\lambda}_2\|^2\right)+(1-\alpha)\|w_1\lambda_1-w_1\hat{\lambda}_1+w_2\lambda_2-w_2\hat{\lambda}_2\|^2\\
&= \alpha\left(\|\lambda_1-\hat{\lambda}_1\|^2+\|\lambda_2-\hat{\lambda}_2\|^2\right)+(1-\alpha)\|w_1(\lambda_1-\hat{\lambda}_1)+w_2(\lambda_2-\hat{\lambda}_2)\|^2\\
&= \alpha\left(\|\lambda_1-\hat{\lambda}_1\|^2+\|\lambda_2-\hat{\lambda}_2\|^2\right)+(1-\alpha)\mathbf{1}(\gamma\mathbf{w})
\end{aligned}$$

Where:

$$\gamma=\begin{bmatrix}
\lambda_{1,1}-\hat{\lambda}_{1,1} & \lambda_{2,1}-\hat{\lambda}_{2,1}\\
\lambda_{1,2}-\hat{\lambda}_{1,2} & \lambda_{2,2}-\hat{\lambda}_{2,2}\\
\lambda_{1,3}-\hat{\lambda}_{1,3} & \lambda_{2,3}-\hat{\lambda}_{2,3}\\
\vdots & \vdots\\
\lambda_{1,T}-\hat{\lambda}_{1,T} & \lambda_{2,T}-\hat{\lambda}_{2,T}
\end{bmatrix}$$

Or:

$$\begin{aligned}
L &= \alpha\left(\|\lambda_1-\hat{\lambda}_1\|^2+\|\lambda_2-\hat{\lambda}_2\|^2\right)+(1-\alpha)\sum_{i=1}^{T}\left[w_1(\lambda_{1,i}-\hat{\lambda}_{1,i})\cdot w_2(\lambda_{1,i}-\hat{\lambda}_{1,i})\right]\\
&= \alpha\left(\|\lambda_1-\hat{\lambda}_1\|^2+\|\lambda_2-\hat{\lambda}_2\|^2\right)+(1-\alpha)\left\langle w_1(\lambda_1-\hat{\lambda}_1),w_2(\lambda_2-\hat{\lambda}_2)\right\rangle
\end{aligned}$$

(for more assets, can maybe use the matrix format and minimize the forbenius norm of the matrix)

The estimation works as:

1. We define initia values for w_1, w_2.
2. Optimize the function and get λ_i for $\forall\ i$.
3. Plug-in $\lambda_{i,t+1}$ in the tangency portfolio formula and get $\mathbf{w}$.
4. Use the values of $\mathbf{w}$ as the new w_1, w_2.
5. Repeat step 2 to 4 until change in $\Delta\mathbf{w} < \varepsilon$

4 How to compare?

Compare the results with the ones that only use

$$L=\|\lambda_1-\hat{\lambda}_1\|^2+\|\lambda_2-\hat{\lambda}_2\|^2$$

to predict $\hat{\lambda}_{i,t+1}$ and directly get the weights of the tangency portfolio.