

广发基金陈少平:

# 机器学习赋能 构建多元量化体系

□本报记者 万宇



陈少平,工程学硕士。证券从业年限20年,投资管理年限17年。现任广发基金总经理助理,曾任北京海问投资公司研究员,泰达宏利基金经理、研究总监,嘉实基金公司基金经理、研究总监、联席CIO。

## 捕捉长期有效因子

**中国证券报:**近一两年,主动投资和量化投资逐渐融合,从你的研究体会来看,这两者各有什么特色?

**陈少平:**从投资方法看,量化与主观投资被认为是两种泾渭分明的方式。主观投资是基于主观判断和逻辑推演的基本面投资,而量化投资则是依赖于大数据分析、数理统计和自动化技术处理海量信息。

根据我对历史行情的观察,主动投资整体在大盘占优、成长占优的市场风格下表现更出色;量化产品则整体在小盘占优、价值占优时能获得较好表现。可以说,量化投资在风格和赛道上天然适合作为主动投资的补充。

**中国证券报:**相比主动投资,量化投资如何获取超额收益?

**陈少平:**相对来说,量化投资擅长融合来自不同维度的信息,凭借信息广度的优势从众多潜在的市场机会中捕捉到胜率更高的部分,迅速做出交易决策,从而获取长期稳健的超额收益。

在量化选股的过程中,传统因子用得越多,有效性就越差。因此,量化投资要做得好,核心是捕捉到市场上长期有效的因子,就像是挖矿一样,要求使用有逻辑、有规律、有基本面支撑的思路去设计策略。也就是说,量化投资的超额收益并不是来自模型或者算力的简单堆积,而是来自于

对金融数据的深刻理解。

我认为,量化投资要取得长期稳健的超额收益,需要做到以下三点:一是投研紧密配合,研究员在细分方向深入挖掘因子;二是不求极致,而是以稳致胜;三是要有耐心,积累长期业绩。

**中国证券报:**在具体实践中,广发基金量化投资部如何体现“对金融数据的深刻理解”?

**陈少平:**首先,我们坚持每一个因子和模型的设计都必须有逻辑、可解释,注重构建逻辑和金融学意义,对过度数据挖掘产生的因子保持谨慎,这样也会降低模型风险。

其次,我们坚持的是价值投资理念,对基本面赋予较高权重,在因子选用和模型设计上均侧重稳定性。例如,团队在机器学习模型的开发中,配置了较高比例的基本面因子,使得模型的中长期预测更为稳定,能够达到较高的胜率。

第三,通过对因子风格特征的钻研,我们开发了多种具有独创性和差异性的量化选股模型,超额信息来源涵盖基本面数据、交易面数据、另类数据等,为策略提供了长期而持续的生命力。基于策略的独创性和可解释性,在面对市场回撤时,团队能够保持冷静,深入剖析策略效果受到影响的原因和优化方法。

## 多维度提升综合能力

**中国证券报:**这些年,资管行业的竞争很激烈,你认为做好量化投资需要具备哪些核心能力?

**陈少平:**在量化投资中,数据源的多元化、低相关因子的充分储备、模型的多样性都非常重要,能够提升选股模型的准确度,更好地应对不同的市场环境,也是我们持续不断投入和完善的方向。

在数据源层面,我们自主设计了大数据平台,尽可能地追求更广的覆盖面,包括公司财报、分析师预期数据、财务附注、日频、分钟频等不同颗粒度的价量类数据。更多的数据来源,为我们寻找更多低相关的阿尔法提供了可能。

在因子库层面,我们对海量的市场数据进行收集、筛选、处理和提炼,并在此基础上形成了公司的特色因子库。2011年以来,团队自主研发因子超过2600个,因子储备充足,且非常重视对因子细节的打磨,大多数具有比较好的选股效果。

在模型的设计上,我们对机器学习模型的经典论文进行了深度研读,在模型的设计上针对量化投资做了深度的定制和改进,对模型的各项特征进行了广泛而深入的研究。目前选股模型和AI选股模型,实际有效开发代码超12万行。

**中国证券报:**在模型设计上,广发基金量化投资部有哪些迭代创新?

**陈少平:**经过近几年的探索,广发基金量化投资部从经典的多因子模型出发,逐

步进化到多因子模型、量化基本面模型和机器学习模型并行的多元量化体系。多元量化体系的构建,既有助于应对不同风格的市场,也能为不同类型的产品提供更匹配的策略选择。

目前团队已经积累了多个不同的因子合成模型,既包括传统线性模型,也有非线性的机器学习模型。在训练过程中,机器学习模型倾向于给予价量类因子更高权重,而线性模型的因子权重则较为均衡,这导致不同模型在不同的市场环境下具备一定的差异。通过多模型的方式,能在一定程度上平滑波动,提高超额收益的稳定性。

**中国证券报:**团队未来会在哪些方向进一步展开深入研究?

**陈少平:**未来团队探索的一个方向是对机器学习及高频因子做进一步的深度研究和探索。经过前期针对性的改造和适配,我们发现,引入机器学习模式及高频因子,对于公募量化业绩的提升有明显的效果,后续也将继续推进在量化平台上对机器学习模型及高频因子的深入设计。

另一个方向是对量化因子的持续挖掘和扩充,以及因子配置的不断优化。量化因子库是团队投资和研究的基础,根据市场情况不断从新的数据源挖掘新因子,以及在现有数据基础上改进和重构形成新因子,都是未来持续获取超额收益的重要来源。另外,团队将继续研究从宏观中观维度、因子微观维度以及AI赋能等多角度优化因子配权模型。

量化投资,是一种使用数据和模型算法来进行投资决策的投研方法。和主动选股以深度取胜不同,量化投资通常凭借广泛的覆盖和分析来总结普适性的规律。有两个因素至关重要:一是有多年经验的“船长”把握方向,知道哪里“有鱼”;二是有来自相关领域的专才,共同搭建高效的“捕鱼网”。

广发基金总经理助理陈少平,就是一位善于把握航行方向的“船长”。作为一名经验丰富的投研老将,她先后在研究员、研究总监、联席CIO、基金经理等岗位历练。如今,她凭借着善于研究和总结证券市场发展规律的优势,基于量化业务先进经验,带领广发基金量化部持续探索前沿的发展方向。

“通过研究量化业务的发展趋势,我发现高频因子、另类数据以及人工智能(AI)技术,已经在业内得到越来越广泛的应用。”陈少平介绍,广发基金量化投资部较早布局AI和机器学习方面,逐渐构建起多元量化体系。她相信,未来量化投资将更加智能化、精准化,团队也将继续向另类因子及机器学习的方向拓展。

## 去“有鱼”的地方

**中国证券报:**你在基本面研究方面深耕多年,分管量化投资后,过往的这些经验如何赋能量化团队?

**陈少平:**确实,我在主观研究方面做了很多年,在研究、总结市场的发展规律方面积累了一些经验,可以运用在量化研究上。一方面,为团队提供方向性的判断。从多年投研经验来看,我认为研究事物的发展规律很重要。因此,我深入研究了量化业务在海内外的发展路径,并结合自身在基本面研究上的积累,争取帮助团队把握大方向,就像芒格说的“去有鱼的地方捕鱼”。比如在年初,我建议团队要加大对于红利因子的研究力度,上半年红利资产表现优异,为我们的产品业绩做出了正向贡献。

另一方面,积极引进先进技术和人才。通过研究量化业务的趋势,我发现,高频因子、另类数据以及人工智能技术,已经在业内得到越来越广泛的应用。因此,广发基金量化投资部较早对AI和机器学习方面进行布局,并积极进行实践探索,为提升策略有效性带来了较好的效果。

**中国证券报:**根据你的研究,目前业内如何运用AI和机器学习技术?

**陈少平:**具体来看,目前业内运用机器学习技术的量化策略主要在于两个方面:一是协助挖掘因子,显著提升工作效能;二是整合因子打分,为制定策略权重提供支持。

不同团队在AI技术的细节处理上存在差异,包括模型设计、架构搭建、数据预处理和目标函数选择等方面,都需要深入研究和经验积累。在上述方面,广发基金量化投资部已经逐步积累了自身的优势。

**中国证券报:**能否展开说说,广发基金量化投资部如何在投研中体现先进性?

**陈少平:**一方面,团队在数据源方面引入了另类数据。随着技术的发展和进步,可获取的数据越来越丰富,另类数据的使用将极大提升模型的选股能力,充分利用人工智能技术和先进的计算架构,可以提升选股模型的预测准确度,充分拓展量化分析框架的广度和深度。

另一方面,在2023年以来的AI科技风口之下,团队对机器学习进行密切关注和深度研究。机器学习等技术更有助于挖掘一些非线性的因子,对于之前的技术来说,这些因子是比较难以捕捉的。此外,传统的因子复合方法往往采用线性加权的方式,而机器学习则采用非线性的“脑回路”,能提供更高效率的整合策略,有助于获得超额投资收益。