高清『无码』! 因子模型与机器学习

Original QIML编辑部 量化投资与机器学习 2021-08-13 16:25

收录于合集

#深度研读系列 21 #因子投资系列 51



量化投资与机器学习微信公众号,是业内垂直于**量化投资、对冲基金、Fintech、人工智能、大数据**等领域的主流自媒体。公众号拥有来自**公募、私募、券商、期货、银行、保险、高校**等行业**20W+**关注者,连续2年被腾讯云+社区评选为"年度最佳作者"。

量化投资与机器学习公众号独家解读

量化投资与机器学公众号 QIML Insight——深度研读系列 是公众号今年全力打造的一档深度、前沿、高水准栏目。



公众号<mark>遴选</mark>了各大期刊前沿论文,按照理解和提炼的方式为读者呈现每篇论文最精华的部分。QIML希望大家能够读到可以成长的量化文章,愿与你共同进步!

今天跟大家分享一篇机器学习用于因子模型的论文,一句话概括文章的内容:

作者首先利用了几类机器学习的模型对多个因子进行降维(或者称为特征提取),再利用处理后的因子结合组合优化Minimum-Variance的框架进行组合构建。

首先,对原始因子处理的模型使用了以下五大类,只有第5种是跟深度学习相关的,为了行文一致性,我们后文都称为机器学习模型。

- Principal component analysis
- Sparse principal component analysis
- Partial least squares
- Sparse partial least squares
- Autoencoders

利用以上机器学习模型对因子进行处理后,通过时间序列线性回归计算出因子暴露及特质收益项,再分别得出因子收益的协方差矩阵及特质收益矩阵,并通过这两个风险矩阵结合组合优化模型确定组合权重。

文章内容本身很长,但作者利用滚动窗口把数据分为Training、Validation及Out of Sample,并结合基于Minimum-Variance的损失函数,这个整体的框架及步骤很有借鉴意义。这里把整个框架的步骤梳理出来跟大家分享,其他实证性的内容,请大家参考原论文。

需要注意的是,作者在文中采用的是基于时序的因子模型,也就是说,因子收益是已知的,如Fama-French的因子,因子本身代表组合的收益率。并不是如Barra体系下,截面回归中的因子暴露。所以回归时,求解的是因子暴露。

Training: 192 Months

Validation: 48 Months

① 【Training】选取某个机器学习的模型,并使用某组参数,对<mark>某组</mark>因子X进行因子特 征提取,得到降维后的因子 X_tW^T

② 【Training】选取机器学习模型的某一组参数,使用Training的数据,计算出下列 式的 β 因子暴露及 α 截距 $r_{i,t} = a_i^T + \beta_i^T (\mathbf{X}_t \mathbf{W}^T) + u_{i,t}, \text{ for } t = 1, ..., T_0^T,$

③ 【Validation】使用步骤1的β、α及同样的参数,计算Validation集的协方差矩阵

 $\Sigma_r^{\mathcal{V}} = B^{\mathcal{T}'} \Sigma_f^{\mathcal{V}} B^{\mathcal{T}} + \Sigma_u^{\mathcal{V}},$

其中: Σ_f^{ν} 是应用步骤1的参数及机器学习模型,在Validation的因子数据进行

 $u_{i,t} = r_{i,t} - (a_i^T + \beta_i^T (X_t W^T)), \text{ for } t = 1, ..., T_0^V.$

③ 损失函数:使用Validation计算的协<mark>方差矩阵,计算最小方差,并不断调整步骤</mark> 1的参数,使损失函数最小的参数组,将用于样本外的数据

$$\mathcal{L}_{\text{MVP}}(\widehat{\omega}, \widehat{\Sigma}_r) = \widehat{\omega}' \widehat{\Sigma}_r \widehat{\omega}.$$

④ 步骤3得到的参数,应用到样本外的数据集中,计算组合的收益。 (公) 量化投资与机器学习

通过Training的数据求出因子暴露,并利用Training得到的因子暴露、截距及对应的机器学习模型参数,应用到验 证数据中求得的风险矩阵。

将验证数据中求得的风险矩阵,放到损失函数里,得到一组参数使用损失函数最小。这里的创新点就是巧妙的结合 了训练数据与验证数据及损失函数。这正是我们可以借鉴的部分。

需要 论文 的小伙伴,请在公众号后台回复:

收录于合集 #深度研读系列 21

く上一篇

下一篇〉

波动率目标策略,没有想象的那么简单!

基于Order Book的深度学习模型: 预测多时间段收益度例

People who liked this content also liked

北大满哥与奥迪的罗生门

量化投资与机器学习

