

# 趋势策略的深度学习增强

文巧钧 S0260517070001  
邮箱：wenqiaojun@gf.com.cn

广发证券金融工程

2017年9月

01

I

背景

>

02

II

循环神经网络模型

>

03

III

策略与实证

>

04

IV

总结

>

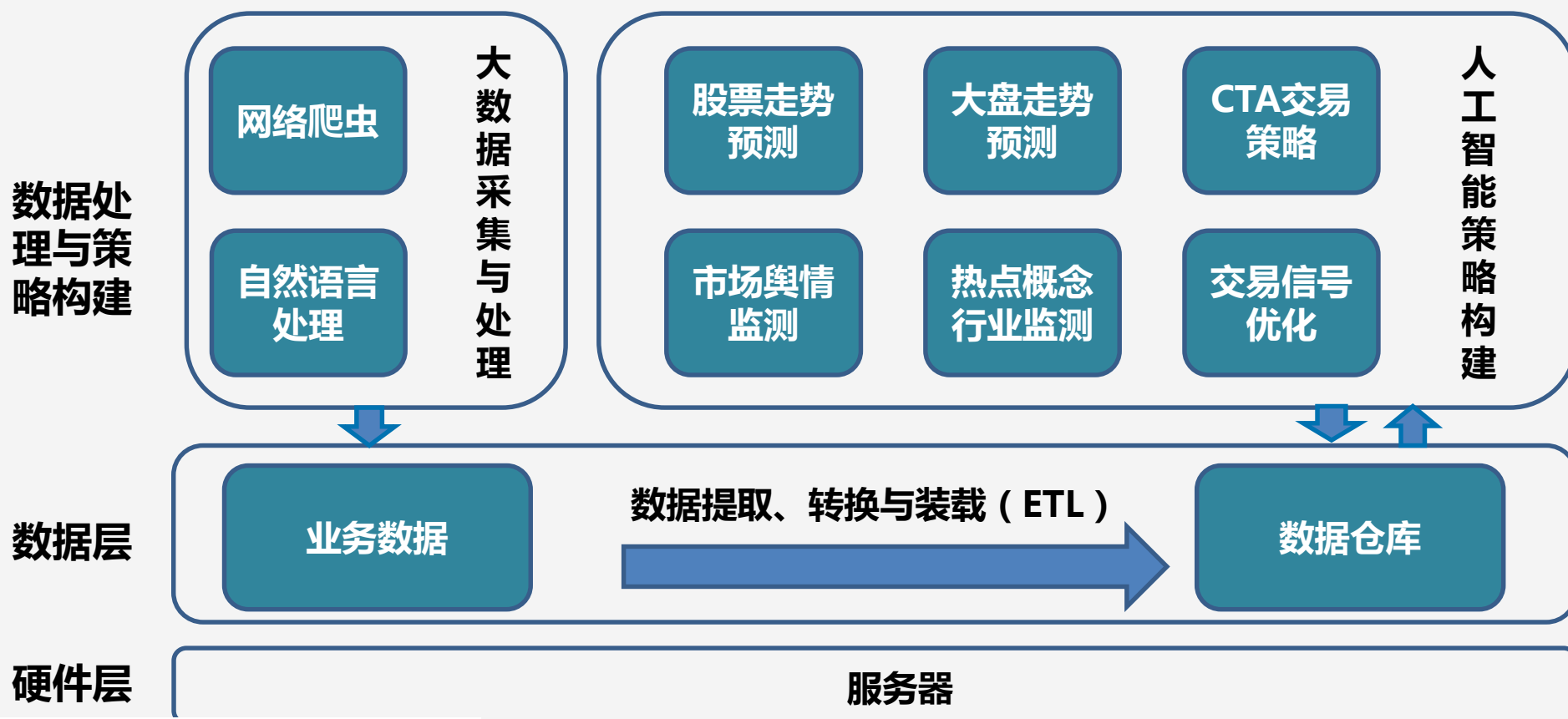


# 01

| 背景 |

>

人工智能方法：扩展信息的获取源，从海量数据中发现知识。



趋势交易中，我们可以筛选交易信号，获取盈利概率较高的趋势交易机会。

- EMDT策略：《另类交易策略系列之十四：经验模态分解下的日内趋势交易策略》
- 成份股一致性策略：《另类交易策略系列之三十二：从成份股一致性来衡量市场的趋势》

## EMDT策略

对沪深300指数期货的早盘走势进行分析，判断趋势是否强烈，实证表明趋势策略在早盘趋势强的市场容易获得收益。

### 经验模态分解趋势提取示意图

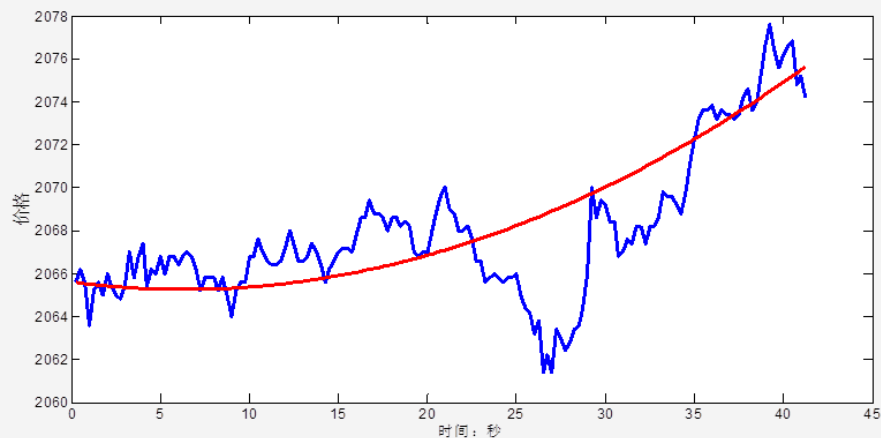
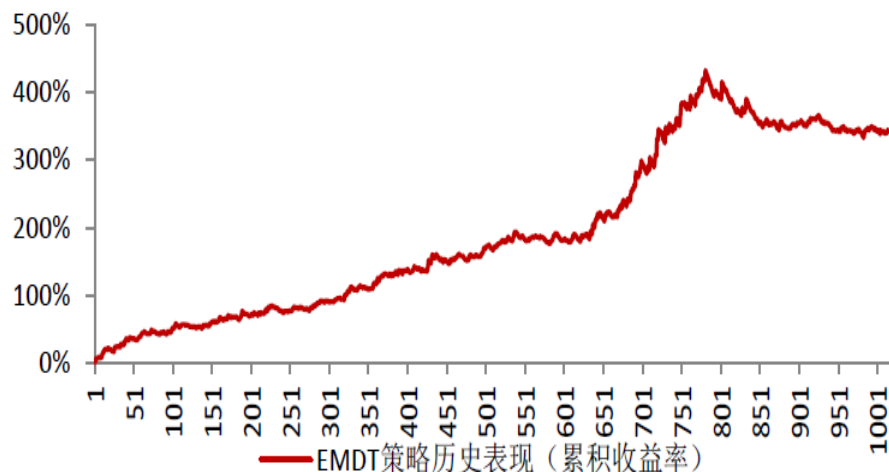


图4: EMDT策略历史表现 (自2010年4月16日起)

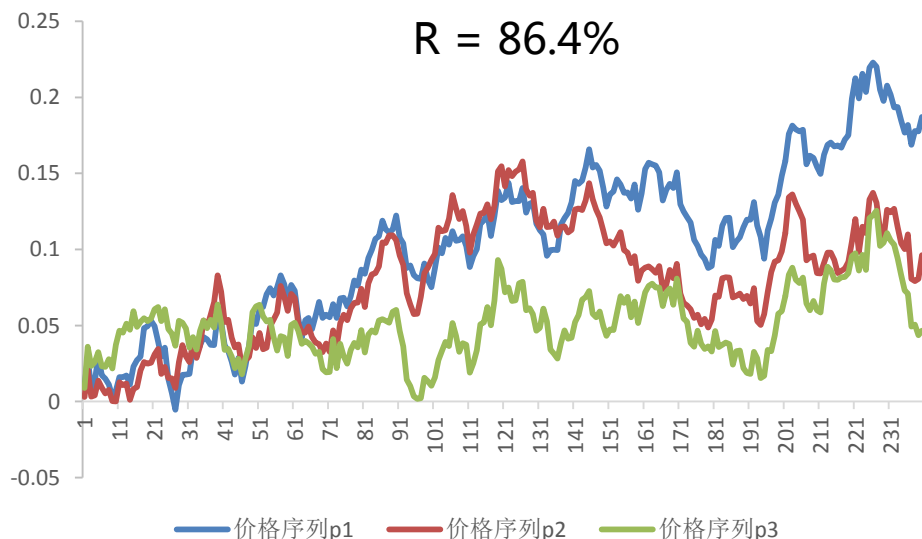


数据来源: 广发证券发展研究中心

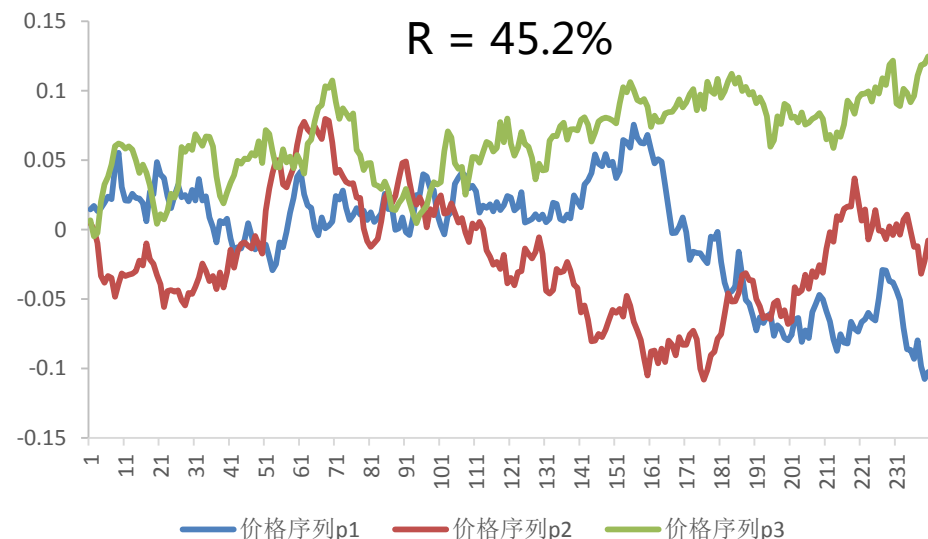
## 成份股一致性策略：

对沪深300指数成份股早盘走势进行分析，判断指数成份股走势是否一致，实证表明趋势策略在早盘成份股一致性强的市场容易获得收益。

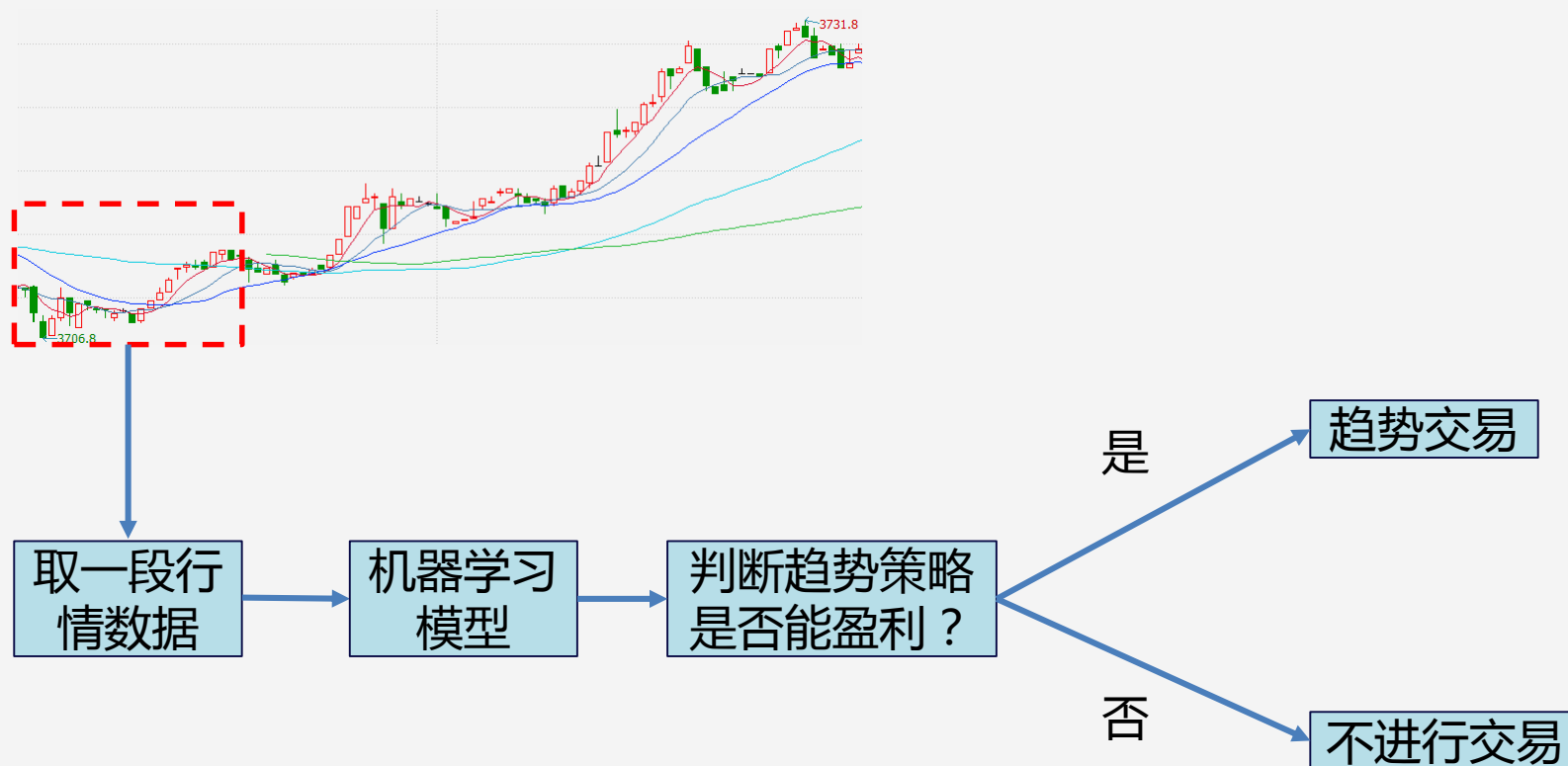
### 成份股一致性强



### 成份股一致性弱



能否通过机器学习的方法，判断市场是否适合进行趋势交易？







01

02

03

04

# 02

| 循环神经网络 |

>

## 线性模型

例如  $y = w_0 + w_1x_1 + w_2x_2$

问题是不能表示 X 和 Y 的非线性关系



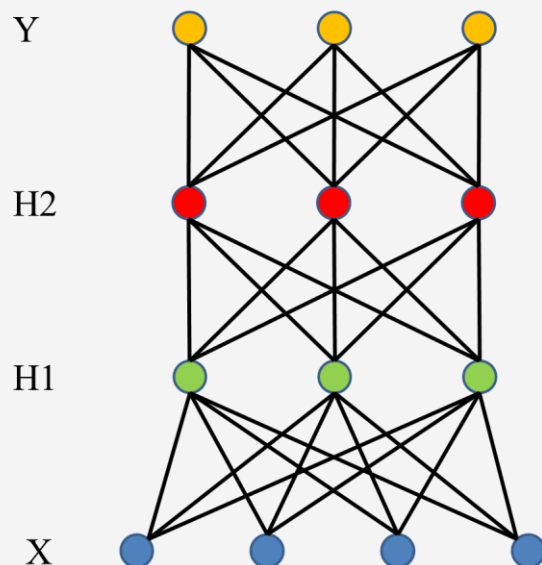
$\varphi: x \rightarrow \varphi(x)$  非线性映射, 例如  $\varphi(x) = x^2$

$$y = w_0 + w_1\varphi(x_1) + w_2\varphi(x_2)$$

通过非线性映射, 可以建立起 X 和 Y 的非线性关系

神经网络、支持向量机可以实现这种非线性映射, 映射  $\varphi$  称为特征映射

## 深层神经网络



$$y_k = \sigma_o \left\{ \sum_{j=1}^{N_2} (w_{kj}^{(2)} h_j^{(2)} + w_{k0}^{(2)}) \right\}$$

$$h_j^{(2)} = \sigma_h \left\{ \sum_{i=1}^{N_1} (w_{ji}^{(1)} h_i^{(1)} + w_{j0}^{(1)}) \right\}$$

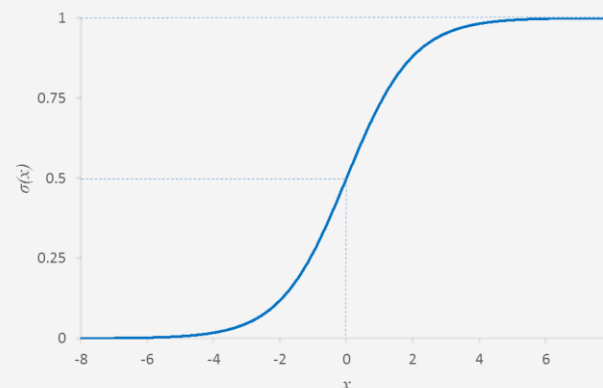
$$h_j^{(1)} = \sigma_h \left\{ \sum_{i=1}^{N_x} (w_{ji}^{(0)} x_i + w_{j0}^{(0)}) \right\}$$

隐层激活函数： $\sigma_h$

输出层激活函数： $\sigma_o$ ，对于分类问题，一般用 Sigmoid函数或者Softmax函数

二分类问题 
$$\sigma(\theta^T x) = \frac{e^{\theta^T x}}{1 + e^{\theta^T x}}$$

Sigmoid函数



待优化网络参数： $\mathbf{w}$

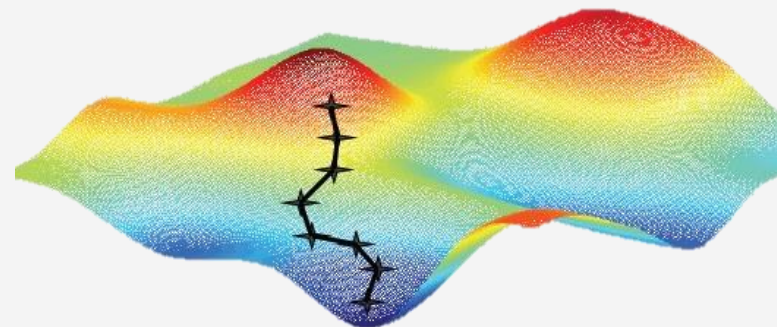
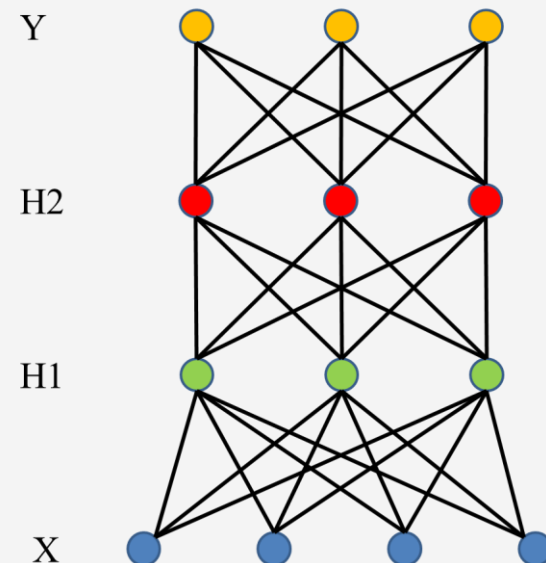
优化目标：最小化均方误差（MSE）

$$E(\mathbf{w}) = \frac{1}{NK} \sum_{n=1}^N \sum_{k=1}^K \left( \underset{\substack{\text{预测输出} \\ /}}{y_{nk}} - \underset{\substack{\text{实际标签} \\ \backslash}}{z_{nk}} \right)^2$$

参数优化方法：BP算法，迷你批量梯度下降算法

➤ 集成了梯度下降法和随机梯度下降法的特点

$$w_{ij}^{(n)} := w_{ij}^{(n-1)} - \alpha \frac{\partial}{\partial w_{ij}} \sum_{n_k \in \text{Batch}(n)} E_{n_k}(\mathbf{w}^{(n-1)})$$



有时候，机器学习任务需要神经网络具有“记忆力”

自然语言处理问题：

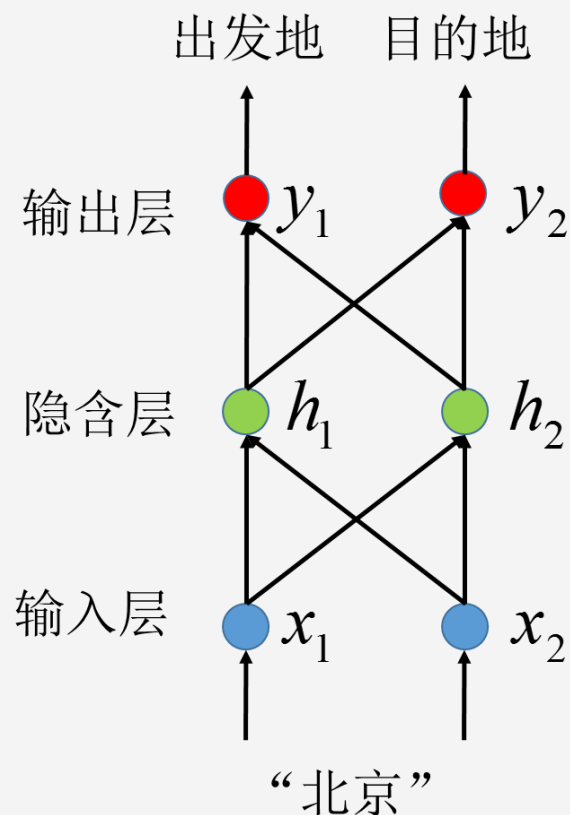
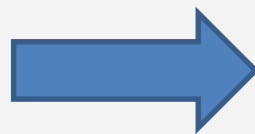
语义的识别需要考虑上下文关系。

例如，识别出发地和目的地

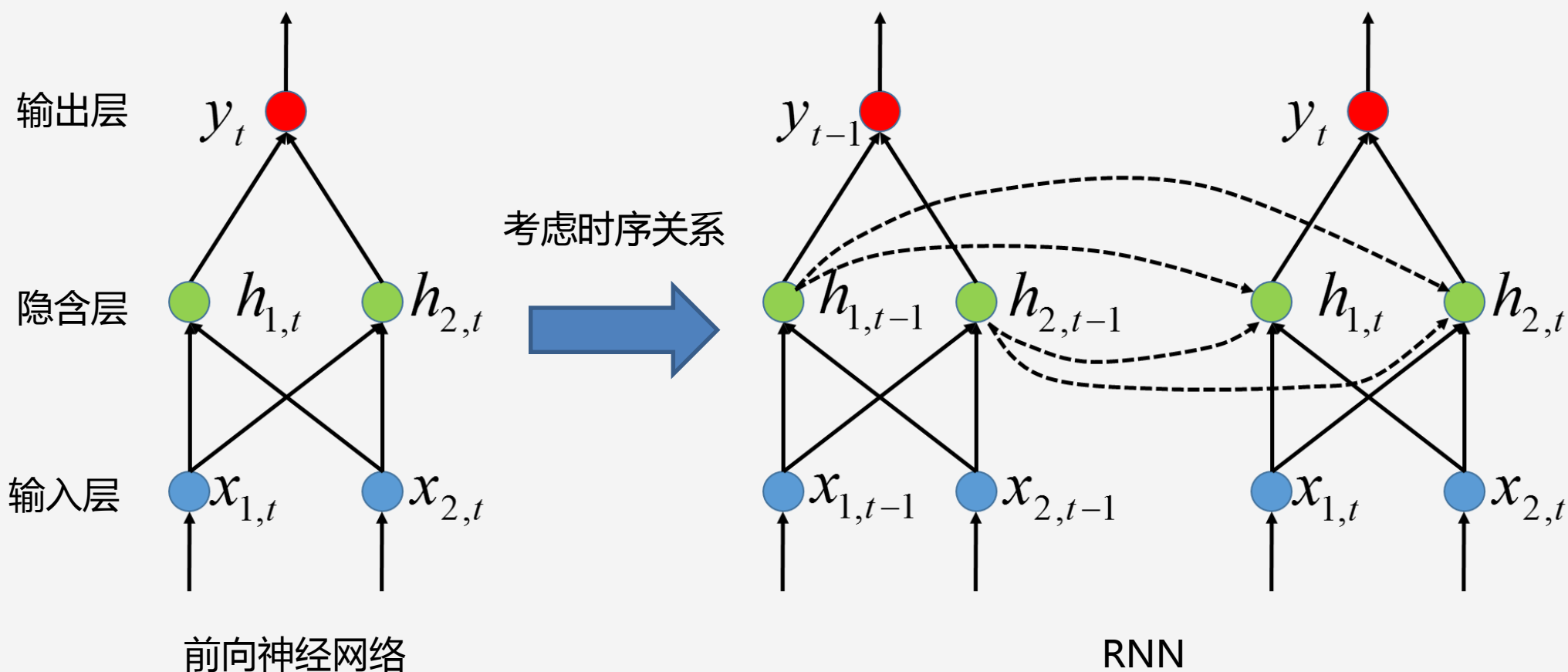
- “小明/今天/到达/北京。”
- “小红/今天/离开/北京。”

机器在理解语言的时候怎么判断“北京”  
是出发地还是目的地呢？

金融市场数据中同样存在时间依赖关系。



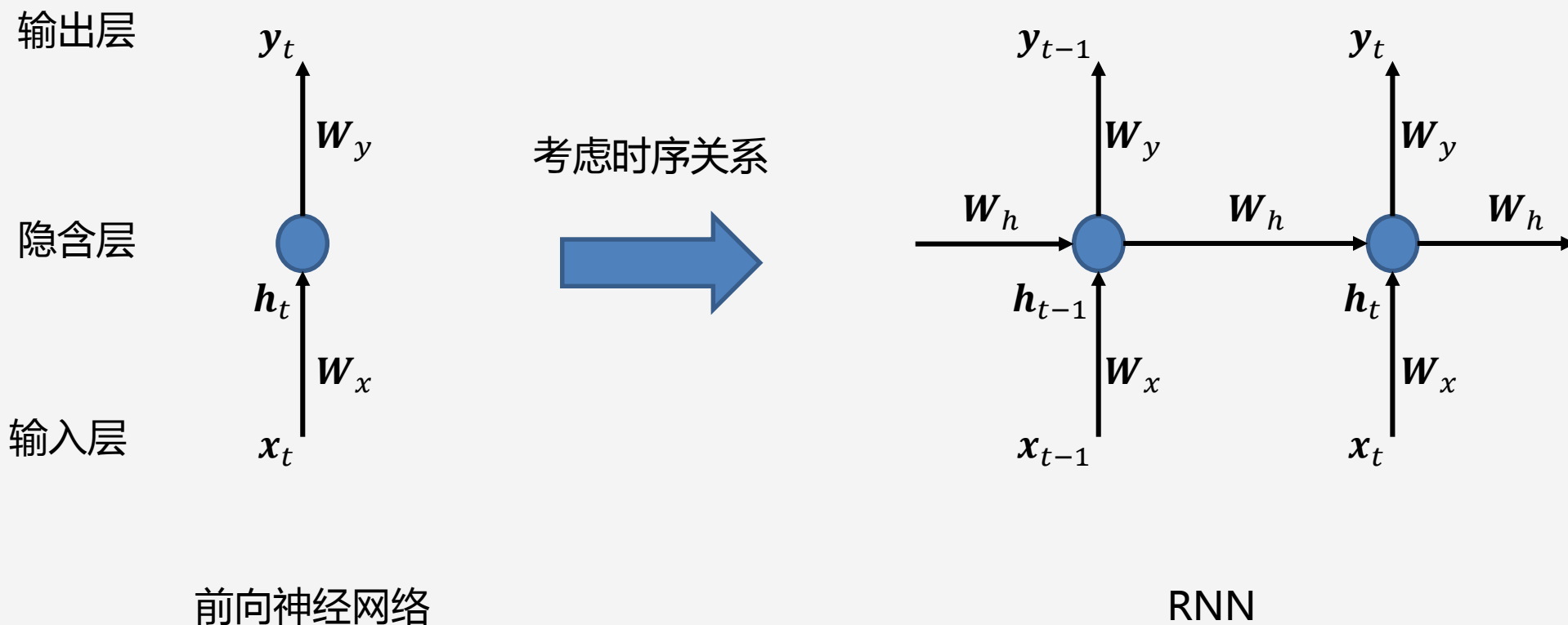
循环神经网络 ( RNN )：通过隐含层的状态来描述数据的时序关系



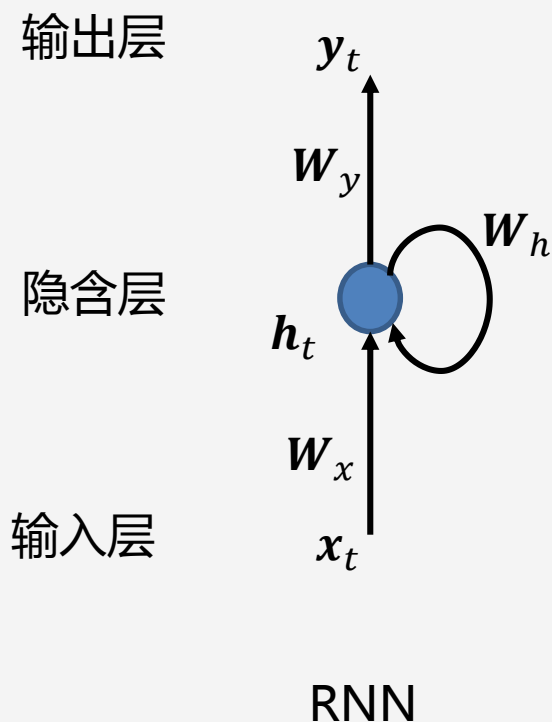
循环神经网络（RNN）：权重矩阵 $W_x$ ， $W_y$ ， $W_h$ 基于时间共享

通过RNN的结构，可以对输入输出数据的时间依赖关系建模。

RNN在结构上类似隐马尔科夫模型（HMM）。



## 循环神经网络 ( RNN ) 与传统时间序列模型的关系



输出： $y_t = \sigma_o(W_y h_t + b_o)$

隐状态： $h_t = \sigma_h(W_x x_t + W_h h_{t-1} + b_h)$

考虑单变量的情况，

记 $x_t = x_t$ ，记 $W_h = \beta$ ， $W_x = \alpha$ ，令输出 $y_t = y_t = h_t$ ，而且令 $\sigma_h$ 为恒等变换，

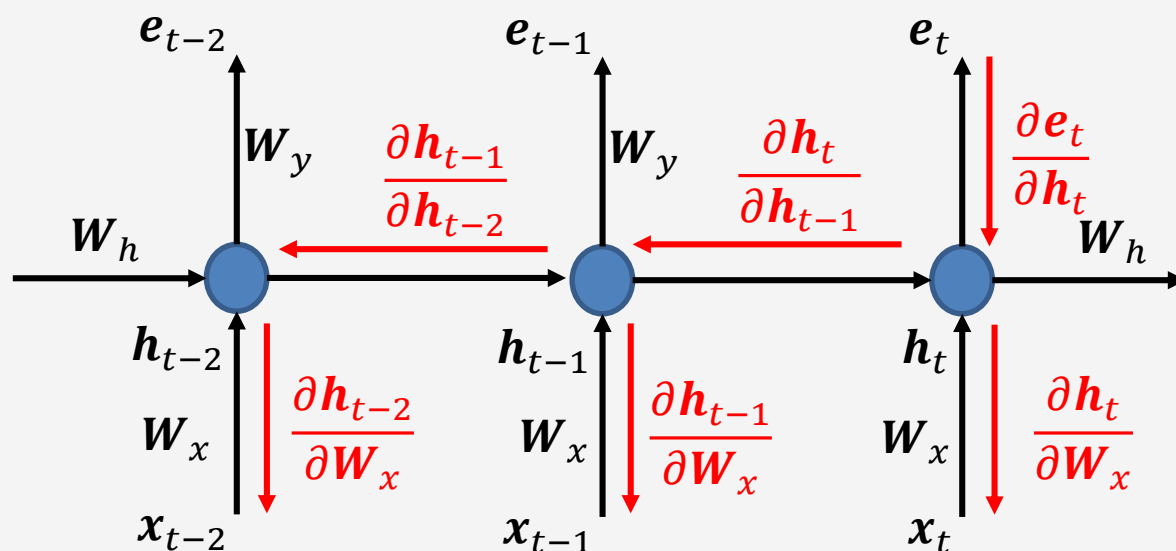
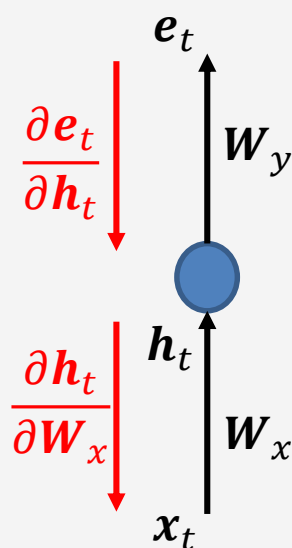
则RNN可写成

$$y_t = \alpha x_t + \beta y_{t-1} + b$$

这是一个带外生变量的一阶自回归模型。



## RNN的参数优化算法：BPTT ( Back Propagation Through Time )



### 前向神经网络的BP算法

$$\frac{\partial e_t}{\partial W_x} = \frac{\partial e_t}{\partial h_t} \frac{\partial h_t}{\partial W_x}$$

### RNN的BPTT算法

$$\begin{aligned} \frac{\partial e_t}{\partial W_x} = & \frac{\partial e_t}{\partial h_t} \frac{\partial h_t}{\partial W_x} + \frac{\partial e_t}{\partial h_t} \frac{\partial h_t}{\partial h_{t-1}} \frac{\partial h_{t-1}}{\partial W_x} \\ & + \frac{\partial e_t}{\partial h_t} \frac{\partial h_t}{\partial h_{t-1}} \frac{\partial h_{t-1}}{\partial h_{t-2}} \frac{\partial h_{t-2}}{\partial W_x} + \dots \end{aligned}$$

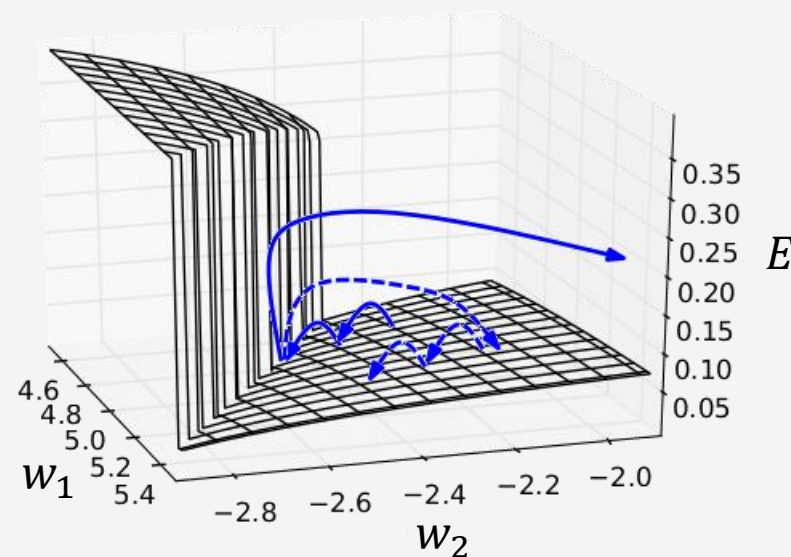
RNN中参数优化的问题：梯度消失和梯度爆炸

因为梯度的爆炸与消失问题，RNN的参数学习困难：

假设激活函数为线性函数，则

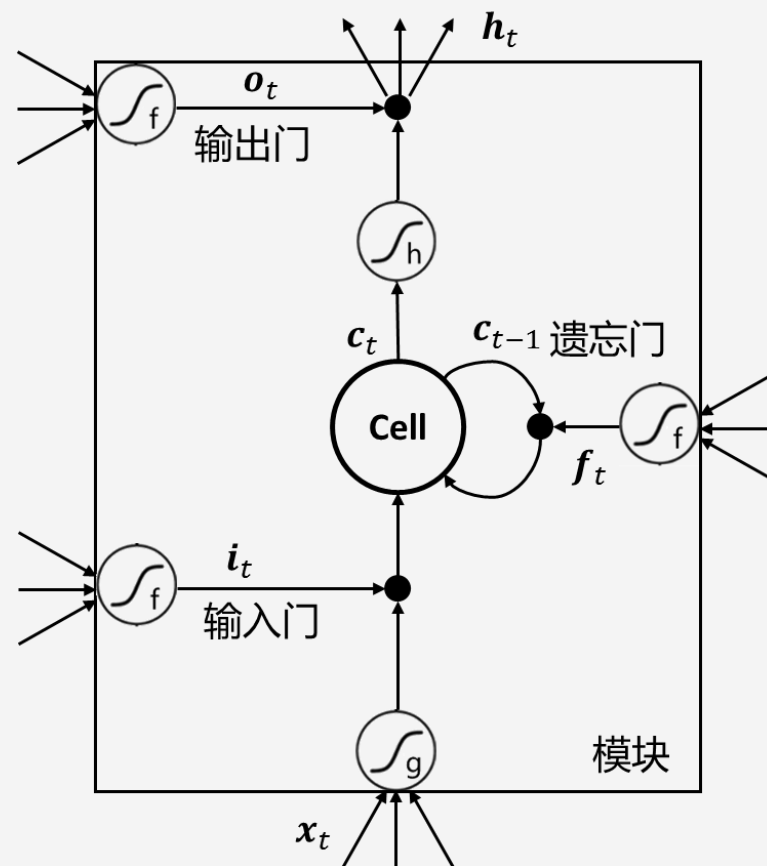
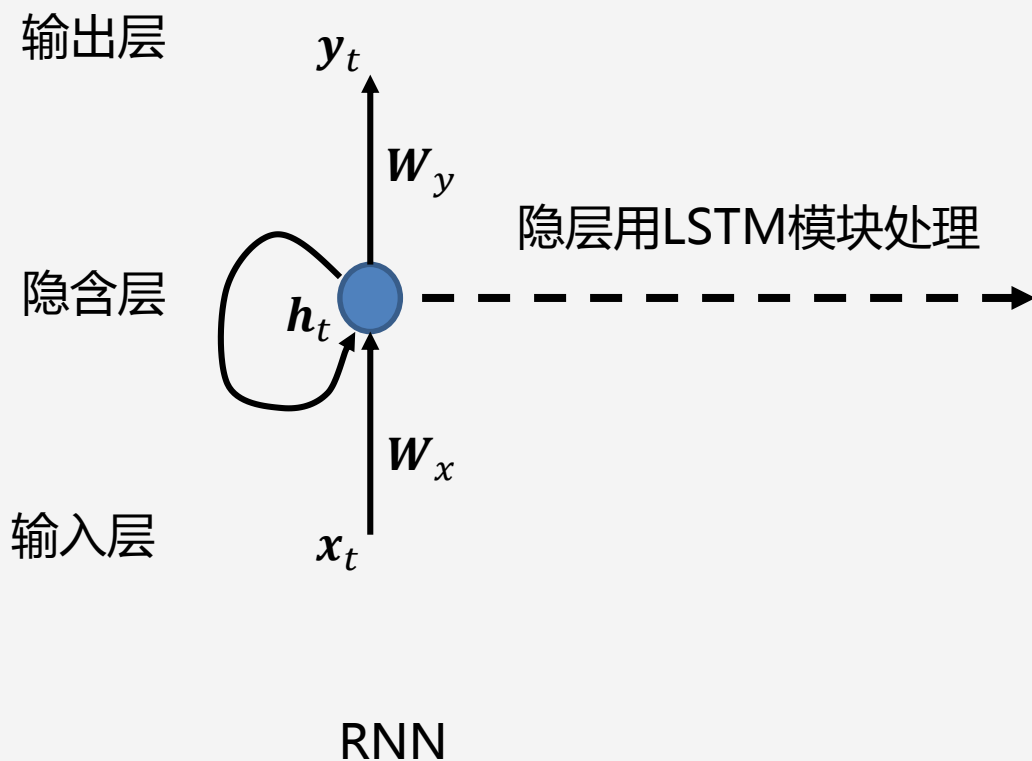
- 当 $\|W_h\| > 1$ 时，梯度爆炸
- 当 $\|W_h\| < 1$ 时，梯度消失

在sigmoid等其他激活函数下，也有类似的问题。其中 $\|W_h\|$ 为 $W_h$ 的L2-范数。



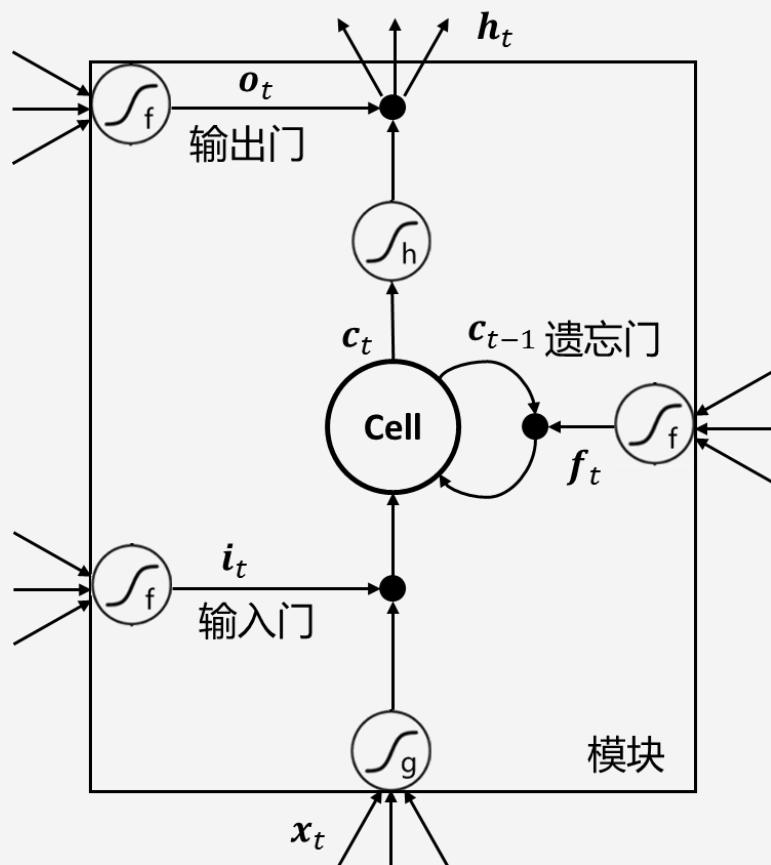
**结论：普通的RNN很难对数据的长周期时序依赖关系进行建模！**

解决方案：长短期记忆单元LSTM ( Long Short-Term Memory )



LSTM模块示意图

LSTM使用“门”，有选择的让一部分信息通过。



输入门：决定什么样的新信息存入细胞状态

$$i_t = \sigma(W^{(xi)}x_t + W^{(hi)}h_{t-1} + b^{(i)})$$

遗忘门：决定原细胞状态是否需要保留

$$f_t = \sigma(W^{(xf)}x_t + W^{(hf)}h_{t-1} + b^{(f)})$$

Cell状态更新：

$$c_t = f_t \otimes c_{t-1} + i_t \otimes \tanh(W^{(xc)}x_t + W^{(hc)}h_{t-1} + b^{(c)})$$

输出门：决定输出哪些细胞状态信息

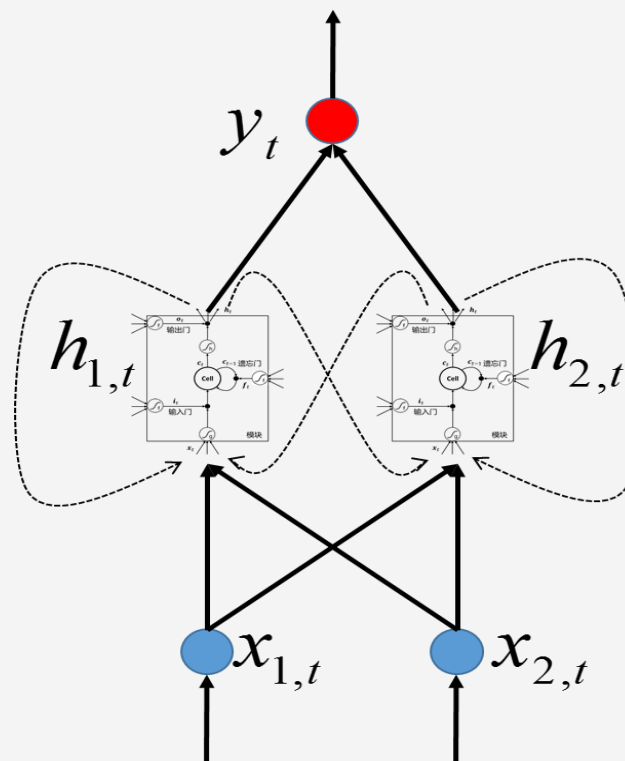
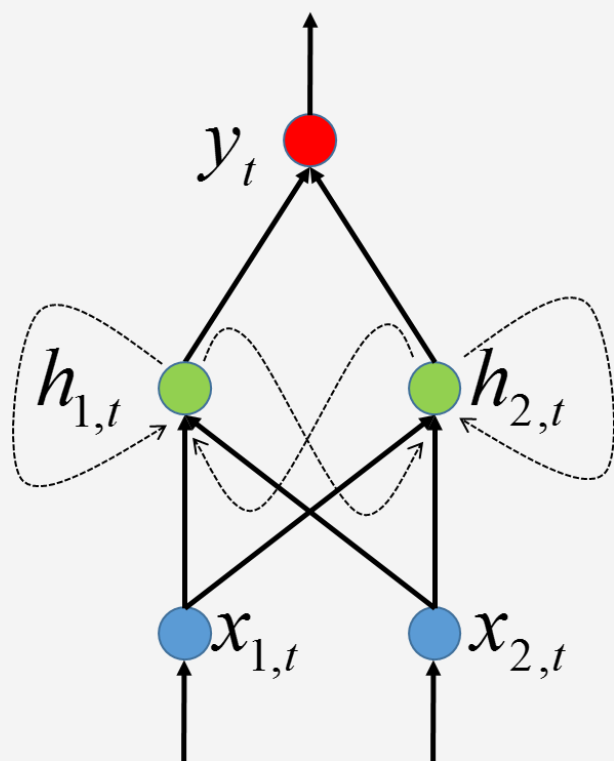
$$o_t = \sigma(W^{(xo)}x_t + W^{(ho)}h_{t-1} + b^{(o)})$$

LSTM单元输出

$$h_t = o_t \otimes \tanh(c_t)$$

带LSTM的RNN示意图：用LSTM单元取代原来RNN中的隐层节点

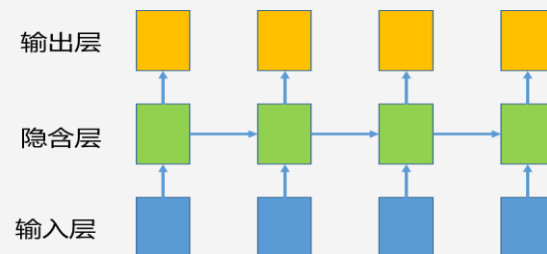
参数优化：BPTT算法



带LSTM单元的RNN

## LSTM应用场景示例

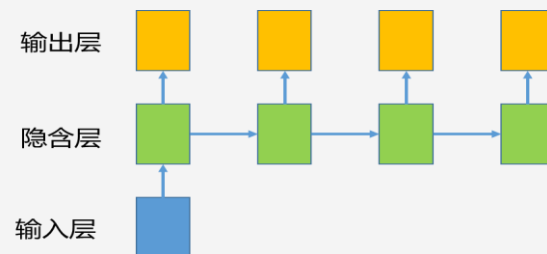
时间序列分析——输入输出序列长度一致



文本情绪识别——多个输入对应一个输出



场景描述——一个输入对应多个输出

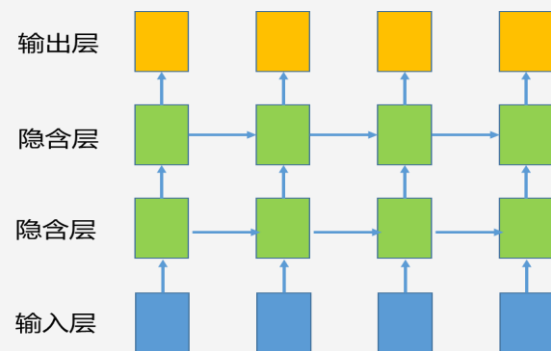
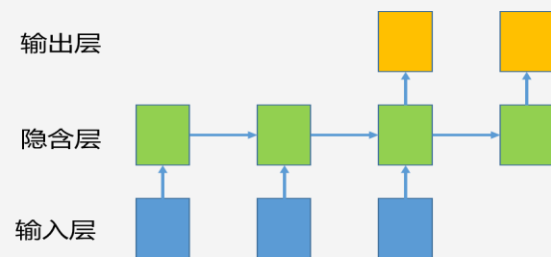
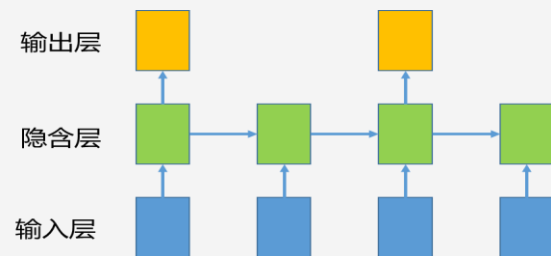


## LSTM应用场景示例

语音识别——许多帧的语音输入，对应少量文本输出

机器翻译——多个输入对应多个输出（长度不确定，输出序列可能比输入序列长或者短）

深层的LSTM





# 03

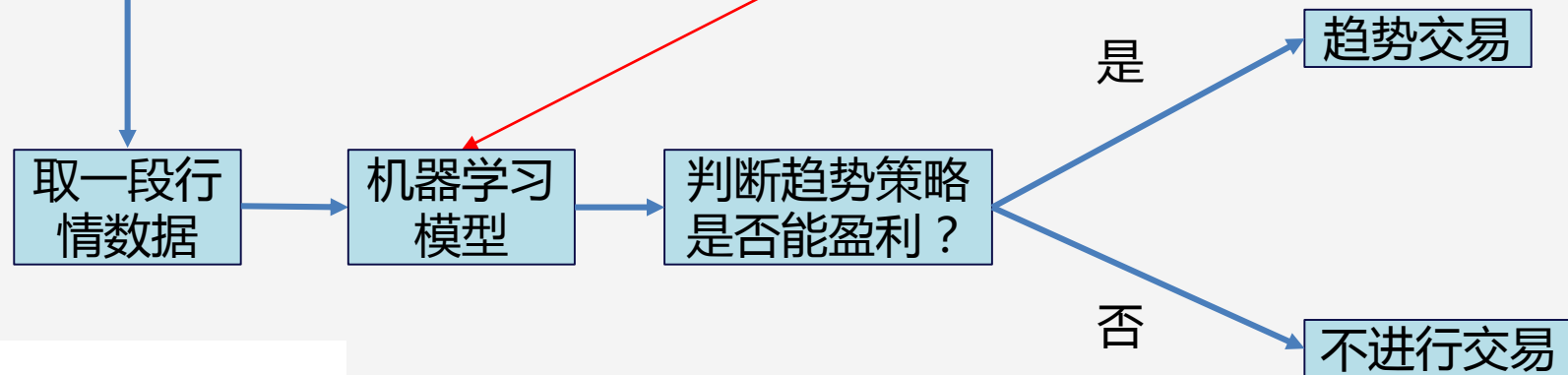
| 策略与实证分析 |

>



## 策略原理

每个交易日，取一段早盘行情数据，判断当日趋势策略能否盈利。  
若判断可以盈利，则按照早盘走势进行趋势跟踪，否则不开仓交易。

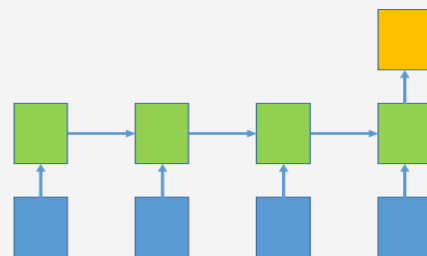


RNN模型

输出层

隐含层

输入层



样本内：如何衡量当天趋势策略能否盈利？

方法1：回测，用开盘区间突破策略等方法回测，计算当天趋势策略是否盈利。

方法2：信号处理，如经验模态分解、计算Hurst指数等。

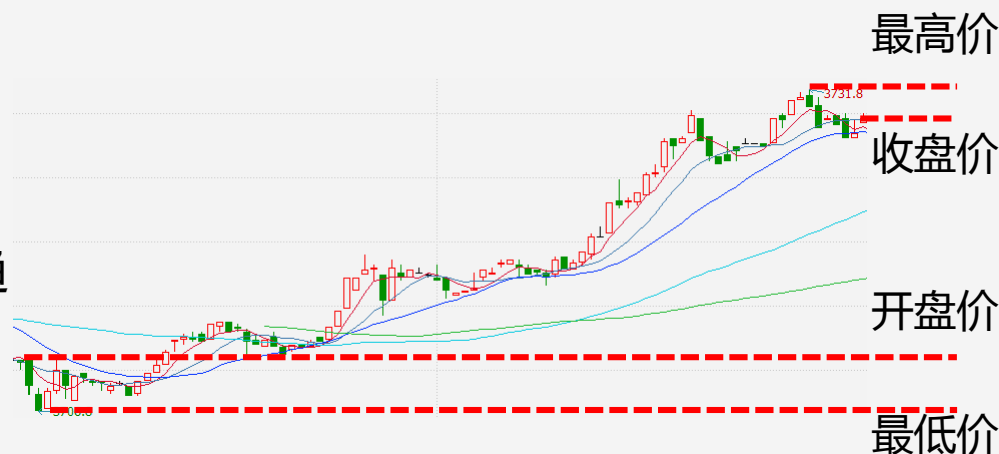
方法3：其他方法，如K线中实体的占比

注意：样本内处理时，可以用到全天的行情信息；样本外回测和实际交易时只有早盘信息！

定义趋势策略盈利指标  $R = \frac{|Close - Open|}{High - Low}$

当  $R$  越大时，表示当天趋势策略容易盈利。

按照  $R > 0.5$  和  $R < 0.5$  将不同交易日期的趋势策略盈利状态分成两类。



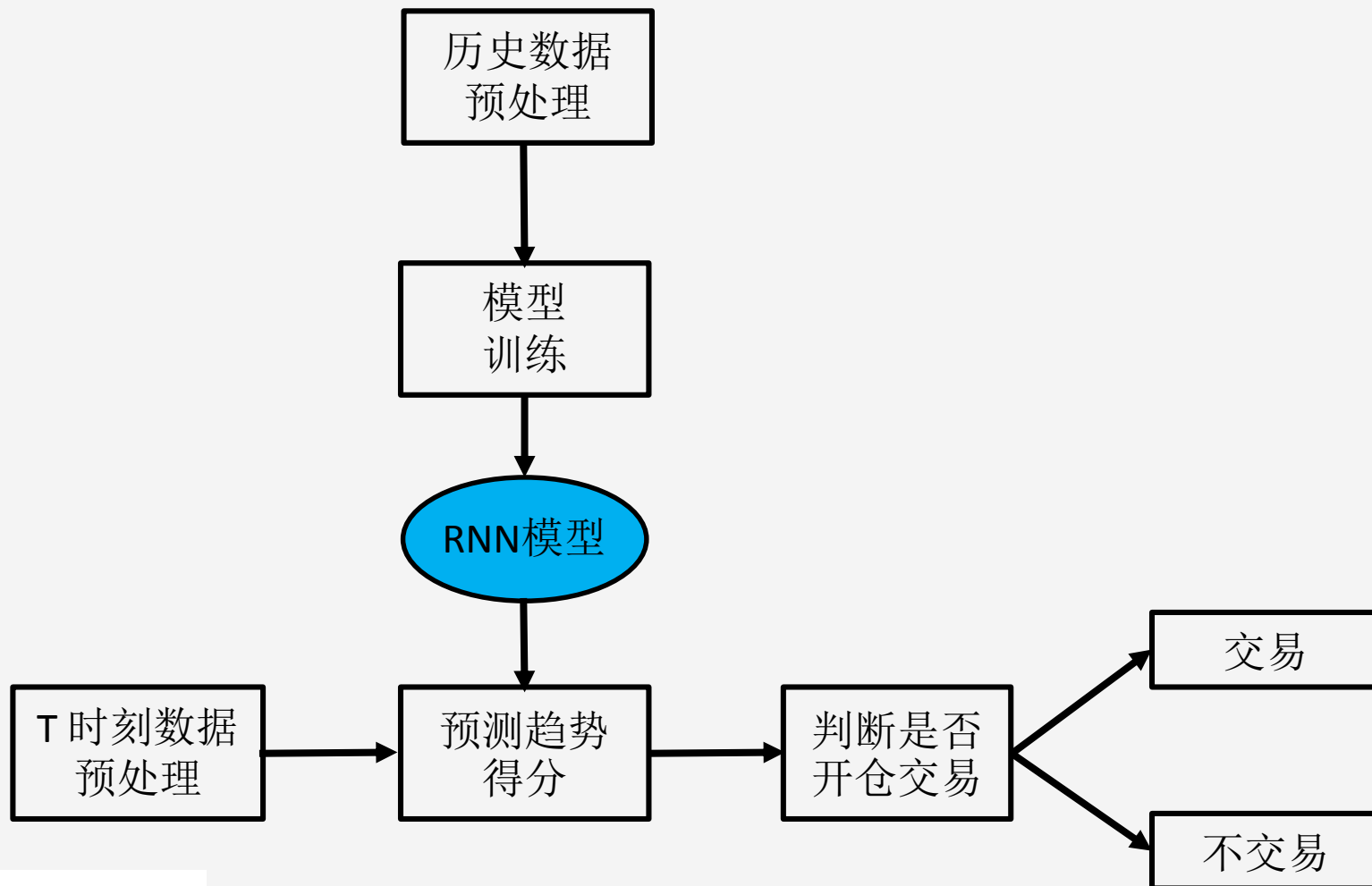
## 回测设置

- **交易标的**：IF股指期货主力合约
- **RNN模型训练**：2010年4月16日至2013年12月31日
- **回测期**：2014年1月1日至2017年7月31日
- **特征数据序列**：分钟行情，取分钟行情的开盘价、收盘价、最高价、最低价、成交量、主买量、主卖量、收盘价变化率、收盘价序列的二阶差分、主买主卖量之比、成交量变化率、主买量变化率、主卖量变化率
- **回测保证金设置**：100%，即不考虑杠杆
- **交易成本**：万分之二

## 策略流程

模型训练

交易



模型训练：通过训练样本训练，优化模型参数

预测输出  $Y$  的维度：1

输入特征  $X$  的维度：13维时间序列

数据预处理：异常值处理、数据标准化

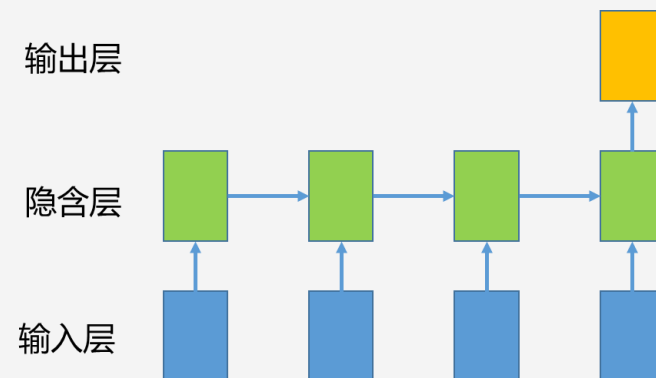
选取模型结构为：

13（输入层）-200（LSTM隐层）-1（输出层）

即用3层的RNN，包括1个输入层，1个输出层和一个隐层，隐层节点个数为200。

输出层激活函数：sigmoid函数

优化目标函数：交叉熵



## 模型表现：

从开盘后33分钟来预测全天的  $R$ ，样本外预测准确率59.1%。

RNN模型对 $R$ 的预测结果（盈利概率  $p$ ）如右图所示。

交易策略：  
每天计算  $T$  时刻概率  $p$  的120日均线，若  
$$p > MA120(p)$$
  
则进行趋势交易。

比较对应时段的股指期货涨跌幅，确定开仓方向

$$p(T) < p(\text{开盘})$$

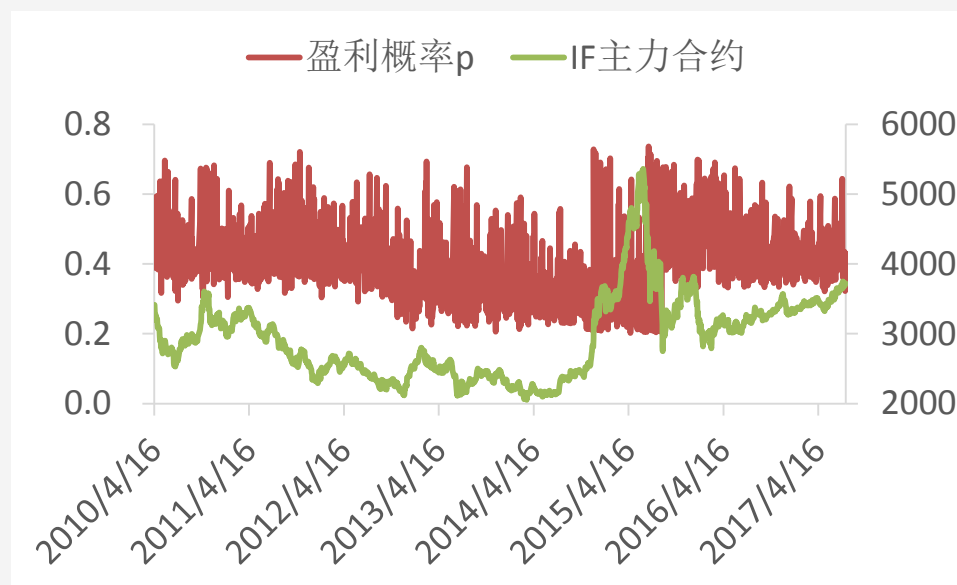


做空

$$p(T) > p(\text{开盘})$$



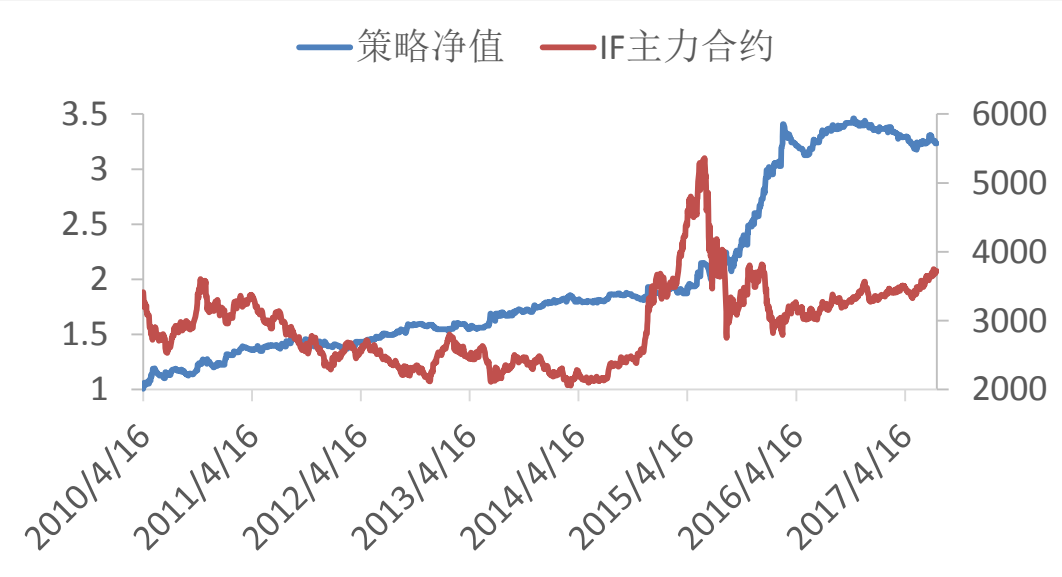
做多



盈利概率  $p$  的走势

策略表现（全样本）

2010年以来策略净值

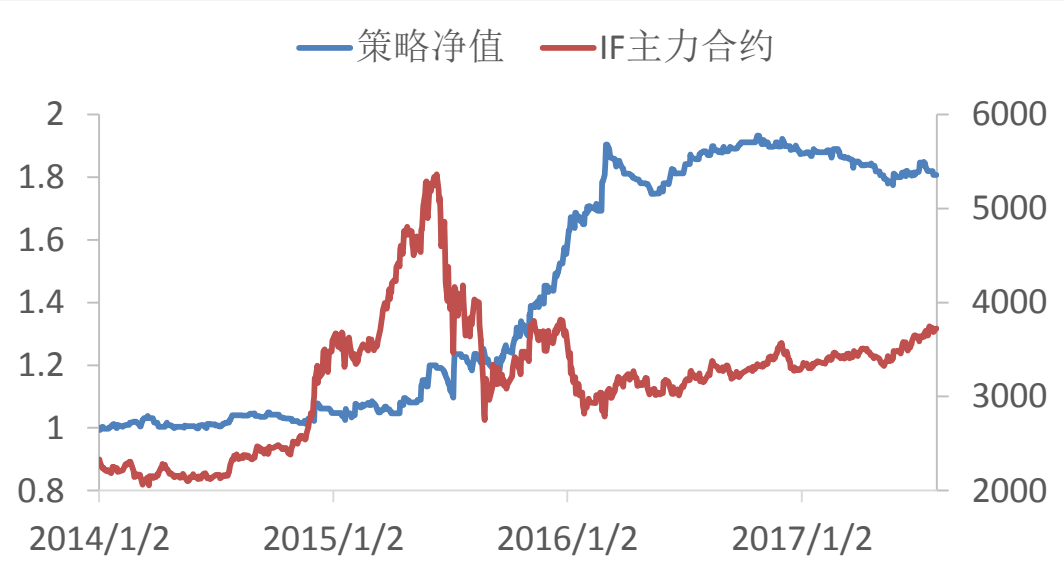


2010年以来策略表现

年化收益率	18.01%
累计收益率	223.46%
最大回撤	-8.63%
交易次数	707
胜率	40.74%
盈亏比	2.17
单次交易收益率	0.17%

策略表现（样本外）

2014年以来策略净值



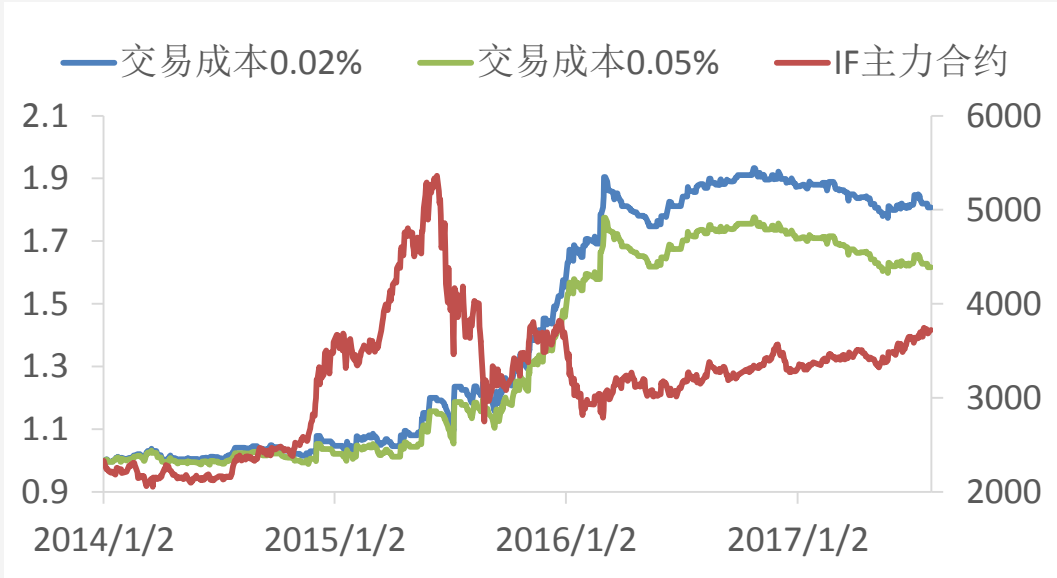
2014年以来策略表现

年化收益率	18.47%
累计收益率	80.72%
最大回撤	-8.63%
交易次数	372
胜率	39.52%
盈亏比	2.27
单次交易收益率	0.17%



交易成本：2015年9月之后，股指期货的流动性变差，实际交易成本提高。

不同交易成本下的策略净值



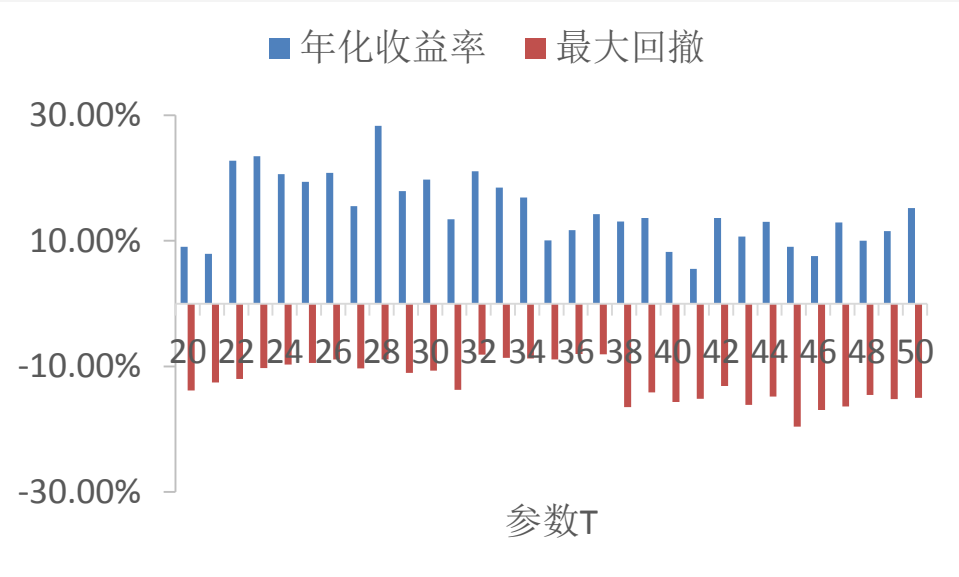
不同交易成本下策略比较

	交易成本 0.05%	交易成本 0.02%
年化收益率	14.71%	18.47%
累计收益率	61.66%	80.72%
最大回撤	-10.00%	-8.63%
交易次数	372	372
胜率	38.44%	39.52%
盈亏比	2.20	2.27
单次交易 收益率	0.14%	0.17%

不同参数下的策略表现（样本外）

参数 (分钟)	年化 收益率	最大 回撤	参数 (分钟)	年化 收益率	最大 回撤
20	9.06%	-13.84%	36	11.69%	-8.02%
21	7.92%	-12.54%	37	14.25%	-8.06%
22	22.77%	-11.98%	38	13.07%	-16.50%
23	23.47%	-10.29%	39	13.64%	-14.13%
24	20.61%	-9.71%	40	8.23%	-15.65%
25	19.40%	-9.47%	41	5.51%	-15.14%
26	20.79%	-8.89%	42	13.63%	-13.11%
27	15.53%	-10.31%	43	10.68%	-16.14%
28	28.32%	-8.87%	44	13.04%	-14.81%
29	17.92%	-11.03%	45	9.06%	-19.58%
30	19.75%	-10.66%	46	7.59%	-16.92%
31	13.41%	-13.72%	47	12.91%	-16.40%
32	21.09%	-8.15%	48	10.03%	-14.55%
33	18.47%	-8.63%	49	11.54%	-15.18%
34	16.88%	-8.71%	50	15.23%	-15.00%
35	10.07%	-8.90%			

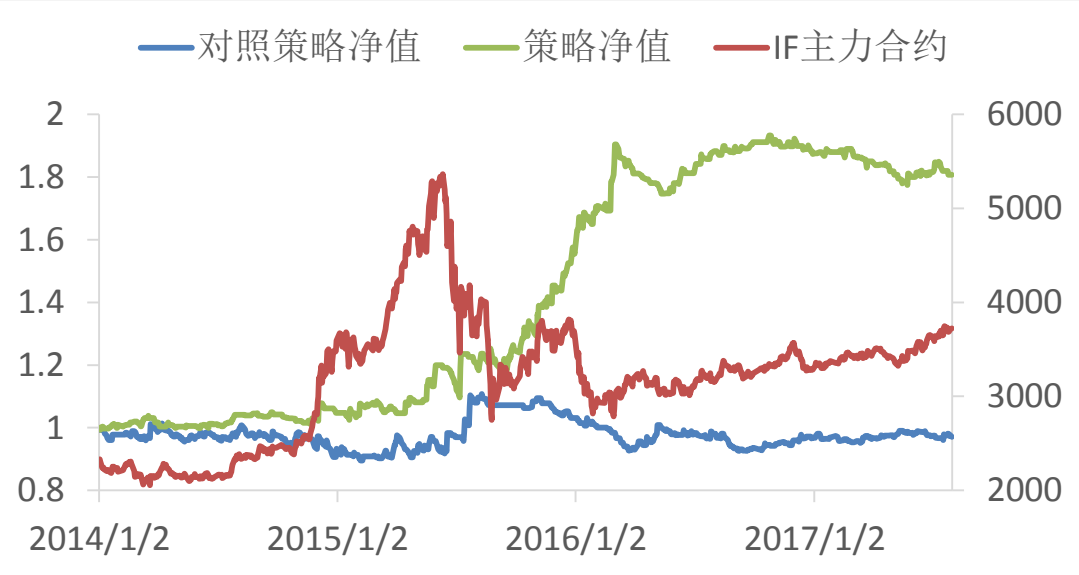
样本外年化收益与最大回撤



对照策略表现：

对照策略：当  $p$  小于120日均线时，进行趋势交易

2014年以来对照策略表现



2014年策略表现对比

	对照策略	交易策略
年化收益率	-0.84%	18.47%
累计收益率	-2.90%	80.72%
最大回撤	-16.33%	-8.63%
交易次数	501	372
胜率	0.387226	39.52%
盈亏比	1.58	2.27
单次交易收益率	0.00%	0.17%



# 04

| 总结 |

>

- 通过循环神经网络，对市场的状况进行评估，判断当天是否适合进行趋势交易。
- 从2014年以来的样本外回测表明，策略在样本外的年化收益18.47%，最大回撤为-8.63%。
- 实证表明，盈利机会较小的交易信号被过滤。即使提高交易成本，策略依然有不俗的表现。

本文旨在对所研究问题的主要关注点进行分析，因此对市场及相关交易做了一些合理假设，但这样会导致建立的模型以及基于模型所得出的结论并不能完全准确地刻画现实环境。而且由于分析时采用的相关数据都是过去的时间序列，因此可能会与未来真实的情况出现偏差。本文内容并不是适合所有的投资者，客户在制定投资策略时，必须结合自身的环境和投资理念。

广发证券股份有限公司（以下简称“广发证券”）具备证券投资咨询业务资格。本报告只发送给广发证券重点客户，不对外公开发布，只有接收客户才可以使用，且对于接收客户而言具有相关保密义务。广发证券并不因相关人员通过其他途径收到或阅读本报告而视其为广发证券的客户。本报告的内容、观点或建议并未考虑个别客户的特定状况，不应被视为对特定客户关于特定证券或金融工具的投资建议。本报告发送给某客户是基于该客户被认为有能力独立评估投资风险、独立行使投资决策并独立承担相应风险。

本报告所载资料的来源及观点的出处皆被广发证券股份有限公司认为可靠，但广发证券不对其准确性或完整性做出任何保证。报告内容仅供参考，报告中的信息或所表达观点不构成所涉证券买卖的出价或询价。广发证券不对因使用本报告的内容而引致的损失承担任何责任，除非法律法规有明确规定。客户不应以本报告取代其独立判断或仅根据本报告做出决策。

广发证券可发出其它与本报告所载信息不一致及有不同结论的报告。本报告反映研究人员的不同观点、见解及分析方法，并不代表广发证券或其附属机构的立场。报告所载资料、意见及推测仅反映研究人员于发出本报告当日的判断，可随时更改且不予通告。

本报告旨在发送给广发证券的特定客户及其它专业人士。未经广发证券事先书面许可，任何机构或个人不得以任何形式翻版、复制、刊登、转载和引用，否则由此造成的一切不良后果及法律责任由私自翻版、复制、刊登、转载和引用者承担。

Thanks !  
谢谢