

中国人民大学本科毕业论文（设计）

中国期货市场技术分析有效性检验—— 基于动态时间扭曲(Dynamic Time Warping)算法

作者： 胡志祥
学院： 财政金融学院
专业： 金融工程
年级： 2014 级
学号： 2014201414
指导教师： 郭彪
论文成绩： A
日期： 2018 年 3 月 28 日

摘要

本文利用数据挖掘技术中用于比较时间序列相似度的工具动态时间扭曲 (Dynamic Time Warping, DTW) 算法对技术分析在中国期货市场的有效性进行了实证检验。本文在 TsinasInaidis (2018) 的研究框架的基础上构建了一个基于价格和成交量序列匹配的自动化交易策略—Subsequence Dynamic Time Warping (SDTW) 策略, 并通过使用中国期货市场 3 个股指期货合约的行情数据, 将其表现与构建的单向市场策略在上涨、下跌和震荡三种不同市场行情下进行了实证对比。通过对 SDTW 策略的表现在 5 个参数的不同取值下进行敏感性分析与最优化拟合, 本文得出了 SDTW 策略 (或技术分析) 的有效性与用于回测的合约品种挂钩且 SDTW 策略相对而言更能够识别出市场中下跌行情的结论。希望这些策略和结论可以为诸多期货市场投机、套保者提供一种新思路与方法。

关键词: 动态时间扭曲 技术分析 自动化交易

Abstract

This paper conducted an empirical test on the effectiveness of technical analysis in the Chinese futures market. Based on the major research framework of Tsinaslnaidis (2018), we developed an automated trading strategy named SDTW strategy. We then utilized the quotes data from 3 stock index futures contracts and conducted an empirical validation on the profitability of this strategy. After this, we conducted some sensitivity analysis and optimization on the performance of the strategy w.r.t. a wide range of different values of 5 input parameters. Based on all the results we got, we reached to the conclusion that SDTW strategy (or technical analysis) is more capable of identifying signs of bearish market and its performance varies a lot on different contracts. We hope these strategies and conclusions could provide some inspirations for the practitioners in the Chinese futures market.

Key Words: Dynamic Time Warping Technical Analysis Automated Trading Strategy

目录

| | |
|---|----|
| 1 引言..... | 1 |
| 2 技术分析简介..... | 3 |
| 2.1 技术分析基本原则..... | 3 |
| 2.1.1 Market Action Discounts Everything..... | 3 |
| 2.2.2 Prices Move in trends: | 4 |
| 2.2.3 History tends to be repeated: | 5 |
| 2.2 本文所用技术分析..... | 6 |
| 3 Dynamic Time Warping (DTW) 算法及其延伸..... | 8 |
| 3.1 DTW 算法原理..... | 8 |
| 3.2 Subsequence DTW (SDTW) 算法..... | 11 |
| 3.3 SDTW 算法与技术分析..... | 13 |
| 4 实证分析..... | 14 |
| 4.1 数据描述..... | 14 |
| 4.2 参数与假设..... | 14 |
| 4.3 交易策略..... | 14 |
| 4.4 结果分析..... | 18 |
| 4.5 敏感性分析..... | 22 |
| 5 结论..... | 27 |
| 6 参考文献..... | 28 |
| 7 附录..... | 30 |

图表

| | |
|--|----|
| 图 2.1 钾肥指数（申万）2018 年 03 月 09 日 Belaruskali 公司坍塌事故的冲击 | 4 |
| 图 2.2 上证综指走势 | 5 |
| 图 2.3 几种常见的技术图像：头肩底型、双顶型、旗型、三角旗型 | 6 |
| 图 2.4 使用技术分析的逻辑 | 7 |
| 图 3.1 不等长的序列 P 与 Q | 9 |
| 图 3.2 序列 P 与 Q 之间的距离矩阵 | 10 |
| 图 3.3 输出结果 | 11 |
| 图 4.1 价格序列：从左至右分别对应 $HS300$ 、 $SZ50$ 及 $ZZ500$ | 19 |
| 图 4.2 上涨行情模拟数据 | 20 |
| 图 4.3 震荡行情敏感性分析： RW & pre_RW | 22 |
| 图 4.4 震荡行情敏感性分析： NPQ & NVQ | 23 |
| 图 4.5 上涨行情敏感性分析： RW & pre_RW | 23 |
| 图 4.6 上涨行情敏感性分析： NPQ & NVQ | 24 |
| 图 4.7 下跌行情敏感性分析： RW & pre_RW | 24 |
| 图 4.8 下跌行情敏感性分析： NPQ & NVQ | 24 |
| 图 4.9 真实行情敏感性分析： RW & pre_RW | 25 |
| 图 4.10 真实行情敏感性分析： NPQ & NVQ | 25 |
| 表 3.1 $SDTW$ 算法 | 12 |
| 表 4.1 $SDTW$ 策略 | 15 |
| 表 4.2 市场多头策略 | 16 |

| | |
|--------------------------|----|
| 表 4.3 市场空头策略..... | 17 |
| 表 4.4 回测参数..... | 18 |
| 表 4.5 交易参数..... | 18 |
| 表 4.6 回测收益统计表..... | 18 |
| 表 4.7 回测收益率统计表：上涨行情..... | 20 |
| 表 4.8 回测收益率统计表：下跌行情..... | 21 |
| 表 4.9 回测收益率统计表：震荡行情..... | 21 |
| 表 4.10 最优测试结果一览..... | 26 |

1 引言

期货市场作为一个投机与套保功能并存的市场，为诸多投资者与对冲者提供了一个优越的平台。近期随着中美关税问题及贸易战的升温，诸多农产品、金属原料及利率的价格面临更大的不确定性。因此当今利用好期货市场及时做好相应的套期保值工作以对冲产品价格大涨大跌的风险显得至关重要。技术分析是诸多期货从业人员的必备基础知识与有利工具，有效地利用技术分析能给投资者和套保者带来更好的投资/套保绩效。本文使用了数据挖掘技术中的动态时间扭曲（Dynamic Time Warping, DTW）技术对技术分析在中国期货市场中的有效性进行了检验，并构造出了一个基于序列识别的自动化交易策略，希望可以为诸多投资、套保者提供一个新的思路。

动态时间扭曲（Dynamic Time Warping, DTW）是一种用于判断序列之间相似性的算法。DTW 背后的逻辑在于，给定两条时间序列，通过将它们进行拉伸或压缩以使它们尽可能的更加相似（或“对齐”），在此之后再计算它们的相似度或距离。“对齐”后的序列间的距离则由它们对应的个体点之间的累计距离加和给出。DTW 算法在 1970 年被发明，起初主要用于语音识别领域。如在 70 年代初期，Velichko and Zagoruyko (1970) 和 Sakoe and Chiba (1971) 均将这一算法应用到了语音文字识别领域。^{[22][26]} 在此之后，随着多种变体 DTW 算法的发明（如 Subsequence DTW, Derivative DTW 等），DTW 技术被应用到了更加广泛的领域，其中不乏金融领域的身影。例如 Huang and Kinsner (2002)、Syeda-Mahmood et al. (2007) 和 Tuzcu and Nas (2005) 等将此技术应用到了心电图序列的分析。^{[14][24][25]} 再比如 Aach and Church (2001)、Hermans and Tsiorkova (2007) 等将此技术应用到了基因序列匹配当中。^{[9][12][13]} DTW 在金融领域中被用来挖掘价格序列趋势以及验证技术指标和图像的有效性等。

技术分析主要是一种在金融市场中利用图表分析来对交易数据进行分析的方法，它基于历史通常会重现的原则，企图在价格和成交量等行情数据的图表中挖掘其趋势性。^[19] 技术分析的有效性一直是学术界的研究主题之一。国内外学者对技术分析有效性的研究大体可分为两类：技术指标型和价格形态识别型。在第一类中，Pruitt (1988) 结合价格、成交量和相对强弱指标（RSI）三个因数构建了一个交易策略，并验证了这一策略可以获取超额收益，从而证明了技术分析的有效性。^[21] Brock (1992) 利用 bootstrap 技术，对移动平均规则和交易区间突破这两个著名的策略在道琼斯指数上的有效性进行了验证，结果表明技术分析能够获取超额收益且优于基于统计方法（AR、GARCH 等）的策略。^[10] Kwon (2002) 利用 NYSE 指数对移动平均线、价格动量和成交量三种技术指标进行了实证分析，且结果表明基于这些指标的交易策略相对买入并持有策略可以获得更高收益。^[16] 国内学者中，对于技术指标的检验大多集中于移动平均线及其变体当中。茆田杨 (1995) 使用了移动平均规则对中国股市的行为特征及技术分析有效性进行了检验，得出了股市运动中可能存在某种隐含规律的结论。^[5] 马向前 (2003)、李莎 (2009) 和王志刚 (2009) 采用移动平均线对价格趋势进行了预测，结果表明技术分析可提高预测准确率。^{[3][4][7]} 孙碧波 (2001) 以上证指数为样本，



对移动平均规则和交易区间突破规则进行了检验，并得到这基于这两个技术规则的策略可以获取超额收益的结论。^[6]

在另一类中，学者们主要关注对特定价格形态的识别。Lo A. W. (2000) 首次采用局部极值点的方法对头肩形、三角形、双顶型等十种价格形态进行了数学化的定义，并使用了核回归平滑估计的非参方法和滑动窗口方法实现了价格形态的自动化识别。^[17] Dong M. (2002) 提出了模糊逻辑方法对此前的模式识别算法进行了优化，解决了查询序列可能归属于多种模式定义的问题。^[11] Suh (2004)、Zapranis (2012) 和 Maneesilpk (2004) 分别提出了对技术分析中的杯柄形态、圆形态和头肩形形态的识别方法。^{[18][23][27]} Tsinaslanidis (2018) 则没有对具体的特殊形态进行识别，他提出了使用当前的价格形态作为模板并利用动态时间弯曲 (Dynamic Time Warping, DTW) 技术去寻找最佳匹配的历史交易序列。^[19] 国内学者在这方面的研究并不多，且起步较晚。陈卓思等 (2005) 参考了 Lo A. W. (2000) 和 Brock 等 (1992) 的研究框架，通过将技术分析所使用的图形进行分解，设计出了一种衡量图形技术分析信息含量的检验途径，并提出了一个基于 bootstrap 原理的可用于检验各种交易策略的统计检验方法。^[2] 陈静 (2010) 对技术分析中的三角形态进行了分解，并通过识别分解后的序列进行形态判断。^[1] 张华 (2011) 则在 Lo A. W. (2000) 的基础上增加了“突破”的概念，并使用汇率交易数据实际验证了技术分析在外汇市场上的有效性。^[8]

总的来说，目前学者在这一方面的现有研究大多集中在具体的技术指标和图形的挖掘与识别。然而，正如 Tsinaslanidis (2018) 指出的，现有的技术指标与形态模板的数量众多，我们也没有准确的理由去证明究竟其中哪一些是最好的。^[19] 因此，与其使用特定的模板，不如将当前的形态作为参考模板（或搜索序列）并在历史成交记录中进行搜索匹配。本文在 Tsinaslanidis (2018) 的研究框架的基础上构建了一个基于价格和成交量序列匹配的自动化交易策略--SDTW 策略，这是本文的主要创新之处。通过利用并将其与构建的单向市场策略在上涨、下跌和震荡三种不同市场行情下进行了对比，得出了 SDTW 策略（或技术分析）的有效性与合约种类挂钩且相对而言更能识别市场中下跌行情的结论。的结论。本文的不足在于 1) 样本的选取过于单一，仅覆盖了股指期货市场的三个产品的研究；2) 研究范围仅限于静态历史样本数据，未能给出样本动态更新时的策略绩效评估等；3) 本文未能给出策略参数精确的最优化取值。

本文的行文框架为：第二章首先介绍技术分析的概念与基本原则，并提出本文所使用到的技术分析的范畴与内容。第三章将介绍 DTW 算法及其拓展 Subsequence DTW 算法的原理及简单应用；第四章将是本文的实证分析部分。在这一章本文将首先提出基于序列识别的 SDTW 策略的构建算法并给出与其进行对照的市场策略的算法，接着本文将使用沪深 300 股指期货、中证 500 股指期货和上证 50 股指期货的实际和模拟的价格和成交量数据对所构建策略进行回测。在这一基础上，本文将进一步尝试策略中所涵盖参数的不同取值，并对策略的回测收益做敏感性分析。此外，本文还将利用敏感性分析中获得的“最优”参数再次对策略收益率进行检测，以验证技术分析在期货市场的有效性。第五章则是全文的总结部分，本文将在这一章归纳所得发现与结论。

技术分析源于对金融市场大量历史数据的分析与趋势预测，沿用 John J. Murphy 的描述：
“*Technical Analysis is the study of market action, primarily through the use of charts, for the purpose of forecasting future price trends. The term “market action” includes the three principal sources of information available to the technician—price, volume and open interest. Chart reading is largely an art*”。
[7]

技术分析的一些要素最早出现在阿姆斯特丹 17 世纪商人 Joseph de la Vega 的 *Confusión de Confusiones* 一书中关于荷兰金融市场的报道中。在亚洲，技术分析一般被认为起源于 Homma Munehisa 在 18 世纪初开发的蜡烛图分析方法，这种方法至今仍非常流行，《日本蜡烛图技术》一书被视作技术分析几大经典之一。在 19 世纪后期的美国，Charles H. Dow 通过自己对市场行为的理解与分析开发出了一系列用于金融市场分析、交易的原则，这些原则此后被后人不断验证与补充，最终汇总形成了广为流传的道氏理论（*Dow's Theory*）。早期的技术分析主要专注于对价格、成交量图表的分析，这也是为什么技术分析通常也被称作 *charting*。现代的技术分析纳入了更多数学和计算机技术的辅助分析，除了一些常用的技术指标 MA、MACD、KDJ、ADL 等，更加复杂的数学工具被得到应用并极大的丰富了技术分析的范畴，这些工具包括时间序列分析、神经网络模型和数据挖掘技术等。现代的技术分析被广泛应用于交易策略的研发与交易系统的构建。^①

技术分析主要对历史价格（成交量）形态进行分析，通过挖掘其前进涨跌的趋势以做出交易决策。在这一过程中，分析和判断的逻辑主要基于三个基本原则：

“*Market Action Discounts Everything*（市场行为容纳一切信息）”、“*Prices Move in Trends*（价格以趋势方式前行）”和“*History tends to be repeated*（历史通常会重现）”。^[7]

2.1 技术分析基本原则

2.1.1 Market Action Discounts Everything

市场行为容纳一切信息。技术分析相信，所有可能影响产品价格的背后的信息：基本面，政治面，市场心理及其它因素都已经反应在了产品的价格中。如 2018 年 3 月 13 日白俄罗斯钾肥公司 *Belaruskali* 于 2018 年 3 月 9 日晚的坍塌事故对于全球钾肥市场供给的冲击已很迅速的反映到了诸多相关公司和合约的价格形态中，它伴随着 *钾肥指数*、*盐湖股份* 等合约和公司的三连阳。因此，研究产品或合约的价格（成交量）形态就已足够，因为它们已经包含了背后的许多信息。下图 2.1^②显示了 2018 年 3 月 9 日 *Belaruskali* 公司坍塌事故的冲击迅速被市场价格消化的图像：

^① 资料来源：https://en.wikipedia.org/wiki/Technical_analysis。

^② 资料来源：Wind 资讯。



图 2.1 钾肥指数（申万）2018 年 03 月 09 日 Belaruskali 公司坍塌事故的冲击 1

2.2.2 Prices Move in Trends:

价格以趋势方式前行。技术分析认为，价格总是以一小段一小段的趋势而前行，相同小趋势的叠加形成了合约价格的长期大趋势，而短期内相反趋势的叠加则形成了价格的波动。因此，技术分析通常要做的是通过图表分析去挖掘价格形态中的趋势与反转点，从而在适当的时点确定交易机会。下图 2.2^③给出了上证综指的价格序列，我们可以看出，价格总体呈现大的趋势，且在一个大的趋势内，包含许多方向各异的小趋势。

^③ 资料来源：Wind 资讯。



图 2.2 上证综指走势 1

2.2.3 History Tends to be Repeated:

历史通常会重现。技术分析相信历史的趋势在今天会重现，这成为了研究历史价格信息的价值所在。因此，如果这一准则是有效的，那么技术分析通过对一个合约历史的成交信息进行挖掘，可以寻找出与当前价格形态类似的历史区间，从而通过判断历史的后继走势来预测未来的涨跌趋势。这一准则似乎与有效市场理论(*Efficient Market Hypothesis, EMH*)所提出的结论相违背，依据 *EMH*，如果市场至少是弱势有效的，那么单纯的基于历史信息是无法对未来做出有效预测的，这意味着技术分析必定是失灵的。反之，如果我们验证了技术分析确实可以稳定获取超额收益，则可以反推出市场并不是有效的。

2.2 本文所用技术分析

技术分析中涉及到了很多价格的形态与技术指标，比如头肩底型、双顶双底型、三角旗型、MA、MACD、KDJ、ADL 等等。^[7]

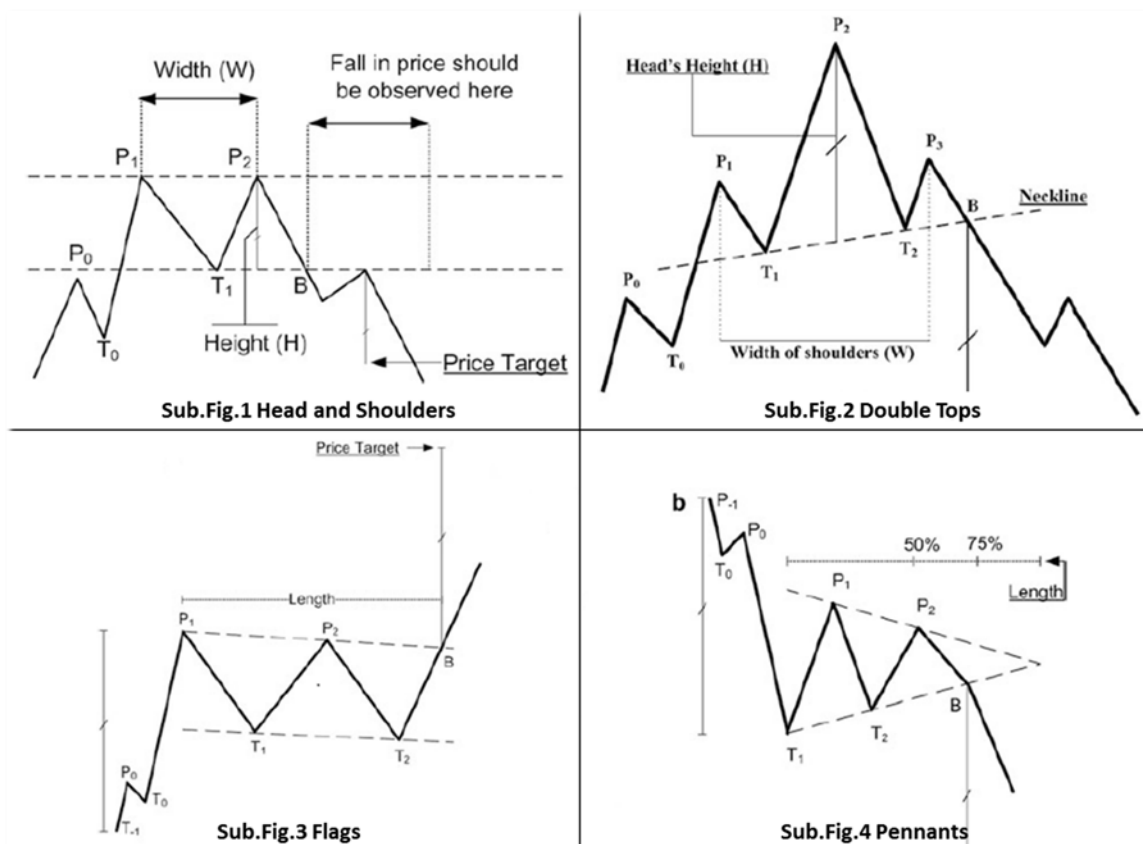


图 2.3 几种常见的技术图像：头肩底型、双顶型、旗型、三角旗型 1

这些图形与指标对于做出判断起到了很好的辅助。然而，流通于市场上的技术指标非常多，我不太确定选择哪几个能够很好地代表技术分析的主要思想，而将所有这些指标收集起来也似乎是一个不太可能实现的工程。但是我觉得这些技术图像和指标本质上是对价格、成交量形态的数学变形。因此，如果能够在历史交易数据中直接对价格、成交量的原始形态进行匹配（而不是特殊的、变形后的数据），同样也能够达到获取交易信号的目的。以著名的头肩底型为例，当价格出现这种形态时，技术分析师会预测价格将会上涨，这其中的逻辑在于在大量历史交易中，价格出现这种形态时通常会上涨。因此，如果有一段价格形态，它并不属于技术分析师常用的标准形态中的一种，但是在历史区间中每当出现相类似的形态时价格总会上涨，那么现在出现了这种形态同样能够给我们带来买入的信号。因此，在技术分析中，我觉得价格的“形状”并不是关键的，关键是出现“形状”后价格的走势。再比如以成交量形态的分析为例，近期成交量的异常突破通常确认了一段价格趋势的形成，在这一过程中成交量图像的形态会突变，以成交量为基础的各种技术指标的值也会发生突变，两者都反映了买盘信号的来临，它们包含着相同的信息。

所以，我在这篇文章中并没有使用到具体的形态或指标，我使用了原始的价格、成交量信息代替了它们。下面是本文在使用技术分析时的主要逻辑：

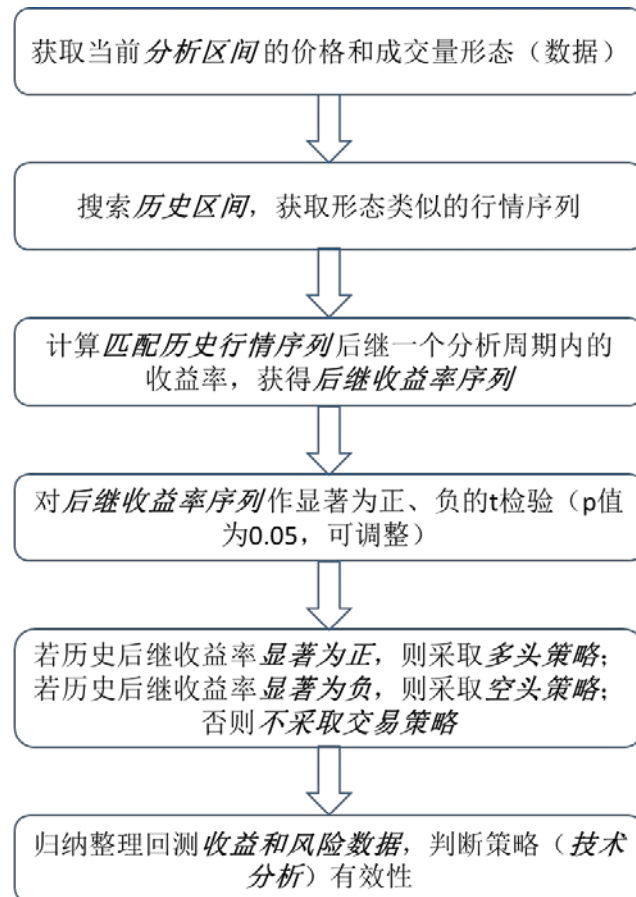


图 2.4 使用技术分析的逻辑 1

3 Dynamic Time Warping (DTW) 算法及其延伸

承前所述，技术分析主要是对产品当前的价格和成交量形态进行分析，通过去寻找历史交易数据中出现的类似行情及其后继走势以对当前行情的未来趋势做出判断，从而获取交易信号。价格和成交量数据均属于金融时间序列的范畴，因此对当前与历史行情数据进行匹配将会涉及到两段时间序列的比较。如果两段时间序列的长度是一致的，那么比较它们的相似度便可通过直接计算各点之间的累积距离（如欧氏距离）即可。但是通常，我们所获取的当前和历史行情数据的长度并不是一致的，趋势与波段有大有小，它们所横跨的时间段并不会特别规整地一一对齐。因此，在这里便涉及到了两段相似但长度通常不同的时间序列的比较。*Dynamic Time Wapring (DTW)* 算法为我们提供了一种很好的解决方案。

3.1 DTW 算法原理

对于两段长度不同但包含信息相似的时间序列^④，*DTW* 主要通过将其中的一段时间序列进行压缩或拉伸，与另一段时间序列进行匹配，最后计算匹配后的序列之间的累积距离以作为它们的相似度。一个简单的例子如下：

给定两段序列 $P = \{p_1, p_2, p_3, \dots, p_N\}$, $Q = \{q_1, q_2, q_3, \dots, q_M\}$, 其中 $N = 100$, $M = 50$ 。 P 与 Q 均为两段 $[0, 2\pi]$ 内的正弦序列叠加一个高斯白噪声过程 $\{\varepsilon_i\}_{i=1:NorM}$ ，即：

$$p_n = \sin(x_n) + 0.2\varepsilon_n, x_n \in [0, 2\pi], \varepsilon_n \sim N(0, 1) \text{ and } N = 100$$

$$q_m = \sin(x_m) + 0.2\varepsilon_m, x_m \in [0, 2\pi], \varepsilon_m \sim N(0, 1) \text{ and } M = 50$$

两条序列的图像如下图所示：

^④ 一个很好的例子是：Tom 和 Jerry 两个人同时说 Egg 这个单词，两人表达的意思是一样的，都想要吃鸡蛋。但 Tom 来自德州，说话慢，而 Jerry 来自纽约，说话急促。可能 Tom 仍处在发出 Egg 的元音阶段时，Jerry 已经说完了。因此两人最终发音得到的音序列长度并不是一致的，但两段序列包含的信息却是相同的。

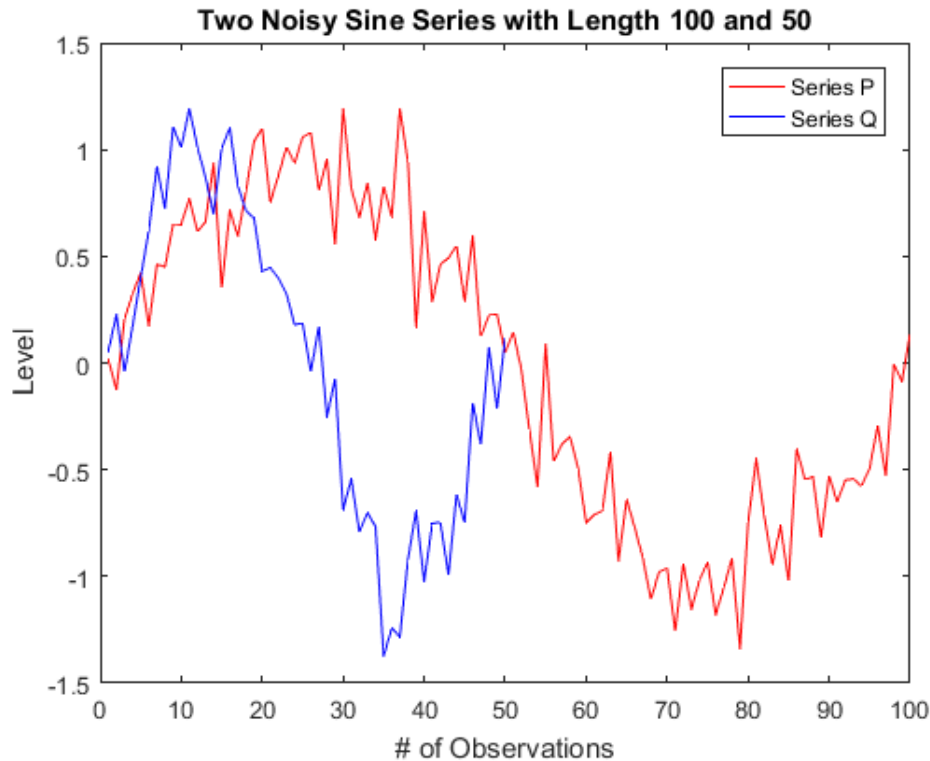
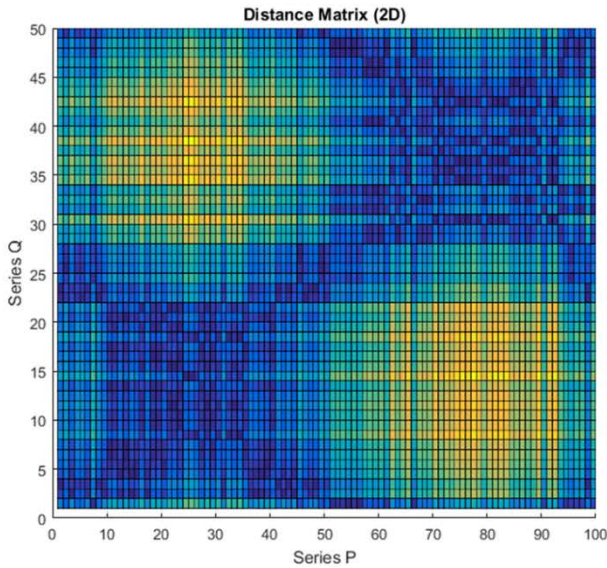


图 3.1 不等长的序列 P 与 Q 1

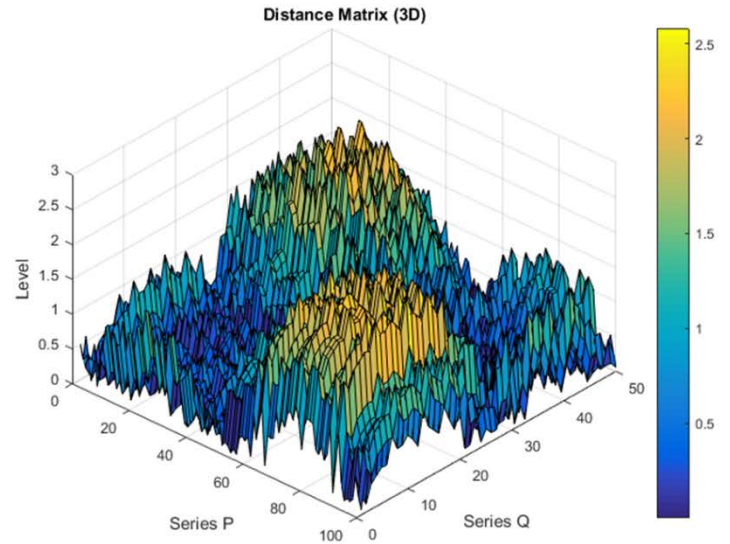
现在要计算 P 和 Q 之间点距离（相似度），由于 P 和 Q 长度不同，无法直接计算它们各点之间的累积距离。比如 P 中的第 5~10 个点包含的信息可能对应着 Q 中的第 3 个点，直接计算 P 中第 5~10 个点和 Q 中第 5~10 个点之间的距离并不是最优的。因此在计算 P 和 Q 之间的距离之前，需要先将它们“对齐”，也即含有相似信息的点放在一起计算距离。而这一“对齐”过程实际上是一种动态规划求得两序列间累积距离最小的过程，也即我们有着多种计算距离的方法（或不同的观测点对应方法），而当其中一种方法获得的最终累积距离最短时，这种方法便最有可能实现了将包含相似信息的点归类在一起。通过这种方法计算出来的距离也就是我们最终获得的衡量两序列之间的相似度的距离。

具体来看，这一动态规划求得最短距离的过程一共包含 3 个子步骤：

1) 计算两序列间的距离矩阵 $D_{(n \times m)}$: $d(n, m) = f(p(n), q(m)), f(\cdot)$, 其中 $f(\cdot)$ 为距离函数, 即计算两个点之间距离的方法, 一般可取计算欧氏距离^⑤。距离矩阵 D 给出了两个序列中各个点之间的欧氏距离, 最终得到的矩阵形式如下:



Sub.Fig. 1 距离矩阵 (2D)



Sub.Fig. 2 距离矩阵 (3D)

图 3.2 序列 P 与 Q 之间的距离矩阵 1

距离矩阵给出了后面计算累积距离矩阵的基础。

2) 定义累积距离矩阵: $\tilde{D}_{(n \times m)}$, 其中 \tilde{D} 中元素由以下状态转移方程及初值条件给出:

状态转移方程:

$$\tilde{d}(n, m) = \begin{cases} d(1, m) + \tilde{d}(1, m - 1), & n = 1 \text{ and } m \neq 1 \\ d(n, 1) + \tilde{d}(n - 1, 1), & m = 1 \text{ and } n \neq 1 \\ d(n, m) + \min\{\tilde{d}(n - 1, m), \tilde{d}(n, m - 1), \tilde{d}(n - 1, m - 1)\}, & n \neq 1 \text{ and } m \neq 1 \end{cases}$$

初值条件:

$$\tilde{d}(1, 1) = d(1, 1).$$

通过上述条件, 便可正向向前计算出各点的累积距离, 最终抵达终点 (N, M) 。

^⑤ 两个 K 维序列之间元素的欧氏距离 $d(n, m) = \sqrt{\sum_{k=1}^K ((p(n, k) - q(k, m)) * (p(n, k) - q(k, m)))}$, 当 $K = 1$ 时, 距离表达式简化为 $d(n, m) = |p_n - q_m|$ 。

3) 反向搜索获取最优路径：在得到累积距离矩阵后，便可通过反向搜索获取各个阶段取得最小累积距离时的点的坐标，也即获得 P 与 Q 之间点的对应规则。定义对应规则序列 $Z = \{z_1, z_2, z_3, \dots, z_K\}$ ，状态转移方程由以下给出：

$$z_{k-1}^* = \begin{cases} (1, m-1), & \text{if } n = 1 \\ (n-1, 1), & \text{if } m = 1 \\ \operatorname{argmin}\{\tilde{d}(n-1, m-1), \tilde{d}(n-1, m), \tilde{d}(n, m-1)\}, & \text{otherwise.} \end{cases}$$

其中，初值条件为 $z_k = (N, M)$ ，上述计算过程在 $N = M = 1$ 时终止。此时我们便得到了 P 与 Q 之间的最优对应规则（*Optimal Alignment*），且 P 与 Q 之间的距离已在第二步中给出。

以上述 P 、 Q 序列为例，通过 DTW 算法计算所最优路径和“对齐”后的序列分别如下：

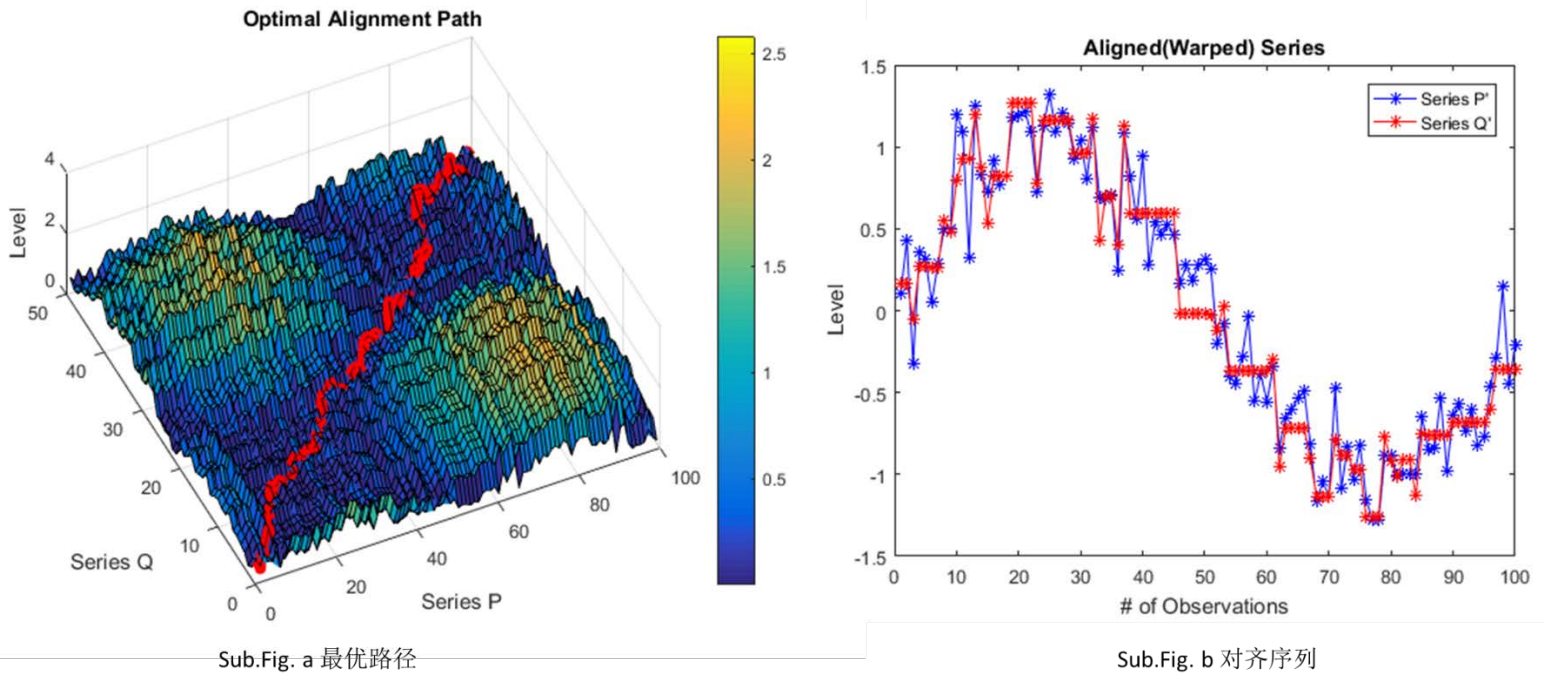


图 3.3 输出结果 1

图 3.3.a 中的红线给出了最优对齐路径，依据这一对齐规则，我们将序列 P 、 Q 分别进行了“拉伸”，即 P 中的某几个点对应着 Q 中的一个点（或反之）。拉伸后的序列如图 3.3.b 所示。最后，依据累积距离矩阵，我们还得到了 P 与 Q 之间的最短累积欧式距离为 13.8580。

3.2 Subsequence DTW (SDTW) 算法

许多时候，我们想要实现的是，给定了一小段序列 Q ，我们想要在一大段历史序列 P 中去寻找有没有类似的情况出现过。这个时候查询序列 Q 与参考序列 P 的长度相差甚远，即 $M \ll N$ 。为了



实现查询的目的，我们需要在参考序列 P 中滚动搜索相似长度的子序列（*Subsequence*）来与查询序列 Q 进行匹配，并从中选出相似度最高（累积距离最短）的序列作为我们的结果。

在这种情况下，上述 DTW 算法调整为：

表 3.1 SDTW 算法 1

SDTW 算法: $SDTW(P, Q)$

输入: P -参考序列, Q -查询序列

输出: A^* 最优路径矩阵

1. 计算查询序列与参考序列的距离矩阵 $D_{(n \times m)}$: $d(n, m) = f(q(n), y(m))$;
2. 计算累计距离矩阵 $\tilde{D}_{(n \times m)}$ ，其中 \tilde{D} 中元素由以下状态转移方程和边值条件给出：

状态转移方程：

$$\begin{cases} \tilde{d}(n, 1) = d(n, 1), m = 1 \text{ and } n \neq 1 \\ \tilde{d}(1, m) = d(1, m) + \tilde{d}(1, m - 1), n = 1 \text{ and } m \neq 1 \\ \tilde{d}(n, m) = d(n, m) + \min\{\tilde{d}(n - 1, m), \tilde{d}(n, m - 1), \tilde{d}(n - 1, m - 1)\}, n \neq 1 \text{ and } m \neq 1 \end{cases}$$

3. 定义距离临界值 τ ，获取 $\tilde{d}(n, M) < \tau$ 的 n 坐标值向量 $\Omega^* = \{\omega_l^*\}_{l=1}^L$;
4. 对于 Ω^* 内的每一元素，反向搜索获取最优路径矩阵 $A^* = \{\alpha_l^*\}_{l=1}^L$ ，其中对于每一个 $\alpha_l^* = \{z_1, z_2, z_3, \dots, z_K\}$ ，且获取 α_l^* 中元素的状态转移方程由以下给出：

$$z_{k-1}^* = \begin{cases} (1, m - 1), \text{ if } n = 1 \text{ and } m \neq 1 \\ \text{argmin}\{\tilde{d}(n - 1, m - 1), \tilde{d}(n - 1, m), \tilde{d}(n, m - 1)\}, \text{ if } n \neq 1 \text{ and } m \neq 1 \\ \text{Stop}, \text{ if } m = 1 \end{cases}$$

返还值: $A^* = \{\alpha_l^*\}_{l=1}^L$.

注意在上述调整后的算法中，出现了以下几个变化：

(1) 计算累计距离矩阵 \tilde{D} 时，初边值条件发生了变化，在此时我们没有再明确给出初值条件 $\tilde{d}(1,1)$ 的值 s ，且边值条件 $\tilde{d}(n, 1) = d(n, 1) + \tilde{d}(n - 1, 1)$ 变为了 $\tilde{d}(n, 1) = d(n, 1)$ 。这两个变化实际上是保证了，匹配序列的起点不一定是从 $(1,1)$ 点开始，而是从任何 $\alpha: 1 \leq \alpha \leq M$ 开始；

(2) *Step 4*) 中坐标状态转移方程中不再包含关于 $m = 1$ 时的状态转移，因为此时得到的匹配子序列不一定是从 $n = 1$ 处出发，所以只要保证查询序列的完整性（即 $m = 1$ 时结束搜索）即可。

通过上述调整算法，我们最终得到了最优路径矩阵 $A^* = \{\alpha_l^*\}_{l=1}^L$ 。 A^* 中的最后一行包含了 L 条最优路径的起始坐标值（即最优匹配序列的起点）， A^* 中的第一行包含了 L 条最优路径的终端坐标值（即最优匹配序列的终点）。这样我们就得到了由距离临界值 τ 所限定的 L 条最优匹配子序列。

3.3 SDTW 算法与技术分析

承前所述，*SDTW*算法能够实现在历史参考序列中去搜索与查询序列相似的子序列。在这里，如果我们定义查询序列为特殊的技术图像（如头肩底型），或者为当前分析的一段时间内的价格和成交量序列，历史参考序列为历史交易行情数据，那么我们便能实现给定分析序列，根据历史经验来确定它的未来走势的功能。价格形态的匹配找出了可能的多空机会，而成交量（放量）的进一步匹配则确认了机会的出现。

4 实证分析

4.1 数据描述

因作者缺乏能力覆盖中国期货市场所有产品合约的研究，这篇文章只选取了市场中几个具有代表性的股指期货用作技术分析有效性检验。具体地，所选数据为：（1）沪深 300 指数期货成立日（2010 年 4 月 16 日）至 2018 年 3 月 16 日的收盘价和日成交量序列；（2）中证 500 股指期货成立（2015 年 4 月 16 日）至 2018 年 3 月 16 日的收盘价和日成交量序列；（3）上证 50 股指期货成立（2015 年 4 月 16 日）至 2018 年 3 月 16 日的收盘价和日成交量序列。三者的数据来源均为 Wind 金融数据库。

4.2 参数与假设

（1）沪深 300 股指期货初始保证金率 15%，中证 500 股指期货保证金率 30%，上证 50 股指期货保证金率 15%；维持保证金比例 0.75；

（2）初始资金 1000 万，每次交易合约数量为 10 手，交易手续费率为 0.023%，不考虑其它成本（如冲击成本，保证金的时间价值等）；

（3）交易无延迟瞬间达成，且无其它技术性故障。

4.3 交易策略

整个交易策略的主要流程为：利用 *SDTW* 技术对当前价格和成交量序列在历史样本中进行搜索与匹配，将获得的最佳匹配序列的历史后继走势作为当前序列的预测走势，若历史后继走势收益率序列显著为正（ p 值 < 0.05 ）则多头建仓，若显著为负（ p 值 < 0.05 ）则空头建仓，否则不建仓。

交易策略的主要逻辑源于技术分析原则三：历史通常会重现。通常，在合约价格不出现异动（如被主力操控）的情况下，价格形态预示着趋势的来临，而成交量的突增则确认了该趋势的形成。所以，如果当前行情的价格和成交量形态均与历史的某段行情非常相似，那么根据技术分析的原则我们会相信它们的未来走势将会是类似的。所以便可以通过历史的后继行情来预测当前未来的走势，从而采取对应交易策略。

交易策略的具体细则由以下算法给出：



表 4.1 SDTW 策略 1

SDTW 策略: $SDTW(data, refLength, RW, pre_RW, NPQ, NVQ, pVa, rate_RiskFree)$

输入: $data$ -合约交易数据, $refLength$ -历史区间长度, RW -滚动搜索区间长度, pre_RW -预测区间长度, NPQ -最大价格序列查询匹配数, NVQ -最大成交量序列查询匹配数, $pVal$ -收益率建仓显著度 p 值, $rate_RiskFree$ -无风险利率

输出: $PFValue$ -仓位净值序列

1. 将整个回测区间依据 $refLength$ 划分为历史区间 HI 和查询区间 QI ;

for $bkt = 1:(length(data)-RW+1)$:

2. 在 QI 中从起始点开始滚动获取长度为 RW 的查询序列 QS ;

3. 价格查询: 在 HI 中搜索价格序列与 QS 相似的子序列, 获取相似度最高 (距离最短) 的前 NPQ 个子序列;

4. 成交量查询: 在上述获取的 NPQ 个子序列中, 再此搜索成交量序列与 QS 相似的子序列, 获取相似度最高 (距离最短) 的前 NVQ 个子序列;

5. 对于获取的 NVQ 个最佳匹配子序列, 在历史区间中分别计算它们后继 pre_RW 天的收益率, 剔除异常值 (子序列位于历史区间的末尾), 获得后继收益率序列 ref_Rets ;

6. 对获取的 ref_Rets 的均值分别作显著为正、负的 t 检验 (显著度 p 值为 $pVal$):

if ref_Rets 均值显著为正:

7. 平掉所有空头仓位;

8. 多头建仓 10 手。

else if ref_Rets 均值显著为负:

9. 平掉所有多头仓位;

10. 空头建仓 10 手。

else:

11. 不平仓、不建仓。

end

13. for $i = bkt:(bkt+RW+pre_RW-1)$:

14. 记录保证金账户情况:

if 遭遇 *Margin Call*:

if 爆仓:

15. break;

16. print(“爆仓, 结束。”);

```

else:
    17. 补充保证金;
    18. 记录仓位净值  $PFValue(bkt)$ ;
end
else:
    19. 记录仓位净值  $PFValue(bkt)$ ;
end
end
20.  $bkt = bkt + pre\_RW$ 
end
21. 平掉所有多、空头仓位，并记录仓位净值  $PFValue(length(PFValue))$ 。

```

返回値: $PFValue$.

对于与上述策略作为对照实验的“市场策略”，这篇文章采取了如下交易策略（多/空头并持有至到期）：

表 4.2 市场多头策略 1

市场多头策略: $longMKT(QI)$

输入: QI -查询区间序列

输出: $PFValue$ -仓位净值序列

```

1. 多头开仓 10 手;
for  $bkt = 1:length(QI)$ :
    2. 记录保证金账户情况:
        if 遭遇 Margin Call:
            if 爆仓:
                3. break;
                4. print (“爆仓，结束。”);
            else:
                5. 补充保证金;
                6. 记录仓位净值  $PFValue(bkt)$ ;
            end
        end
    end
end

```

else:

7. 记录仓位净值 $PFValue(bkt)$;

end

8. $bkt = bkt + 1$;

end

9. 平掉所有仓位, 并记录仓位净值 $PFValue(length(PFValue))$ 。

返回值: $PFValue$.

表 4.3 市场空头策略 1

市场空头策略: $shortMKT(QI)$

输入: QI -查询区间序列

输出: $PFValue$ -仓位净值序列

1. 空头开仓 10 手;

for $bkt = 1:length(QI)$:

2. 记录保证金账户情况:

if 遭遇 *Margin Call*:

if 爆仓:

3. *break*;

4. *print*("爆仓, 结束。");

else:

5. 补充保证金;

6. 记录仓位净值 $PFValue(bkt)$;

end

else:

7. 记录仓位净值 $PFValue(bkt)$;

end

8. $bkt = bkt + 1$;

end

9. 平掉所有仓位, 并记录仓位净值 $PFValue(length(PFValue))$ 。

返回值: $PFValue$.

4.4 结果分析

我们首先来看当上述参数取如下一组值时，两组不同策略的表现（其中市场策略取多、空中表现更好的一只）：

表 4.4 回测参数 1

| 回测参数 | <i>refLength</i> | <i>RW</i> | <i>pre_RW</i> | <i>NPQ</i> | <i>NVQ</i> | <i>pVal</i> | <i>rate_RiskFree</i> [®] |
|------|--------------------|-----------|---------------|------------|------------|-------------|-----------------------------------|
| | (2/3)*length(data) | 30 | 7 | 100 | 30 | 0.05 | 3.46% |

表 4.5 交易参数 1

| 交易参数 | 初始资金 | 初始保证金比率 | 维持保证金比例 | 单次建仓手数 |
|------|------|--|---------|--------|
| | 1 亿元 | 沪深 300 股指期货：15% 中证 500 股指期货：30% 上证 50 股指期货：15% | 0.75 | 10 |

用于比较的市场策略取多、空策略中的收益率较大者。两组策略在 3 个不同品种的回测收益、风险数据的比较如下表 4.6 所示：

表 4.6 回测收益统计表 1

| 合约：沪深 300 股指期货 | 策略 | |
|----------------|-------------|------------|
| | <i>SDTW</i> | <i>MKT</i> |
| 年化收益: | -4.44% | 0.97% |
| 年化波动: | 6.36% | 2.16% |
| 夏普比率: | -1.3363 | -1.1565 |
| 最大回撤: | 14.04% | 3.04% |
| 合约：中证 500 股指期货 | 策略 | |
| | <i>SDTW</i> | <i>MKT</i> |
| 年化收益: | 10.67% | 1.24% |
| 年化波动: | 7.02% | 3.38% |
| 夏普比率: | 0.9399 | -0.6553 |
| 最大回撤: | 3.37% | 3.56% |
| 合约：上证 50 股指期货 | 策略 | |
| | <i>SDTW</i> | <i>MKT</i> |
| 年化收益: | -5.91% | 2.00% |
| 年化波动: | 2.17% | 1.34% |
| 夏普比率: | -4.3214 | -1.0903 |
| 最大回撤: | 4.88% | 1.23% |

[®] 选取交易所 7 天回购利率从沪深 300 股指期货成立日 2010 年 4 月 16 日至 2018 年 3 月 16 日的平均值 3.4619%。

从上表可看出，在上述给定参数取值与回测区间下，*SDTW* 策略仅在中证 500 股指期货（以下简称 *ZZ500*）这一个品种中跑赢了市场策略，年化收益为 10.67%，夏普比率为 0.9399，最大回撤为 3.37%。同期市场策略收益仅为 1.24%，低于无风险收益率，夏普比率为-0.6553，最大回撤略高于 *SDTW* 策略。这表明在所考虑回测区间下，对于中证 500 股指期货这一品种，基于技术分析的交易策略能够获得较低风险的超额收益。

然而在同样的回测区间，对于其它的两个期货品种，*SDTW* 策略的表现明显要逊于市场策略。在这两个品种中，*SDTW* 策略均实现了较大亏损，年化收益率达到-5%左右，同期市场策略均录得正收益；波动率数值相对较大，为市场策略的 2 倍左右；最大回撤则为市场策略的 4 倍左右。这表明在回测区间，对于沪深 300 股指期货（以下简称 *HS300*）和上证 50 股指期货（以下简称 *SZ50*）这两个品种，基于技术分析的交易策略并不能奏效。

这一结果并不那么清晰，其一，它告诉了我们技术分析的有效性似乎与合约的品种有关。不同品种的市场参与者数量及交易量可能存在差异，它们的市场有效性亦随之有所不同。其二，这一结果可能与所考察行情数据的特殊形态相关。为了弄清楚这个问题的答案，我们首先需要来观察一下三个合约价格序列分别呈现的形态：

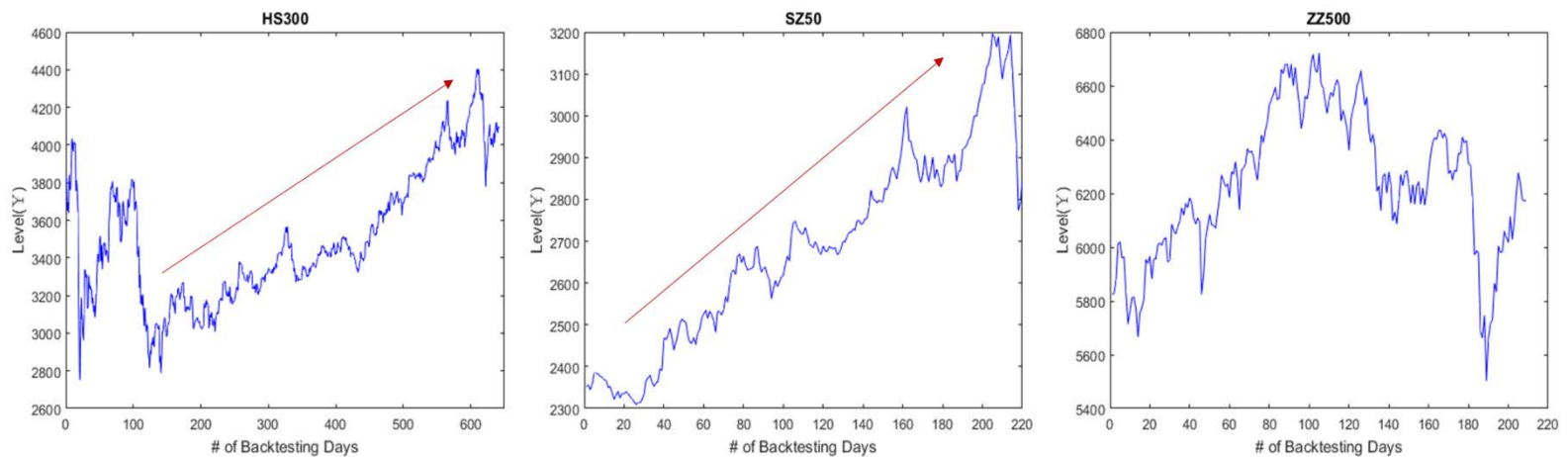


图 4.1 价格序列：从左至右分别对应 HS300、SZ50 及 ZZ500

上图8中从左至右分别给出了*HS300*、*SZ50*和*ZZ500*在回测区间内的价格序列。从中我们不难发现，在回测收益糟糕的两个品种中，它们的价格序列均呈现出明显的趋势，而在录得较大正收益的*ZZ500*品种中，它的价格序列呈现不规则的震荡态势。这一现象与Prodromos E. Tsinaslanidis（2018）得到的结论类似。Prodromos E. Tsinaslanidis利用*DTW*技术通过对大量纽约证券交易所上市股票在2010年至2015年间的行情的分析，指出技术分析能够显著地预测出市场的下跌行情，而对于上涨行情的预测却并不那么显著。^[19]结合Tsinaslanidis先生的结论，在这里我们不难发现，技术分析的有效性似乎与合约价格所处的走势形态有关。因此，在接下来的分析中，我们将有必要

把我们的分析依据市场行情的走势划分为上涨、下跌和震荡三种情况进行分类讨论。在这里，我们将原始的合约行情数据进行了一定的构造，形成了三个产品分别处于上涨、下跌和震荡三种不同情形下的模拟数据，此后的分析都将基于这些构造数据。

首先来看当市场处于上涨时 *SDTW* 策略的表现：

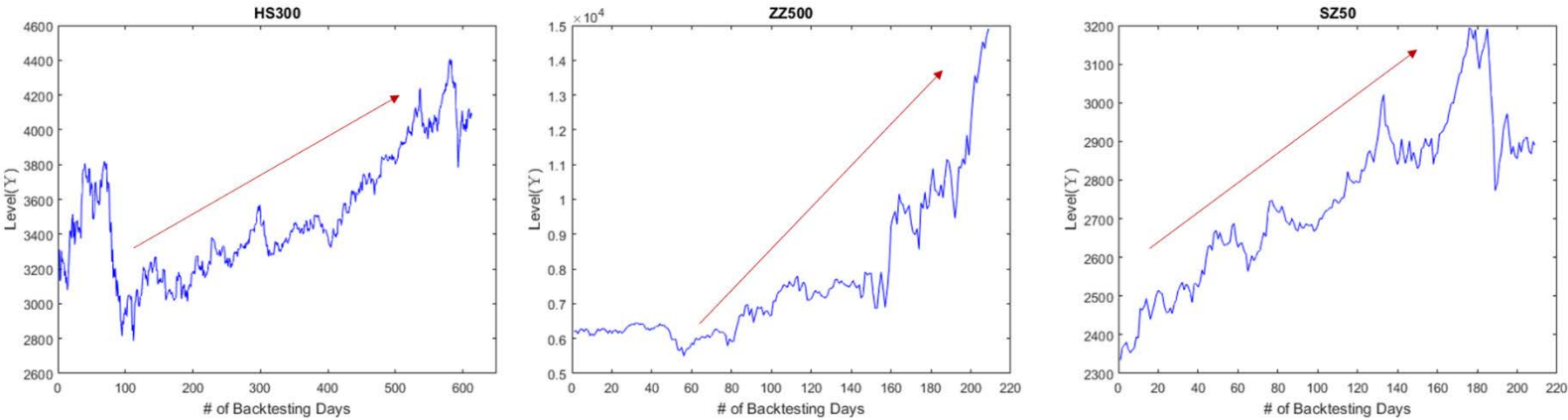


图 4.2 上涨行情模拟数据 1

上图 4.2 给出了构造出的上涨行情数据的走势图像，此时三个合约的价格形态均存在较强的向上趋势。同样的，首先来看在使用表 4.4 给出的回测参数时的回测效果。

表 4.7 回测收益率统计表：上涨行情 1

| 合约：沪深 300 股指期货 | 策略 | |
|----------------|-------------|------------|
| | <i>SDTW</i> | <i>MKT</i> |
| 年化收益: | -4.44% | 0.97% |
| 年化波动: | 6.36% | 2.16% |
| 夏普比率: | -1.3363 | -1.1565 |
| 最大回撤: | 14.04% | 3.04% |

| 合约：中证 500 股指期货 | 策略 | |
|----------------|-------------|------------|
| | <i>SDTW</i> | <i>MKT</i> |
| 年化收益: | 10.99% | 31.39% |
| 年化波动: | 45.96% | 12.89% |
| 夏普比率: | 0.1638 | 2.1671 |
| 最大回撤: | 31.04% | 4.41% |

| 合约：上证 50 股指期货 | 策略 | |
|---------------|-------------|------------|
| | <i>SDTW</i> | <i>MKT</i> |
| 年化收益: | -5.91% | 2.00% |
| 年化波动: | 2.17% | 1.34% |
| 夏普比率: | -4.3214 | -1.0903 |
| 最大回撤: | 4.88% | 1.23% |

从上表可看出，在所给参数条件下，SDTW策略没能跑赢市场策略，在每个合约品种下的收益率与夏普比率均相对较低，且最大回撤大致为市场策略的 4 倍。同理，下表表 4.8 和表 4.9 分别给出了当市场处于下跌和震荡行情时 SDTW 策略与市场策略风险收益绩效的比较。

表 4.8 回测收益率统计表：下跌行情 1

| 合约：沪深 300 股指期货 | 策略 | |
|----------------|---------|---------|
| | SDTW | MKT |
| 年化收益: | -5.16% | 0.51% |
| 年化波动: | 13.31% | 3.20% |
| 夏普比率: | -0.6474 | -0.9206 |
| 最大回撤: | 21.47% | 4.70% |

| 合约：中证 500 股指期货 | 策略 | |
|----------------|---------|--------|
| | SDTW | MKT |
| 年化收益: | -4.43% | 9.79% |
| 年化波动: | 24.25% | 14.79% |
| 夏普比率: | -0.3255 | 0.4283 |
| 最大回撤: | 31.11% | 10.48% |

| 合约：上证 50 股指期货 | 策略 | |
|---------------|---------|---------|
| | SDTW | MKT |
| 年化收益: | -3.47% | 1.58% |
| 年化波动: | 2.19% | 1.03% |
| 夏普比率: | -3.1698 | -1.8209 |
| 最大回撤: | 4.27% | 0.40% |

表 4.9 回测收益率统计表：震荡行情 1

| 合约：沪深 300 股指期货 | 策略 | |
|----------------|---------|---------|
| | SDTW | MKT |
| 年化收益: | 0.03% | 0.12% |
| 年化波动: | 2.64% | 1.62% |
| 夏普比率: | -1.2996 | -2.0684 |
| 最大回撤: | 3.02% | 2.49% |

| 合约：中证 500 股指期货 | 策略 | |
|----------------|--------|---------|
| | SDTW | MKT |
| 年化收益: | 10.67% | 1.24% |
| 年化波动: | 7.02% | 3.38% |
| 夏普比率: | 1.0258 | -0.6553 |
| 最大回撤: | 3.37% | 3.56% |

| 合约：上证 50 股指期货 | 策略 | |
|---------------|---------|---------|
| | SDTW | MKT |
| 年化收益: | 2.95% | 0.93% |
| 年化波动: | 1.92% | 0.87% |
| 夏普比率: | -0.2665 | -2.9016 |
| 最大回撤: | 1.16% | 0.54% |

从上述统计表可看出，当市场存在明显趋势（上行或下行）时，基于技术分析的 *SDTW* 策略均不能跑赢市场策略，而当市场不存在较明显的大趋势（震荡）时，*SDTW* 策略却在大多数情况能够胜过市场（其中 *ZZ500* 合约收益率远高于市场策略）。这一结果似乎与技术分析的原则二相矛盾，原则二表明技术分析相信合约价格以趋势形态前进，且技术分析要做的就是挖掘出这样的趋势以采取交易措施。这一矛盾可能与我们在 *SDTW* 策略中设定的参数取值有关。比如在表 5 给出的交易参数中，我们将 *pre_RW* 设为了 7，这表明我们根据后继 7 天的交易数据来判断未来可能的走势。7 天的数据可能不足以容纳一个完整的趋势，因此基于此的判断便很可能不准确。故而，为了得到更准确的结论，我们需要用这些参数的取值对 *SDTW* 策略做一下敏感性分析。

4.5 敏感性分析

在 *SDTW* 策略中，一共需要传入 8 个参数，它们分别是：*data*, *refLength*, *RW*, *pre_RW*, *NPQ*, *NVQ*, *pVal*, *rate_RiskFree*。其中，*data* 和 *rate_RiskFree* 为给定常量，剩下的 6 个参数均为可变变量。*refLength* 在此处只是用于决定回测的区间，并给出了历史搜索样本的长度，且本文认为这一变量对策略收益率的影响与 *NPQ*、*NVQ* 类似，都在于决定搜索匹配的样本数量。因此我们将不对此变量做敏感性分析。综上，接下来本文将依据 *RW*, *pre_RW*, *NPQ*, *NVQ* 及 *pVal* 这 5 个变量的取值对 *SDTW* 策略的绩效进行敏感性分析。

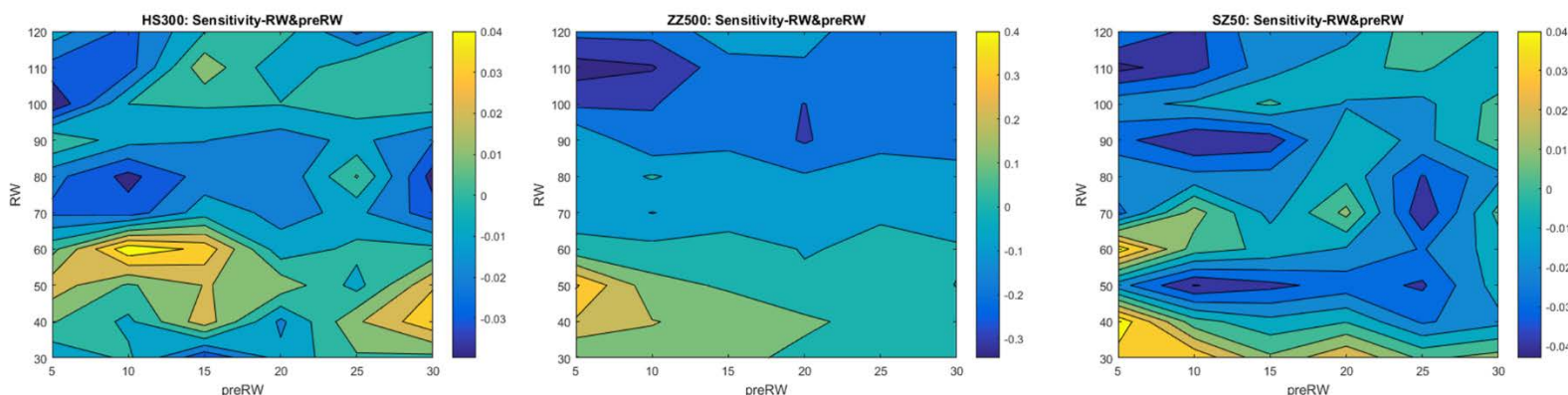


图 4.3 震荡行情敏感性分析：RW & pre_RW 1

上图 4.3 给出了在震荡行情下对参数 *RW* 和 *pre_RW* 分别在 [30, 120] 和 [5, 30] 区间内取值时得到的策略收益率与市场收益率差值的等高图。从图中可看出，策略的超额收益率对于 *RW* 和 *pre_RW* 取值的敏感性较高。以 *HS300* 合约为例，当我们使用历史后继 5~20 天的收益率作为未来的预测基准时，若选取的查询序列长度在 40~65 之间，则策略能够获得超额收益（1%~4%），而若查询长度稍微增加一点，比如在 70~90 之间，则策略的收益将会出现大幅下降（-4%~-1%）。此外，我们还可以发现，三个产品可以获得的超额收益水平有所差异，其中对于 *ZZ500* 合约平均而言策略可以获得更高收益，当 *RW* 在 40~55 区间，*pre_RW* 在 5~10 区间时，获得的超额收益高达 30%~40%。

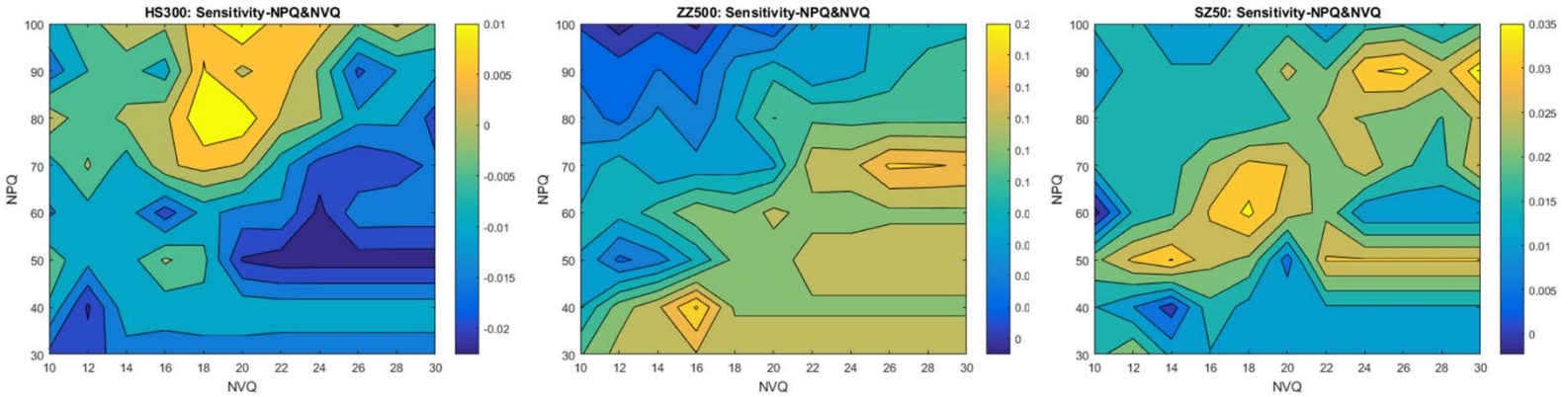


图 4.4 震荡行情敏感性分析：NPQ & NVQ 1

同样的，上图 4.4 给出了震荡行情下的 NPQ 和 NVQ 两个参数的敏感性分析超额收益等高图。从图中可以观察到：其一，策略的收益率对于这两个参数的敏感性较高；其二，并不是匹配的序列数越少越精确，得到的策略收益率越高。可能的原因是，除了价格和成交量序列，还存在影响一个合约价格未来走势的其它因素。因此，策略的收益率将会是匹配数量和其它因素力量的权衡结果。其三，不同品种得到的收益率存在差异， $ZZ500$ 的收益率仍显著高于其它两种产品。

敏感性分析的最后一个变量是匹配序列的历史后继收益率不为零的显著度 $pVal$ 。我们通过回测发现，当 $pVal$ 从 0.005 以 0.005 为梯度增加值 0.1 时，得到的三个合约的回测收益率均没有发生变化。因此，本文认为 $pVal$ 变量的取值对于 $SDTW$ 策略的收益并无较大影响。

关于上涨行情与下跌行情的敏感性分析类似，我们在这里不再一一赘述。下图图 4.5 至图 4.10 分别给出了两类行情及真实行情下的 RW & pre_RW 和 NPQ & NVQ 的敏感性分析结果。

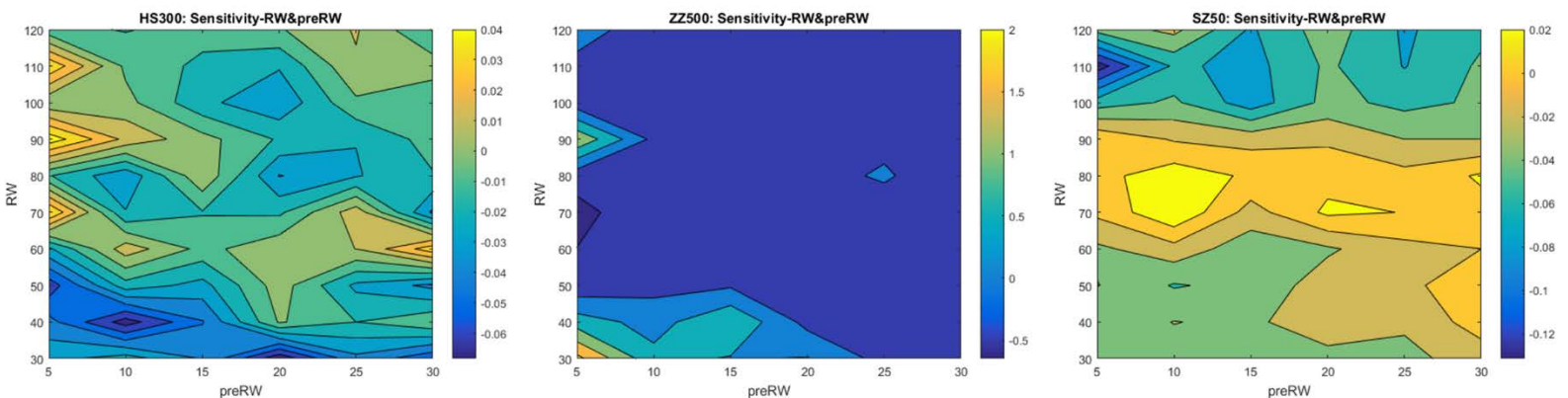


图 4.5 上涨行情敏感性分析：RW & pre_RW 1

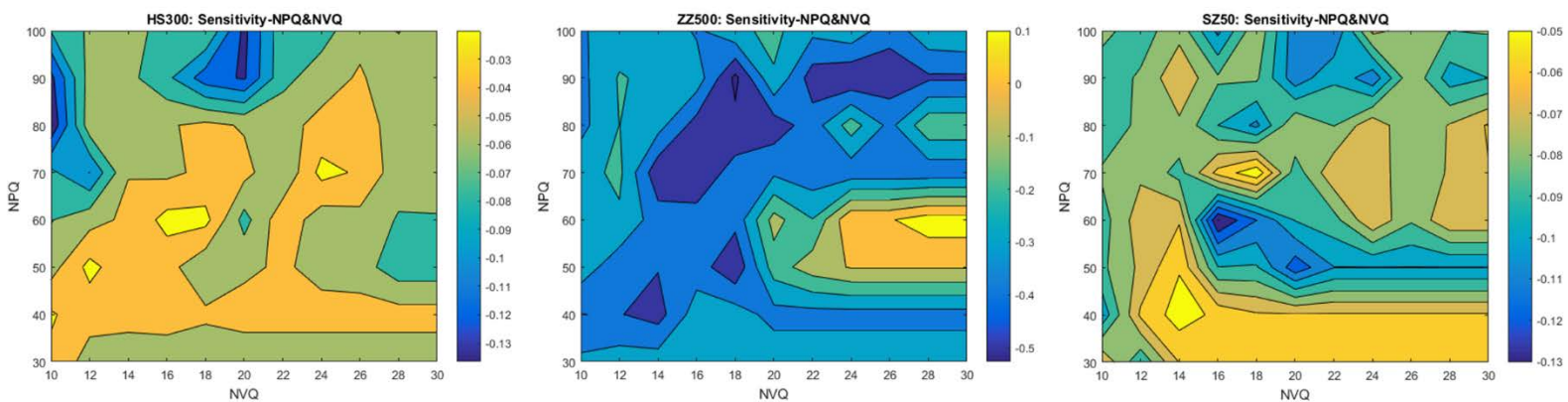


图 4.6 上涨行情敏感性分析：NPQ & NVQ 1

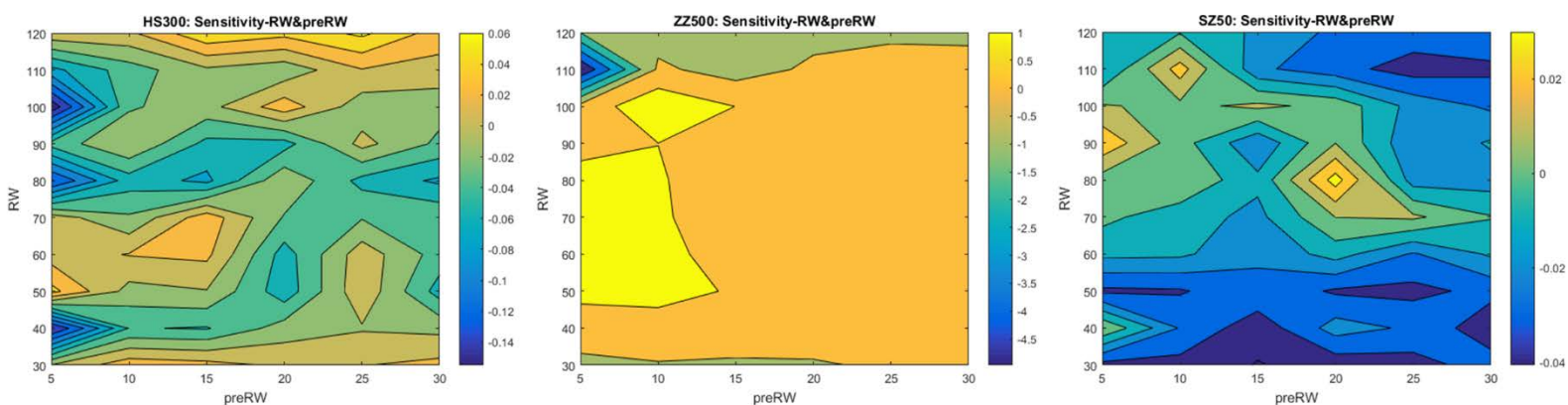


图 4.7 下跌行情敏感性分析：RW & pre_RW 1

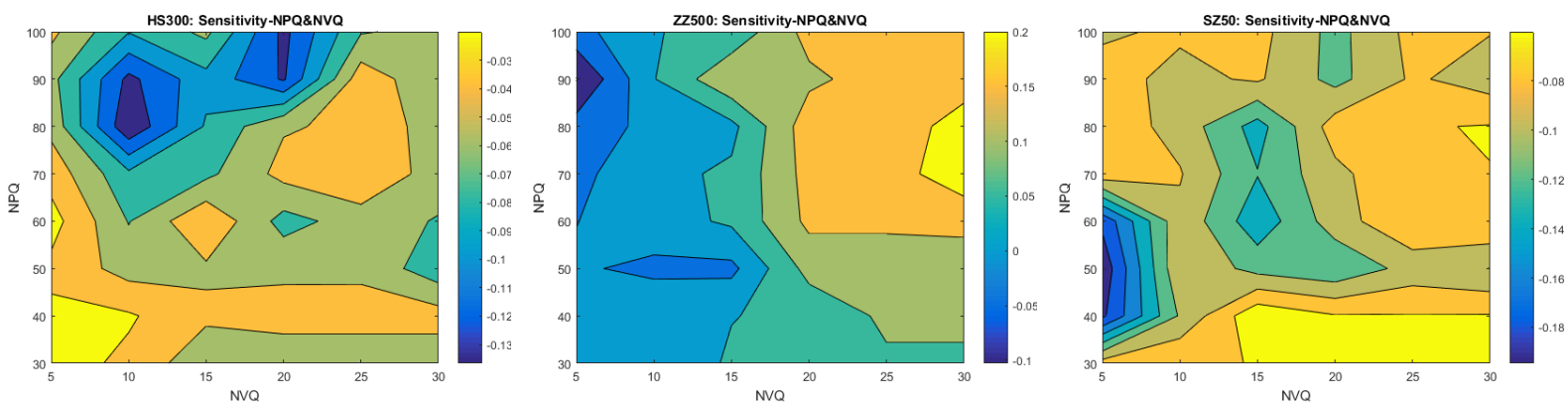


图 4.8 下跌行情敏感性分析：NPQ & NVQ 1

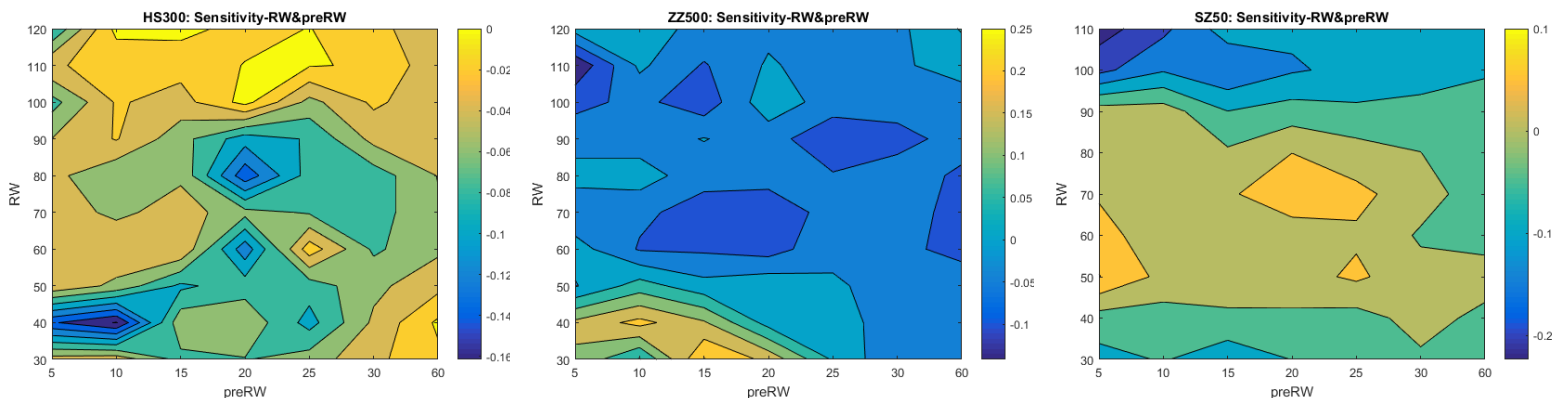


图 4.9 真实行情敏感性分析：RW & pre_RW 1

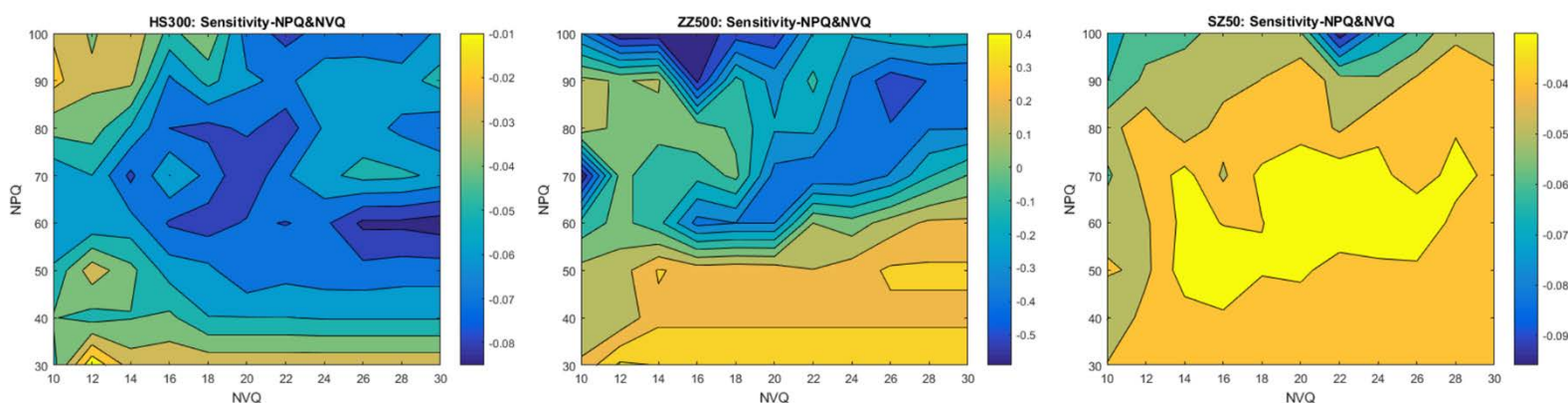


图 4.10 真实行情敏感性分析：NPQ & NVQ 1

由上述图片可看出，策略在不同行情下的表现情况存在差异，且可初步判定，策略似乎更能够识别下跌和上涨行情，因它们在敏感性分析中获得超额收益的个案比例明显更高。这一结论符合技术分析原则二，与我们在未做敏感性分析前得到的结论相反。此外，通过对真实行情的敏感性分析进行观察，策略在 *HS300* 和 *SZ50* 两个合约的表现非常糟糕，少有录得超额收益的情况出现。而相比之下 *ZZ500* 合约获得超额收益的情况出现的更为频繁，且最高时超额收益可达 25%。出现这一问题的原因我们在此前已经指出，真实行情中 *HS300* 和 *SZ50* 这两个合约的价格走势均呈现明显的上涨趋势，而 *ZZ500* 合约则呈现震荡态势。*SDTW* 策略识别下跌和震荡行情能力更强，因此它在 *ZZ500* 合约中的表现更好也就不足为奇了。

在得到上述敏感性分析的结果后，我们将使用所得最优参数再次对不同行情下策略的回测表现进行测试。所得测试结果一览表如下表 4.10 所示：

表 4.10 最优测试结果一览 1

| | <i>HS300</i> | | <i>ZZ500</i> | | <i>SZ50</i> | |
|------|--------------|------------------|--------------|------------------|-------------|------------------|
| | 平均超额收益 | 显著度(<i>p</i> 值) | 平均超额收益 | 显著度(<i>p</i> 值) | 平均超额收益 | 显著度(<i>p</i> 值) |
| 上涨行情 | -3.06% | 0.0000 | 21.51% | 0.0208 | 2.01% | 0.0000 |
| 下跌行情 | 2.87% | 0.0001 | 60.74% | 0.0000 | 1.14% | 0.0060 |
| 震荡行情 | 0.90% | 0.0085 | 20.88% | 0.0000 | 2.51% | 0.0000 |
| 真实行情 | -1.23% | 0.0000 | 22.35% | 0.0000 | 7.16% | 0.0000 |

观察上表，我们可发现如下现象：其一，在三个股指期货合约中，*SDTW*策略在 *ZZ500* 合约中不论在何种行情下均录得最大超额收益，这与我们此前获得的结论一致；其二，基于技术分析的 *SDTW* 策略总体更能识别下跌行情，其中在 *HS300* 与 *ZZ500* 两合约中录得的超额收益要显著高于 *SZ50*。这一结论与 Tsinaslnaidis(2018)利用 *NYSE* 数据所得实证结论一致。

5 结论

本文使用了数据挖掘中常用的用来比较序列之间相似度的算法动态时间扭曲 (*Dynamic Time Warping, DTW*) 对技术分析在中国期货市场的有效性进行了检验。我们在 TsinasIlnaidis(2018)的研究框架的基础上, 构建了一个基于价格和成交量序列匹配的自动化交易策略--*SDTW* 策略, 并通过使用 3 个股指期货合约 (沪深 300 股指期货合约、中证 500 股指期货合约及上证 50 股指期货合约) 的行情数据, 将其表现与构建的单向市场策略在上涨、下跌和震荡三种不同市场行情下进行了实证对比。通过分析所得结果, 我们得到了如下结论:

1) *SDTW* 策略 (或技术分析) 的有效性与合约种类挂钩, 在股指期货市场中, 技术分析尤其能在中证 500 股指期货合约交易中获取较高超额收益;

2) *SDTW* 策略 (或技术分析) 相对而言更能识别下跌行情, 在下跌行情中策略能够相对获得更高超额收益。

本文在做实证检验时使用的是静态历史数据, 后续的研究方向可以将数据的范畴扩展至动态更新的模式, 计算量将会有所增大。此外, 本文并没有给出 *SDTW* 策略参数的精确最优取值, 因此通过最优化算法 (遗传算法、神经网络等) 获取策略最优参数将是另外一个可能的后继研究方向。

6 参考文献

- [1] Aach J., Church G. M. (2001). Aligning Gene Expression Time Series with Time Warping Algorithms. *Bioinformatics*, 17(6), 495-508.
- [2] Brock W., Lakonishok J., LeBaron B. (1992). Simple Technical Trading Rules and the Stochastic Properties of Stock Returns. *The Journal of Finance*, 47(5), 1731-1764.
- [3] Dong M. (2002). Exploring the Fuzzy Nature of Technical Patterns of U.S. Stock Market. *Proceedings of Fuzzy System and Knowledge Discovery*, 1(1), 324-328.
- [4] Giorgino T., Tormene P. (2009). *dtw*: Dynamic Time Warping Algorithms. R package version 1.13-1, URL <http://CRAN.R-project.org/package=dtw>. Accessed date: 2018/03/03.
- [5] Hermans F., Tsiporkova E. (2007). Merging Microarray Cell Synchronization Experiments Through Curve Alignment. *Bioinformatics*, 23(2), 64-70. doi:10.1093/bioinformatics/btl320.
- [6] Huang B., Kinsner W. (2002). ECG Frame Classification Using Dynamic Time Warping. In W. Kinsner, A. Sebak, K. Ferens (eds.), *Proceedings of the Canadian Conference on Electrical and Computer Engineering - IEEE CCECE 2002*, Vol. 2, pp. 1105-1110. IEEE Computer Society, Los Alamitos, CA, USA. doi:10.1109/CCECE.2002.1013101.
- [7] John J. Murphy. Technical Analysis of the Futures Market[M]. New York: New York Institute of Finance. 1986.
- [8] Kwon K. Y., Kish R. J. (2002). Technical Trading Strategies and Return Predictability: NYSE. *Applied Financial Economics*, 12(9), 639-653.
- [9] Lo A. W., Mamaysky H., Wang J. (2000). Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation. *The Journal of Finance*, 55(4), 1705-1770.
- [10] Maneesilp K., Prasatkaew C. (2014). Price Pattern Detection Using Finite State Machine with Fuzzy Transitions. *e-Business Engineering (ICEBE), 2014 IEEE 11th International Conference*, [C].
- [11] Prodromos E. Tsinaslanidis(2018). Subsequence dynamic time warping for charting: Bullish and bearish class predictions for NYSE stocks. *Expert Systems with Applications*, 94, 193-204.
- [12] Pruitt S. W., White R. E. (1988). The CRISMA trading system: who says technical analysis can't beat the market? [J]. *The Journal of Portfolio Management*, 14(3), 55-58.
- [13] Pruitt S. W., White R. E. (1988). The CRISMA trading system: who says technical analysis can't beat the market? *The Journal of Portfolio Management*, 14(3), 55-58.
- [14] Sakoe H., Chiba S. (1971). A Dynamic Programming Approach to Continuous Speech Recognition. In *Proceedings of the Seventh International Congress on Acoustics*, 3, 65-69.
- [15] Velichko V. M., Zagoruyko N. G. (1970). Automatic Recognition of 200 Words. *International Journal of Man-Machine Studies*, 2, 223-234.
- [16] Suh S. C. (2004). A Novel Chart Pattern Recognition Approach: A Case Study on Cup with Handle. *Proc of Artificial Neural Network in Engineering Conf*, St. Louis, Missouri, [C].

- [17] Syeda-Mahmood T., Beymer D., Wang F. (2007). Shape-Based Matching of ECG Recordings. In A. Dittmar, J. Clark, E. McAdams, N. Lovell (eds.), *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pp. 2012-2018. IEEE Computer Society, Los Alamitos, CA, USA. doi:10.1109/IEMBS.2007.4352714.
- [18] Tuzcu V., Nas S. (2005). Dynamic Time Warping as a Novel Tool in Pattern Recognition of ECG Changes in Heart Rhythm Disturbances. In M. Jamshidi, M. Johnson, P. Chen (eds.), *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, volume 1, pp. 182-186 Vol. 1. IEEE Computer Society, Los Alamitos, CA, USA. doi:10.1109/ICSMC.2005.1571142.
- [19] Zapranis A., Tsinaslanidis. P. E. (2012). A Novel, Rule-Based Technical Pattern Identification Mechanism: Identifying and Evaluating Saucers and Resistant Levels in the U.S. Stock Market. *Expert Systems with Applications*, 39(7). 6301-6308.
- [20] 陈静, 张川, 徐成贤. 基于特征提取的金融时间序列形态挖掘[J]. 统计与信息论坛, 2010, 25(6): 13-19.
- [21] 陈卓思, 宋逢明. 图形技术分析的信息含量[J]. 数量经济技术经济研究. 2005(9): 74-83.
- [22] 李莎, 李红刚. 股票市场中技术分析有效性的实证研究. 北京师范大学学报[J]: 自然科学版, 2009, 45(2): 212-214.
- [23] 马向前, 任若恩. 上海股市对技术信号的价格反应模型[J]. 北京航空航天大学学报: 社会科学版, 2003(2): 45-49.
- [24] 茆田杨, 肖焕雄. 中国股票市场行为的可预测性研究[J]. 预测, 1995(3): 55-57.
- [25] 孙碧波, 方健雯. 对中国证券市场弱态有效性的检验——基于技术分析获利能力的实证研究[J]. 上海财经大学学报, 2004, 6(6): 52-57.
- [26] 王志刚, 曾勇, 李平. 中国股票市场技术分析非线性预测能力的实证检验. 管理工程学报[J], 2009, 23(1): 149-153.
- [27] 张华, 任若恩. 外汇市场中非参数核回归方法的技术形态[J]. 哈尔滨工程大学学报, 2011, 32(12): 1642-1648.

7 附录

R codes: *SDTW Strategy*

URL: <https://pan.baidu.com/s/1yJDEdyiH0Cc0hZJg9pGFqA>



中國人民大學
RENMIN UNIVERSITY OF CHINA

作者簽名 胡志強