Methodology of Attribution Analysis

06/18/2020

1 Notations

The return of the stock "i" between time "t" and "t+1": $r_i(t) = \frac{P_i(t+1)-P_i(t)}{P_i(t)}$ where $P_i(t)$ is the price of stock "i" at time "t"

2 Decomposition of Stock Price Moves

For each time "t" the stock returns are decomposed using the factor loadings of the risk model as follows:

$$r_i(t) = \text{intercept}(t) + \sum_{k=1}^{NumberOfRiskFactors} F_{ik}\lambda_k(t) + \epsilon_i(t) \tag{1}$$

where F_{ik} - are risk factor loadings for stock "i" and risk factor "k", and $\lambda_k(t)$ is the return of the risk factor "k". In the equation above $\epsilon_i(t)$ is called the risk model residual. As a next step we decompose the residual $\epsilon_i(t)$ into alpha factors used in the model:

$$\epsilon_i(t) = \sum_{p=1}^{NumberOfAlphaFactors} \alpha_{ip}b_p(t) + \zeta_i(t) \tag{2}$$

where $b_p(t)$ is the return of the corresponding alpha factor α_{ip}. The coefficients b_p and λ_k and "intercept" are estimated from linear regressions.

3 Analysis of Portfolio PnL

The pnl of the stock "i" with position p_i can be written as:

$$\text{Pnl}_\text{i} = p_i(P_i(t+1) - P_i(t)) = \text{mv}_\text{i} r_i(t)$$

where "mv" is the market value of the position $mv_i(t) = p_iP_i(t)$ at time "t". Using equations (1) and (2) we decompose the pnl of the portfolio at time "t" into 3 different buckets:

$$RiskPnl = \sum_{i,k} F_{ik}\lambda_k(t)mv_i(t)$$

$$AlphaPnL = \sum_{i,p} \alpha_{ip}b_p(t)mv_i(t)$$

and the remainder so called residual pnl defined as $TotalPnl - (RiskPnl + AlphaPnl)$

The residual pnl of attribution typically can come from outliers or from factors missing in the original alpha model.

4 Outliers

We use the quantity called "Inverse Participation Ratio" (IPR) to evaluate the concentration of portfolio and to estimate the number of PnL outliers. For a vector with components w_i (where $i = 1....N$) the IPR is defined as:

$$\mathrm{IPR} = \frac{(\sum_{i=1}^{N} {w_i}^2)^2}{\sum_{i=1}^{N} {w_i}^4} \tag{3}$$

To get an intuition about this quantity consider 2 extreme cases:

4.1 Perfect Diversification

In this case all entries of the entries of the vector w_i are equal to the same number. Using equation (3) we will easily find that in this case $IPR = N$

4.2 Highly Concentrated Vector

Assume that $w_1 = C$, while all other entries are 0, i.e. $w_{i>1} = 0$, in this case $IPR = 1$

From 2 examples described above we can see that effectively IPR measures the number of stocks which dominate portfolio weights or portfolio PnL on a given date.