

A Method of Finding Stock Buying Point Based on Moving Average Price and Big Data Analysis

Xiaoxiao Xu, Lizhu Wang, Congdian Cheng

College of Mathematics and Systems Science, Shenyang Normal University, Shenyang Liaoning
Email: xuxiaoxiaosax@163.com, wanglizhu80@126.com, zhiyang918@163.com

Received: Dec. 9th, 2019; accepted: Dec. 26th, 2019; published: Jan. 2nd, 2020

Abstract

Considering that in a certain period of time, how to solve the problem of buying and selling stocks to achieve expected profits in a large probability. An algorithm based on moving average and big data operation to find the stock buying point is established, and its efficiency is tested by experiment. The work we have done can promote the development of stock market operation, stock market research and quantitative technology in China.

Keywords

Average Line, Stock Price, Data, Algorithm, Experiment

一种基于移动平均价和大数据分析的寻找股票买入点的方法

徐峭笑, 王立柱, 程丛电

沈阳师范大学数学与系统科学学院, 辽宁 沈阳
Email: xuxiaoxiaosax@163.com, wanglizhu80@126.com, zhiyang918@163.com

收稿日期: 2019年12月9日; 录用日期: 2019年12月26日; 发布日期: 2020年1月2日

摘要

考虑在一定时期内, 大概率实现预期盈利的买卖股票问题。建立一种基于移动平均线和大数据运算寻找买入点的算法, 并通过实验检验其效率。所做工作可促进我国股市运作, 股市研究和量化技术的发展。

文章引用: 徐峭笑, 王立柱, 程丛电. 一种基于移动平均价和大数据分析的寻找股票买入点的方法[J]. 统计学与应用, 2020, 9(1): 1-6. DOI: 10.12677/sa.2020.91001

关键词

均线, 股价, 数据, 算法, 实验

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

中国股票市场自其产生以来, 经过三十年来的发展目前在我国经济发展中有着举足轻重的地位, 股票市场有拓宽企业融资渠道、解决企业生存发展问题, 为企业融资开辟平台的作用, 同时炒股也可为广大投资者分享企业利润提供一个平台, 也可培养个人艰苦奋斗的精神。总之, 股市投资是国家经济发展的重要途径之一, 也是国民致富的重要途径之一。因此, 股市受到了广泛的关注与研究。例如潘水洋等[1]展示了运用神经网络模型对股票的收益率进行预测的方法。王勇[2]探讨了用布林线指标把握短期的股票买卖点问题。熊伟等[3]研究了一种采用重标度极差分析法对股票的收益率进行长程相关性分析的方法。余万林等[4]考察了货币供给、房地产价格与股票估值波动之间的关系。高竹楼[5]探讨了在股市操作中如何应用均线的问题。冷松和俞雪华[6]给出了一种股票短线启动时机选取的方法并讨论了如何确定其买卖点的问题。

移动均线具有简单、有效、稳定等特点是当今预测股价走势等最重要的技术方法, 如何确定买卖点是股市投资者最关心的问题。如果要在股市中大概率的赚取利润, 就不能不研究这一问题。现已有不少学者从移动均线的角度出发研究股市与确定炒股买卖点问题。例如, 萌生[7]研究了股价走势曲线与移动均线距离的远近获得短期买卖点问题。黄莉等[8]利用短期、中期、长期移动均线之间的交叉建立数学模型, 再结合收益率的相关性, 找到最佳收益的移动均线组合问题。董大勇[9]通过对短期、中期、长期移动均线的位置、位置关系和运动方向加以分析来预测股价的变化趋势。此外, 文献[10]-[15]也从各种均线的角度出发研究了此类问题。受到摄取、存储水平和计算水平的限制, 这些工作所用的数据量相对不足, 所作运算量相对不足, 实用性相对不足。

鉴于上述背景, 本文试基于移动均线, 给出一种通过大数据分析寻找大概率盈利的炒股方法。

2. 数学规划

本节进行符号说明与做必要的定义。

给定一支股票 S , 我们分别以 $P_i^o, P_i^c, P_i^l, P_i^h$ 表示交易日 i 的开盘价, 收盘价, 最低价和最高价。令

$\bar{P}_i = \frac{1}{4}(P_i^o + P_i^c + P_i^l + P_i^h)$, 称其为 i 交易日的四点平均价, 或 F 平均价, 简称 i 交易日的均价。称

$$\bar{P}(t, l) = \frac{1}{l} \sum_{i=1}^l \bar{P}_{t+i}。$$

为 t 交易日的后 l 日 F 平均价, 其中 $t+i$ 表示交易日 t 之前的第 i 个交易日。

在 t 交易日, 购买一股股票, 用 $O(t, f)$ 表示预期往后 f 日的盈利值, 叫做未来 f 日的盈利目标。称考虑如何大概率实现盈利 $O(t, f)$ 的问题为 S-RFPP (f), 简称 RFPP。

下面, 我们以算法的形式给出一种大概率实现未来 f 日的盈利目标的买股方法。

注 1: 由于我们主要是考虑买入后的成功率问题, 简明起见, 所有提及买入均按买入一股理解, 对于多量买入的情况, 可通过乘以买入量系数的方法进行调整处理。

3. 算法

本节建立 RFPO 算法, 通过该算法给出一种争取在未来 f 日内实现预期盈利目标的买股方法。

股民买股票时都希望一定时期内大概率盈利。当价格一定程度低于某种平均值(或说处于某种均线下方)时买入是寻求大概率盈利的一种主要方法, 低少了不利于盈利, 低多了又不容易买到, 究竟以低多少为确定买入点的标准呢? 为了提高收益率, 我们不仅应当考虑买到的成功率, 还应当考虑买到的可能性, 这既是一个具有实际意义, 很值得研究的问题, 又是一个相当复杂, 很不容易研究的问题。为了降低难度, 并利于实施量化技术(通过数量化方式及计算机程序化发出买卖指令, 以获取稳定收益为目的的交易方式), 我们在一定条件下进行研究。限定问题 RFPP, 并假设期望买到后以概率 r 盈利, 且只在 $\bar{P}(t, l)$ 值之下买入。我们试探索能够使得买到后至少盈利 $O(t, f)$ 的概率大于 r , 且买到的可能性较大的方法。设 $O(t, f) = v$, $P_i^o, P_i^c, P_i^l, P_i^h, i = 1, 2, \dots, L + f + 1$, 为最近的 $(L + f + 1)$ 个历史数据, 并确定在当前价低于 $(\bar{P}(t, l) - \delta)$, $\delta > 0$ 时可以买入, 对 $t = f + 1, f + 2, \dots, f + L$ 这 L 个交易日进行复盘, 以 P_i^l 代替当前价, 试看有多少个交易日可以买到, 对于买到的交易日又能够有多少个交易日可以成功盈利 v 。设有 I 个交易日可以买到, 有 N 个交易日可以成功, 且过程具有充分的平稳性, 则当 $\frac{N}{I} > r$ 时, 可以认为在当前按 $(\bar{P}(t, l) - \delta)$ 价格买入股票可获得盈利 v 的概率大于 r 。如何找到满足条件的 δ 呢? 我们以一个适当的 ε 做为间隔尺度, 依次按 δ 为 $\varepsilon, 2\varepsilon, \dots$, 通过实际运算进行寻找, 且为了防止由于 δ 过大而导致买入的可能性太小, 我们以首先满足 $\frac{N}{I} > r$ 的 δ 作为寻找对象。这就是本文寻找买入点的主导思想。

显然 L, l 过小, 数据量过小, 不能充分体现过程的内涵; 另外, 运算速度太慢不能及时给出答案也不行, 所以 L, l 要充分大, 运算速度要充分快, 也就是说我们必须从大数据的视域出发, 做到快速摄取大量数据, 快速处理大量数据。为此我们根据上述主导思想, 设计了如下的算法 RFPO, 以其给出我们的方法。

RFPO 算法

问题: S-RFPP (f)。

输入: $f, l, r, L, v, \varepsilon, P_i^o, P_i^c, P_i^l, P_i^h, i = 1, 2, \dots, L + f + 1; p$ 。

(其中 $f, l, r, L, v, \varepsilon > 0$; f 为正整数, 表示未来日数; l 为正整数, 表示 $\bar{P}(t, l)$ 中的交易日数; r 是一个表示成功率的参变量; $v = O(t, f)$; ε 是一个表示运行尺度的参变量; $P_i^o, P_i^c, P_i^l, P_i^h$ 分别表示交易日 i 的开盘价、收盘价、最低价和最高价; $L + f + 1$ 为运用历史数据的个数; p 为当前股价; 1 表示当前交易日, 2 表示 1 的前一个交易日, \dots , $L + f + 1$ 表示 $L + f + 1 - 1$ 的前一个交易日)。

输出: Y 或 N (其中, Y 表示可以按当前股价 p 购买股票; N 表示不可以购买股票)。

过程:

1. 命 $\delta := 0$ 。
2. 置 $\delta := \delta + \varepsilon$; $I := 0$; $N := 0$ 。
3. 关于 $t = f + 1, f + 2, \dots, f + L$, 进行下面各步。
 - (1) 计算 $\bar{P}(t, l)$ 。
 - (2) 如 $P_i^l < \bar{P}(t, l) - \delta$, 置 $I := I + 1$; 然后考虑

$$\max \{ P_i^h : i = t - 1, t - 2, \dots, t - f \} > \bar{P}(t, l) - \delta + v,$$

不等式成立, 置 $N := N + 1$ 。

4. (1) $I = 0$ 。转步 5。

(2) (i) $\frac{N}{I} \leq r$ 。返回第二步。

(ii) $\frac{N}{I} > r$ 。转步 6。

5. 输出 N 。

6. (1) 计算 $\bar{P}(0, l)$, (0 表示当前交易日)。

(2) (i) $p > \bar{P}(0, l) - \delta$ 。输出 N 。

(ii) $p \leq \bar{P}(0, l) - \delta$ 。输出 Y 。

4. 实验

针对沪市股票中船科技(600072), 选取 2015 年 12 月 24 日至 2016 年 11 月 25 日这期间的 220 个交易日(去掉周六周日和法定节假日)每日的开盘价、收盘价、最高价、最低价为实验数据, 取 $l = 40$, $L = 90$, $f = 30$, $r = 0.6$, $v = 1.4$, $\varepsilon = 0.1$, 依次以 2016 年 8 月 25 日至 2016 年 10 月 14 日这期间的 30 个交易日为当前时刻, 进行实验。在实验中以当前日的最低价为当前价。

1. 以 2016 年 8 月 25 日为当前交易日 0 时, $i = 1$ (2016 年 8 月 24 日), 2 (2016 年 8 月 23 日), ..., $(L + f + l)$ (2015 年 12 月 24 日), 将数据和参量代入算法中, 得到结果: 可以买入, 不成功;

2. 以 2016 年 8 月 26 日为当前交易日 0 时, $i = 1$ (2016 年 8 月 25 日), 2 (2016 年 8 月 24 日), ..., $(L + f + l)$ (2015 年 12 月 25 日), 将数据和参量代入算法中, 得到结果: 可以买入, 不成功;

.....

30. 以 2016 年 10 月 14 日为当前交易日 0 时, $i = 1$ (2016 年 10 月 13 日), 2 (2016 年 10 月 12 日), ..., $(L + f + l)$ (2016 年 2 月 4 日), 将数据和参量代入算法中, 得到结果: 不可买入。

详细结果见表 1。

由表 1 可见, 在 30 个实验交易日中, 有 20 个交易日可以买入, 在可以买入的交易日中有 18 个成功, 成功率达 60% 以上, 可见算法有效。

注 2: 1) 我们的实验是通过运用 MATLAB 语言编出 RFPO 算法的计算机运行程序, 然后在一台型号为 Inspiron1427FT02、CPU 为 Intel(R) Core(TM)2 Duo CPU T6500 @2.10GHz 2.10GHz 的计算机上运行所编出的程序进行大数据运算而完成的。2) 算法 RFPO 不仅参变量较多, 而且可以从炒股软件中快速提取大量数据, 快速进行运算, 及时回答我们是否可以按当前价买入股票的问题, 充分地体现了大数据的特点。3) 由于时间和篇幅等原因, 在此我们暂未讨论怎样选择合适的参量 r , 实验是在 $r = 0.6$ 的条件下进行的, 如何优化地确定 r 是一个很值得继续研究的问题。4) 为了降低难度和便于使用量化技术, 算法 RFPO 的主要意图限于发展一种买入股票后可大概率盈利, 且买入可能性较大的方法, 没有考虑不成功时的盈亏问题; 如何在不成功时控制盈亏, 使得尽量少亏损多盈利, 并进而考虑如何使用资金, 争取较大收益率等, 也是很值得继续研究的问题。

5. 结束语

本文尝试地建立了一种基于大数据与平均线的选股方法, 并通过实验说明了它的效率, 运用沪市股票中船科技(600072) 2015 年至 2016 年的 220 个交易日的的数据, 进行实验, 结果表明算法有效。由于水平有限, 不足之处很多, 敬望各位专家批评指正。如何优化的确定 r , 如何优化的运用资金, 以及收益率如

何等是很值得进一步研究的问题。希望该项工作能够激发大家关于这方面的研究兴趣，希望能够有更多的学者和我们一起从事这方面的研究，共同促进我国量化技术发展。

Table 1. Summary of experiment result

表 1. 实验结果汇集表

当前日期	可否买入	是否成功	当前日期	可否买入	是否成功
2016.8.25	可以买入	不成功	2016.9.19	可以买入	成功
2016.8.26	可以买入	不成功	2016.9.20	可以买入	成功
2016.8.29	可以买入	成功	2016.9.21	不可买入	——
2016.8.30	可以买入	成功	2016.9.22	不可买入	——
2016.8.31	可以买入	成功	2016.9.23	可以买入	成功
2016.9.1	可以买入	成功	2016.9.26	可以买入	成功
2016.9.2	可以买入	成功	2016.9.27	可以买入	成功
2016.9.5	可以买入	成功	2016.9.28	可以买入	成功
2016.9.6	可以买入	成功	2016.9.29	可以买入	成功
2016.9.7	可以买入	成功	2016.9.30	不可买入	——
2016.9.8	不可买入	——	2016.10.10	不可买入	——
2016.9.9	不可买入	——	2016.10.11	不可买入	——
2016.9.12	可以买入	成功	2016.10.12	不可买入	——
2016.9.13	可以买入	成功	2016.10.13	不可买入	——
2016.9.14	可以买入	成功	2016.10.14	不可买入	——

基金项目

辽宁省自然科学基金资助项目(20170540821)。

致 谢

论文最后感谢辽宁省自然科学基金的资助，感谢程丛电导师、王立柱老师的指导和审稿人提出的建议。

参考文献

- [1] 潘水洋, 刘俊玮, 王一鸣. 基于神经网络的股票收益率预测研究[J]. 浙江大学学报, 2019, 46(5): 550-555.
- [2] 王勇. 学会用布林线法把握股票的买卖点[J]. 金融经济, 2006(12): 41.
- [3] 熊伟, 霍玉洪, 胡茂林. 中国股票市场收益率的 R/S 分析[J]. 牡丹江师范学院学报(自然科学版), 2007(3): 1-3.
- [4] 余万林, 裴延华, 时敏. 货币供给、房地产价格与股票估值波动[J]. 山东理工大学学报(社会科学版), 2019, 35(4): 5-12.
- [5] 高竹楼. 在股市操作中如何应用平均线[J]. 南京股友之家, 2004(12): 41-42.
- [6] 冷松, 俞雪华. 股票短线启动时机选择及其买卖点的确定探讨[J]. 扬州教育学院学报, 2009, 27(4): 34-37.
- [7] 萌生. 利用移动平均线预期买卖点[J]. 股市动态分析, 2007(17): 33.
- [8] 黄莉, 李秀茹, 廖鸿余, 陈瑞兴. MA 线在股市中的应用[J]. 南阳理工学院学报, 2011, 3(1): 125-128.
- [9] 董大勇. 均线指标组合上下的上证指数收益实证分析[J]. 孝感职业技术学院学报, 2011, 6(1): 76-80.

- [10] 宋东宇, 王立柱. 移动平均交易规则对股票动态价格模型的影响[J]. 牡丹江师范学院学报, 2018(1): 7-9.
- [11] 姚宏亮, 艾刘可, 王浩, 李俊照. 均线滞后的时序自回归股市态势预测算法[J]. 郑州大学学报(理学版), 2018, 50(3): 60-66.
- [12] 陈珂, 陈伟. 单均线交易策略在 A 股的实证研究[J]. 时代金融, 2016(4): 129-132.
- [13] 宋海峰. 论均线在股市中的应用[J]. 现代商贸工业, 2016(31): 106-107.
- [14] 江瑞侠, 于荣格, 张泽浩. 股票均线几种特性在预示后市中的作用[J]. 中外企业家, 2013(2): 50-51.
- [15] 骆鹏飞. 基于个人投资者的移动平均线理论应用——以上证指数 1A0001 为例[J]. 山东商业职业技术学院学报, 2012, 4(12): 31-34.