

## “中文云”金融文本量化研究平台在股票期权投资中的应用

数量化策略报告

### ◆ 内容提要:

2015年2月9日，A股市场第一只股票期权产品——50ETF期权即将上市交易，标志着A股市场正式进入期权交易时代。在国际成熟证券市场中，股票期权交易拥有高流动性、带杠杆、成本低、T+0和风险-收益的不对称性等多项特征，从而受到广大机构和个人投资者的青睐。

简单地说，股票期权能够实现的金融功能包括资产保值、单边投机和套利。股票期权的推出，将大大丰富A股市场的交易手段，从而提升市场交易的活跃程度，并且提升高频数据的重要性。这对传统的数据源提出了挑战，并且对期权参与者的技术准备提出了非常高的要求。

光大证券“中文云”金融文本量化研究平台（简称“中文云”），是目前业内最为成熟的金融文本研究系统。

从2010年6月到现在，“中文云”积累了36万篇以上的研究报告，300多万篇财经新闻以及2800万条以上论坛帖子，形成了业内最为完整的金融文本数据库储备。以高性能文本数据库为依托，光大金工团队开发出一系列量化策略，主要包括普通投资者情绪择时模型、关注度因子选股模型、概念套利模型、新闻套利模型，以及广义事件套利模型等。

中文云概念套利可以与各种期权交易策略相结合，极大地丰富交易手段和策略内容。比如，看多概念可以买入对应股票的看涨期权，从而在风险可控的前期下博取杠杆化的高收益。当股票期权推出之后，我们还可以根据股票的关注度因子及其波动率的历史关系挖掘规律，从而构建股票的期权交易模型。

### 分析师

冯 剑

(执业证书编号：S0930511090001)

021-22169176

fengj@ebscn.com

为带动国内金融文本研究，光大中文云系统现已开源。7人专业团队，6年专注研究，2万行代码，爬虫、索引、检索、统计、热度、选股一气呵成。共享地址：  
<https://github.com/dmquant/chinesecloud>

### 相关研报

期权策略操作指南——无风险套利篇  
(2014-03-12)

期权无风险套利 EXCEL 模板  
(2014-03-28)

个股期权套利模板  
(2014-05-16)

## 1、A 股进入期权交易时代

中国证监会昨日宣布，已批准上海证券交易所开展股票期权交易试点，试点产品为上证 50ETF 期权，正式上市交易日为 2015 年 2 月 9 日，标志着 A 股市场即将正式进入股票期权交易时代。

上海证券交易所同日公布《上海证券交易所股票期权试点投资者适当性管理指引》，对机构与个人投资者参与期权交易的条件进行了规定。其中个人投资者参与期权交易的条件包括：申请开户时托管在其委托的期权经营机构的证券市值与资金账户可用余额（不含通过融资融券交易融入的证券和资金），合计不低于人民币 50 万元；指定交易在证券公司 6 个月以上并具备融资融券业务参与资格或者金融期货交易经历；或者在期货公司开户 6 个月以上并具有金融期货交易经历。个人投资者申请的交易权限级别分为一级、二级、三级交易权限。

## 2、交易特征与金融功能

在国际成熟证券市场中，股票期权交易拥有高流动性、带杠杆、成本低、T+0 和风险-收益的不对称性等多项特征，从而受到广大机构和个人投资者的青睐。

股票期权具有的金融功能主要包括：

### 1、资产保值

期权产生是一种风险控制的金融工具。比如，买入某只股票的看跌期权，就相当于对投资者目前持有的该只股票多头仓位进行了套期保值；再如，买入某只股票的看涨期权，其成本或风险在交易时就实现了锁定（即期权费），而理论上的潜在收益趋向正无穷，这就是期权交易的风险-收益不对称性。

### 2、单边投机

股票期权具有杠杆，投资者只需付出小量的期权费，就有可能获得股票价格变动带来的潜在收益。随着期权策略的不同，股票价格无论是上涨还是下跌，投资者都将有可能获得收益。如前所述，由于期权的风险-收益不对称性，投资者的成本是可控的，这给确定看好某只股票的投资者提供了绝佳的投资工具，带有杠杆并且风险可控。

### 3、套利工具

与股指期货的期现套利和跨期套利类似，股票期权及其对应股票的价格之间存在着一定的对应关系，并且同一股票往往对应着多个期权产品，因此同样可以实现套利。从股指期货历史上看，在其开始交易的三个月时间内，往往具有较大的套利空间。其后，随着投资者成熟度的不断提高，套利空间会缩小。因此股票期权开始交易后的一定时间内，同样有可能具有一定程度的套利空间，甚至有可能带来丰厚的套利收益。

## 三、股票期权对 A 股市场的影响

股票期权的推出，将大大丰富 A 股市场的交易手段，从而提升市场交易的活跃程度。

具体表现在：

首先，股票期权支持多空双边交易。一直以来，A 股市场的股票都没有一个方便、快捷的做空机制。两融推出后，一定程度上弥补了这一不足，但是由于券源有限，还未能根本解决这一问题。而股票期权的推出，将大大丰富 A 股市场的做空机制。

其次，股票期权支持 T+0 交易。T+0 交易本身就将极大地提高市场的活跃程度，同时也可能带来日内的高波动发生。这对原有的日内交易策略提出了新的挑战，新的金融工具带来新规则，市场的不确定性可能增加。

第三，高频数据重要性大大提升

T+0，高杠杆，期权交易的战场有可能比股票激烈十倍。在期权市场上，传统的低频数据的可用性被大大降低。对于量化研究同样是如此，传统的日频甚至是周频、月频的数据，对于股票期权交易的重要性远远不如高频数据。这对传统的数据源提出了挑战，并且对期权参与者的技术准备提出了非常高的要求。

## 四、中文云与期权策略

### 1、“中文云”简介

光大证券“中文云”金融文本量化研究平台(简称“中文云”),是目前业内最为成熟的金融文本研究系统。“中文云”以光大金工团队独有的互联网文本挖掘技术为基础,包括三大模块,即数据获取模块,数据引擎模块,以及策略开发模块。为带动国内金融文本研究,“中文云 1.0”已于 2014 年 8 月 30 日正式开源,供同行学习参考。共享地址如下:

<https://github.com/dmquant/chinesecloud>

“中文云”平台由网络爬虫逐日从特定信息源自动获取文本信息,形成底层文本数据库。从 2010 年 6 月到现在,“中文云”积累了 36 万篇以上的研究报告,300 多万篇财经新闻以及 2800 万条以上论坛帖子,形成了业内最为完整的金融文本数据库储备。以高性能文本数据库为依托,光大金工团队开发出一系列量化策略,主要包括普通投资者情绪择时模型、关注度因子选股模型、概念套利模型、新闻套利模型,以及广义事件套利模型等。

我们还向市场提供个股的关注度系列因子产品,目前包括关注度因子、关注度变化因子和关注度历史位置因子。历史回测显示,无论是在选股,还是在行业配置方面,关注度系列因子都体现出很强的预测能力,是非常优异的因子。

股票期权推出后,概念热度和股票关注度以其对股票价格的高敏感性和预测能力,能够对股票期权策略提供支持。

### 2、概念热度与波动率

追逐热点是资金的天性,也是投资研究的永恒主题。进入互联网时代以后,随着信息化的高度发达,概念热点已经成为股市投资不可或缺的基本要素。2013 年 3 月份的环保概念、9 月的上海自贸区概念、2014 年上半年的军工概念、航母概念都成为引领市场普遍关注和资金广泛参与的焦点,而彩票、地图、苹果、世界杯、工业 4.0、一带一路等各种的概念更是此起彼伏,不断创造新的投资机会。

与技术指标、基本面等传统投资不同,概念投资的独特性在于,其促发因素往往是国家战略、领导讲话、财经新闻和国内外重大事件。换句话说,概念投资所依赖的信息源是文本信息,而不是传统的数字信息。这实际上就是概念投资量化的根本性难点,因为现有量化模型都是基于数字信息的,文本处理方面几乎是空白。“中文云”系统应运而生,并且成为目前业内最为成熟的金融文本研究系统。

“中文云”系统目前概念数量达到 119 个，基本上涵盖了市场上存在的全部概念。新概念由系统和人工两种方式共同维护，及时添加。

我们用“概念热度”这一统计指标，来刻画概念被市场关注的程度。将概念这种文本信息转化成为概念热度这一量化指标的核心方法是文本信息处理的核心算法——词频统计。之前还会进行数据清洗，后续还有进一步的统计处理，篇幅所限从略。

跟行业轮动相似，概念之间也存在着轮动。并且这种轮动的方向、频率比行业轮动还要复杂。“中文云”根据概念热度进行概念轮动套利。简单地说，当概念热度上升，但股价未有表现的，看多；反之，当概念热度回落，股价横盘或回落，且两者均持续 5 日以上的，看空；此外，当概念热度高位回落，股价高位回落 40% 左右，且两都持续低位，可参与概念第二波。

大部分概念都没有直接标的，而需要进一步对应到个股。传统的个股匹配方法比较单一，即人工匹配。其缺点一是速度慢；二是如果概念过于抽象则很难匹配，如工业 4.0 概念；三是，匹配结果因人而异，带有一定主观性。

“中文云”系统采用创新性的文本相关性算法，以逻辑文本信息源为基础，对概念及其对应个股在邻近文档中出现的概率进行估算，从而精确地给出概念与个股的匹配度，完成概念到个股的匹配。任何概念，“中文云”系统均可以在 10 秒以内给出 10 至 20 支对应股票。

具体至股票期权交易方面，中文云概念套利可以与各种期权交易策略相结合，极大地丰富交易手段和策略内容。比如，看多概念可以买入对应股票的看涨期权，从而在风险可控的前期下博取杠杆化的高收益。

### 3、股票关注度与波动率

众所周知，在量化投资中很大一部分是  $\alpha$  多因子产品。传统因子尽管数量众多，但均属于数字信息源。实际上投资者在作出决策时，往往同时会考虑文本信息。如果在多因子模型中仅仅依赖数字信息源，而没有文本信息源，那么就会造成信息的缺失，最终导致  $\alpha$  的缺失。

“中文云”所提出的关注度因子，就是基于个股文本信息提炼出的有效因子，其周频和月频 IC 均达到 10%，周频多空组合收益率达 1.5%。

当股票期权推出之后，我们还可以根据股票的关注度因子及其波动率的历史关系挖掘规律，从而构建股票的期权交易模型。

## 分析师声明

负责准备本报告以及撰写本报告的所有研究分析师或工作人员在此保证，本研究报告中关于任何发行商或证券所发表的观点均如实反映分析人员的个人观点。负责准备本报告的分析师获取报酬的评判因素包括研究的质量和准确性、客户的反馈、竞争性因素以及光大证券股份有限公司的整体收益。所有研究分析师或工作人员保证他们报酬的任何一部分不曾与，不与，也将不会与本报告中具体的推荐意见或观点有直接或间接的联系。

## 分析师介绍

冯剑，光大证券研究所首席金融工程分析师，团队负责人；复旦大学经济学博士，8年经济金融研究经验，4年证券研究经验，主要带领团队从事“中文云”金融文本量化研究平台策略开发。

## 行业及公司评级体系

买入—未来 6-12 个月的投资收益率领先市场基准指数 15%以上；

增持—未来 6-12 个月的投资收益率领先市场基准指数 5%至 15%；

中性—未来 6-12 个月的投资收益率与市场基准指数的变动幅度相差-5%至 5%；

减持—未来 6-12 个月的投资收益率落后市场基准指数 5%至 15%；

卖出—未来 6-12 个月的投资收益率落后市场基准指数 15%以上；

无评级—因无法获取必要的资料，或者公司面临无法预见结果的重大不确定性事件，或者其他原因，致使无法给出明确的投资评级。

市场基准指数为沪深 300 指数。

## 分析、估值方法的局限性说明

本报告所包含的分析基于各种假设，不同假设可能导致分析结果出现重大不同。本报告采用的各种估值方法及模型均有其局限性，估值结果不保证所涉及证券能够在该价格交易。

## 特别声明

光大证券股份有限公司（以下简称“本公司”）创建于1996年，系由中国光大（集团）总公司投资控股的全国性综合类股份制证券公司，是中国证监会批准的首批三家创新试点公司之一。公司经营业务许可证编号：z22831000。

公司经营范围：证券经纪；证券投资咨询；与证券交易、证券投资活动有关的财务顾问；证券承销与保荐；证券自营；为期货公司提供中间介绍业务；证券投资基金代销；融资融券业务；中国证监会批准的其他业务。此外，公司还通过全资或控股子公司开展资产管理、直接投资、期货、基金管理以及香港证券业务。

本证券研究报告由光大证券股份有限公司研究所（以下简称“光大证券研究所”）编写，以合法获得的我们相信为可靠、准确、完整的信息为基础，但不保证我们所获得的原始信息以及报告所载信息之准确性和完整性。光大证券研究所可能将不时补充、修订或更新有关信息，但不保证及时发布该等更新。

本报告根据中华人民共和国法律在中华人民共和国境内分发，仅供本公司的客户使用。

本报告中的资料、意见、预测均反映报告初次发布时光大证券研究所的判断，可能需随时进行调整。报告中的信息或所表达的意见不构成任何投资、法律、会计或税务方面的最终操作建议，本公司不就任何人依据报告中的内容而最终操作建议作出任何形式的保证和承诺。

在法律允许的情况下，本公司及其附属机构可能持有报告中提及的公司所发行证券的头寸并进行交易，也可能为这些公司提供或正在争取提供投资银行、财务顾问或金融产品等相关服务。投资者应当充分考虑本公司及本公司附属机构就报告内容可能存在的利益冲突，不应视本报告为作出投资决策的唯一参考因素。

在任何情况下，本报告中的信息或所表达的建议并不构成对任何投资人的投资建议，本公司及其附属机构（包括光大证券研究所）不对投资者买卖有关公司股份而产生的盈亏承担责任。

本公司的销售人员、交易人员和其他专业人员可能会向客户提供与本报告中观点不同的口头或书面评论或交易策略。本公司的资产管理部和投资业务部可能会作出与本报告的推荐不相一致的投资决策。本公司提醒投资者注意并理解投资证券及投资产品存在的风险，在作出投资决策前，建议投资者务必向专业人士咨询并谨慎抉择。

本报告的版权仅归本公司所有，任何机构和个人未经书面许可不得以任何形式翻版、复制、刊登、发表、篡改或者引用。

## 光大证券股份有限公司研究所 销售交易部 机构业务部

上海市新闻路1508号静安国际广场3楼 邮编200040

总机：021-22169999 传真：021-22169114、22169134

销售交易总部	姓名	办公电话	手机	电子邮件
北京	郝辉	010-58452028	-	haohui@ebsecn.com
	黄怡	010-58452027	-	huangyi@ebsecn.com
	梁晨	-	-	liangchen@ebsecn.com
	刘公直	010-58452029	-	liugongzhi@ebsecn.com
	严非	021-22169086	-	yanfei@ebsecn.com
上海	周薇薇	021-22169087	-	zhouww1@ebsecn.com
	徐又丰	021-22169082	-	xuyf@ebsecn.com
	李强	021-22169131	-	liqiang88@ebsecn.com
	张弓	021-22169083	-	zhanggong@ebsecn.com
	罗德锦	021-22169146	-	luodj@ebsecn.com
深圳	黎晓宇	0755-83553559	-	lixyl@ebsecn.com
	李潇	0755-83559378	-	lixiao1@ebsecn.com
	张亦潇	0755-23996409	-	zhangyx@ebsecn.com
	王渊锋	-	-	wangyuanfeng@ebsecn.com
	濮维娜(总经理)	021-62152373	13611990668	puwn@ebsecn.com
上海	计爽	021-22167101	18621181721	jishuang@ebsecn.com
北京	朱林	010-59046212	18611386181	zhulin1@ebsecn.com
国际业务	陶奕(副总经理)	021-62152393	18018609199	taoyi@ebsecn.com
	戚德文(执行董事)	021-22169152	13585893550	qidw@ebsecn.com
	顾胜寒	021-22167094	18352760578	gush@ebsecn.com