

WOT
2015



51CTO 传媒

移动互联网 开发者大会

洞察移动互联网用户行为 分享移动应用研发实践



股市中的R语言量化算法模型

均值回归，发现逆市中的投资机会

主讲：张丹





目录



1. 均值回归原理
2. 均值回归模型和实现
3. 量化选股



均值回归原理

WOT
2015 | World of Tech
WOT

51CTO 传媒

- 在股票市场中有两种典型的投资策略：趋势追踪 和 均值回归。
- **趋势追踪策略**：在大行情中波段操作，如[均线模型](#)，不仅简单而且有效，我之前写的一篇文章，[两条均线打天下](#) 就属于趋势追踪策略。
- **均值回归策略**：在震荡行情中，找到[超跌](#)的股票买入，等待上涨后卖出，捕捉小的机会，本次就介绍均值回归的模型。



均值回归原理

- 在金融学中，均值回归是指股票价格无论高于或低于**均衡价格水平(均值)**，都会以很高的概率向均值回归。根据这个理论，股票价格总是围绕其均值上下波动。
- 上涨或下跌的趋势，不管延续多长时间，不能永远持续下去。涨得太多了就会跌，跌得太多就会涨。
- 简单地说，“**涨多必跌，跌多必涨**”。



均值回归原理

WOT
2015 | World of Tech
WOT

51CTO 传媒

- 下面以平安银行(000001)股票日线图为例，截取2005年到2015年7月的股票数据，股价为向前复权的价格。





均值回归原理

WOT
2015 | World of Tech
WOT

51CTO 传媒

均值回归是价值投资理论成立的一个核心理论。

具有3个特性：

- 必然性
- 不对称性
- 政府调控



均值回归原理

WOT
2015 | World of Tech
WOT

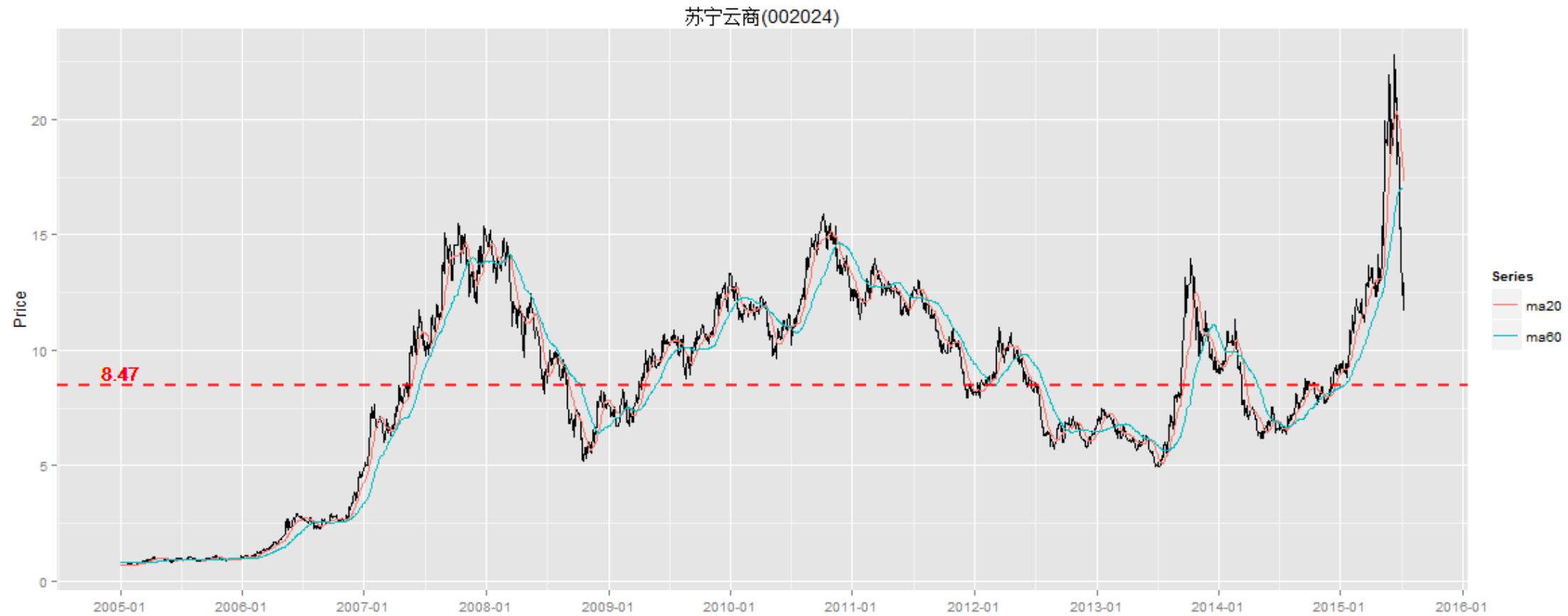
51CTO 传媒

- **必然性**，股票价格不能总是上涨或下跌，一种趋势不管其持续的时间多长都不能永远持续下去。
- 在一个趋势内，股票价格呈持续上涨或下跌，称为**均值回避**。
- 当出现反趋势的情况就是**均值回归**，但回归周期有随机性不能预测。不同的股票市场，回归的周期是不一样的，就算是相同的股票市场，回归的周期也是不一样的。



均值回归原理

- 以苏宁云商(002024)股票日线图为例, 同样截取2005年到2015年7月的向前复权的股价数据。





均值回归原理

WOT
2015 | World of Tech
WOT

51CTO 传媒

- **不对称性**，股价波动的幅度与速度是不一样的，回归时的幅度与速度具有随机性。





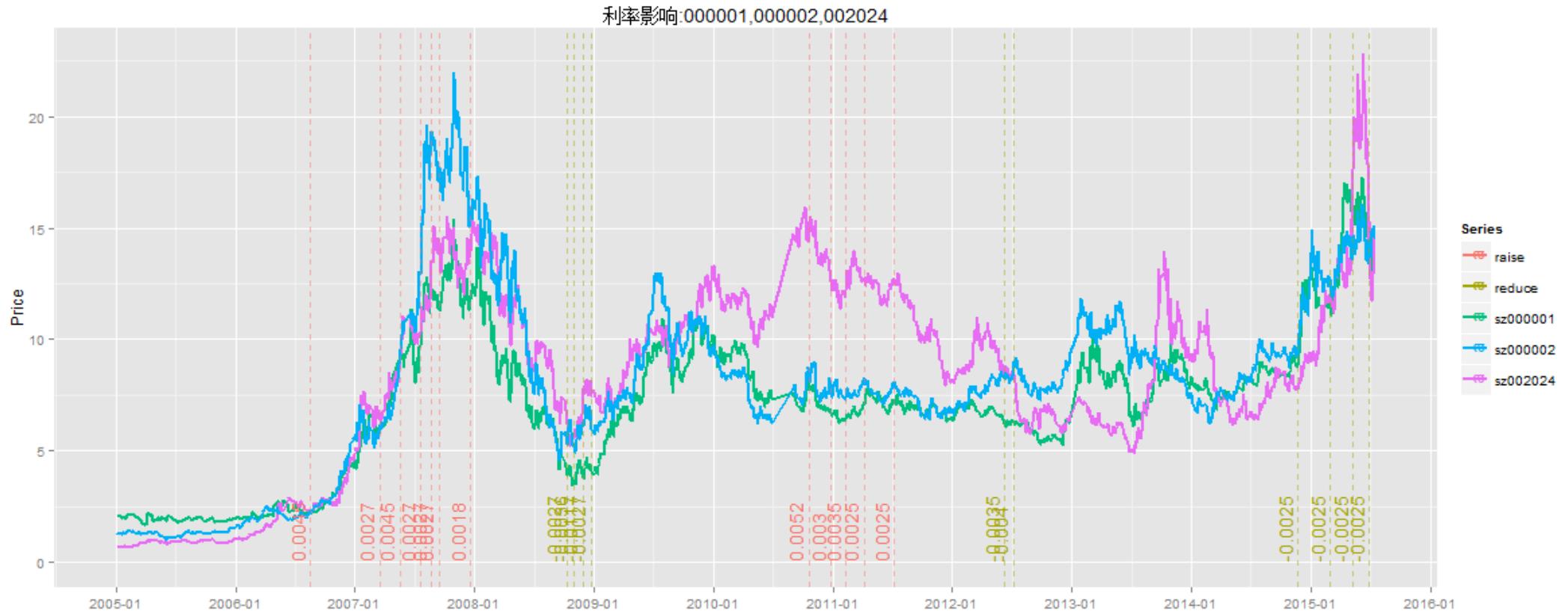
均值回归原理

- 在市场供需影响下，股票收益率不会偏离均值时间太久，股票价格会自然地向均值回归。
- 政府行为，会促进市场的**有效性**。当股价偏离均值后，并等于立即就会向均值回归，很可能会出现持续地均值回避。政府就会通过一些手段进行市场调节。
- 政府行为包括：**升准/降准、升息/降息、购买逆回购等**。在股票市场，地产股、银行股，受到调控政策影响会比较明显。



均值回归原理

- 对于升息/降息，平安银行，万科A，苏宁云商3支股票，在股市中都会有不一样的体现。



红色为升息的时间点和利率变动值，黄色为降息的时间点和利率变动值。



均值回归原理

WOT
2015 | World of Tech
WOT

51CTO 传媒

- 如何应用这个理论，找到投资的切入点呢？



均值回归原理 – 计算公式

- 从价值投资的角度，我们发现股价会在平均值上下波动，但如果考虑到资金的时间成本，把钱都压在股市中，等待几年的大行情，也是很不划算的。那么我们就需要对**均值（均衡价格水平）**进行重新定义。比如：以**20日均值**来代替长期平均值，找到短周期的一种投资方法。
- 计算原理：取日K线，以20日均值做为均值回归的**均值**，计算股价到均值的**差值**，求出差值的20日的标准差，从而判断差值的对于均值的偏离。
- 当偏离超过**2倍**标准差时，我们就认为股价**超涨或超跌**。股价会遵循均值回归的理论，不停地向均值进行修复。



均值回归原理

计算公式：

$$N\text{日平均值} = [T\text{日股价} + (T-1)\text{日股价} + \dots + (T-(N-1))\text{日股价}] / N$$

$$\text{差值} = N\text{日平均值} - T\text{日股价}$$

$$N\text{日差值均值} = [T\text{日差值} + (T-1)\text{日差值} + \dots + (T-(N-1))\text{日差值}] / N$$

$$N\text{日差值标准差} = \sqrt{[(T\text{日差值} - T\text{日差值均值})^2 + \dots + ((T-(N-1))\text{日差值} - (T-(N-1))\text{日差值均值})^2] / N}$$

如果N为20日，则

$$20\text{日平均值} = [T\text{日股价} + (T-1)\text{日股价} + \dots + (T-19)\text{日股价}] / 20$$

计算偏离点

$$T\text{日差值} > T\text{日差值标准差} * 2$$

- 以偏离点作为买入信号点，以均线和股价的下一个交点做为卖出信号点。这样我们就把均值回归的投资理论，变成了一个数学模型。



均值回归模型

WOT
2015 | World of Tech
WOT

51CTO 传媒

2. 均值回归模型和实现



均值回归模型：数据准备



- R语言本身提供了丰富的金融函数工具包，时间序列包zoo和xts，指标计算包TTR，数据处理包plyr，可视包ggplot2等，我们会一起使用这些工具包来完成建模、计算和可视化的工作。
- 关于zoo包和xts包的详细使用，可以参考文章
 - [R语言时间序列基础库zoo](#)
 - [可扩展的时间序列xts](#)
- 我本次用到的数据是从 **况客** 直接导出的，况客会逐步提供各种类型金融数据的API，让开发者可以免费下载。你也可以用quantmod包从Yahoo财经下载数据。



均值回归模型：数据准备

WOT
2015 | WOT
World of Tech

51CTO 传媒

- 本文用到的数据，包括A股日K线(向前复权)数据。
- 从2014年7月到2015年7月，以CSV格式保存到本地文件。

数据格式如下：

```
000001.SZ,2014-07-02,8.14,8.18,8.10,8.17,28604171
000002.SZ,2014-07-02,8.09,8.13,8.05,8.12,40633122
000004.SZ,2014-07-02,13.9,13.99,13.82,13.95,1081139
000005.SZ,2014-07-02,2.27,2.29,2.26,2.28,4157537
000006.SZ,2014-07-02,4.57,4.57,4.50,4.55,5137384
000010.SZ,2014-07-02,6.6,6.82,6.5,6.73,9909143
```

一共7列：

- 第1列，股票代码，code, 000001.SZ
- 第2列，交易日期，date, 2014-07-02
- 第3列，开盘价，Open, 8.14
- 第4列，最高价，High, 8.18
- 第5列，最低价，Low, 8.10
- 第6列，收盘价，Close, 8.17
- 第7列，交易量，Volume, 28604171



均值回归模型：数据准备

WOT
2015 | WOT
World of Tech

51CTO 传媒

- 通过R语言加载股票数据，生成XTS时间序列类型对象。

```
#加载工具包
> library(plyr)
> library(xts)
> library(TTR)
> library(ggplot2)
> library(scales)

# 读取CSV数据文件
> read<-function(file){
+   df<-read.table(file=file,header=FALSE,sep = ",",
+   na.s
+   names(df)<-c("code","date","Open","High","Low","Close"
+   dl<-split(df[-1],df$code)
+
+   lapply(dl,function(row){
+     xts(row[-1],order.by = as.Date(row$date))
+   })
+ }
```

```
# 加载数据
> data<-read("stock.csv")

# 查看数据类型
> class(data)
[1] "list"

# 查看数据的索引值
> head(names(data))
[1] "000001.SZ" "000002.SZ" "000004.SZ" "000005.SZ" "000006.SZ" "000007.SZ"

# 查看包括的股票数量
> length(data)
[1] 2782

# 查看股票000001.SZ
> head(data[['000001.SZ']])
      Open      High       Low     Close   Volume
2014-07-02 8.146949 8.180000 8.105636 8.171737 28604171
2014-07-03 8.171737 8.254364 8.122162 8.229576 44690486
2014-07-04 8.237838 8.270889 8.146949 8.188263 34231126
2014-07-07 8.188263 8.204788 8.097374 8.146949 34306164
2014-07-08 8.130424 8.204788 8.072586 8.204788 34608702
2014-07-09 8.196525 8.196525 7.915596 7.973434 58789114
```



均值回归模型：建模

- 为了能拉近我们对市场的了解，我们取从2015年1月1日开始的数据，来创建均值回归模型。
- 以平安银行(000001)为例，画出平安银行的2015年以来的日K线和均线。

```
# 计算移动平均线
> ma<-function(cdata,mas=c(5,20,60)){
+   if(nrow(cdata)<=max(mas)) return(NULL)
+   ldata<-cdata
+   for(m in mas){
+     ldata<-merge(ldata,SMA(cdata,m))
+   }
+   names(ldata)<-c('Value',paste('ma',mas,sep=''))
+   return(ldata)
+ }

# 日K线和均线
> title<-'000001.SZ'                                # 获得股票数据
> S2000011<-data[[title]]                          # 开始日期
> sDate<-as.Date("2015-01-01")                    # 结束日期
> eDate<-as.Date("2015-07-10")
> cdata<-S2000011[dateArea(sDate,eDate,360)]$Close # 获得收盘价
> ldata<-ma(cdata,c(5,20,60))                      # 选择移动平均指标

# 打印移动平均指标
> tail(ldata)
      Value    ma5    ma20    ma60
2015-07-03 13.07 13.768 15.2545 15.84355
2015-07-06 13.88 13.832 15.1335 15.82700
2015-07-07 14.65 13.854 15.0015 15.79850
2015-07-08 13.19 13.708 14.8120 15.74267
2015-07-09 14.26 13.810 14.6910 15.70867
2015-07-10 14.86 14.168 14.6100 15.67883
```



均值回归模型：建模

WOT
2015 | World of Tech
WOT

51CTO 传媒

- 我们设置3条移动平均线，分别是5日平均线，20日平均线，60日平均线，当然也可以按照自己的个性要求设置符合自己的周期。画出日K线和均线图。





均值回归模型：建模

- 计算不同周期的，股价与移动平均线差值的**标准差**。
- 以5日移动平均线做为均值时，会频繁进行交易，但每次收益都很小，可能都不够手续费的成本。
- 以60日移动平均线做为均值时，交易次数会较少，但可能会出现股票成形趋势性上涨或下跌，长时间不能收敛。
- 我们可以选择**20日**均线作为均值的标的。

```
# 5日平均线的差值、平均标准差
> ldata5<-getMaSd(ldata,5,sDate,eDate)
> head(ldata5)
  Value      ma5       dif       sd   rate
2015-01-05 13.23673 12.78724 -0.4494869 0.1613198 -2.79
2015-01-06 13.03842 12.89961 -0.1388121 0.1909328 -0.73
2015-01-07 12.79055 12.99215  0.2016081 0.3169068  0.64
2015-01-08 12.36089 12.90292  0.5420283 0.4472248  1.21
2015-01-09 12.46004 12.77733  0.3172848 0.3910700  0.81
2015-01-12 12.20390 12.57076  0.3668606 0.2533165  1.45
```

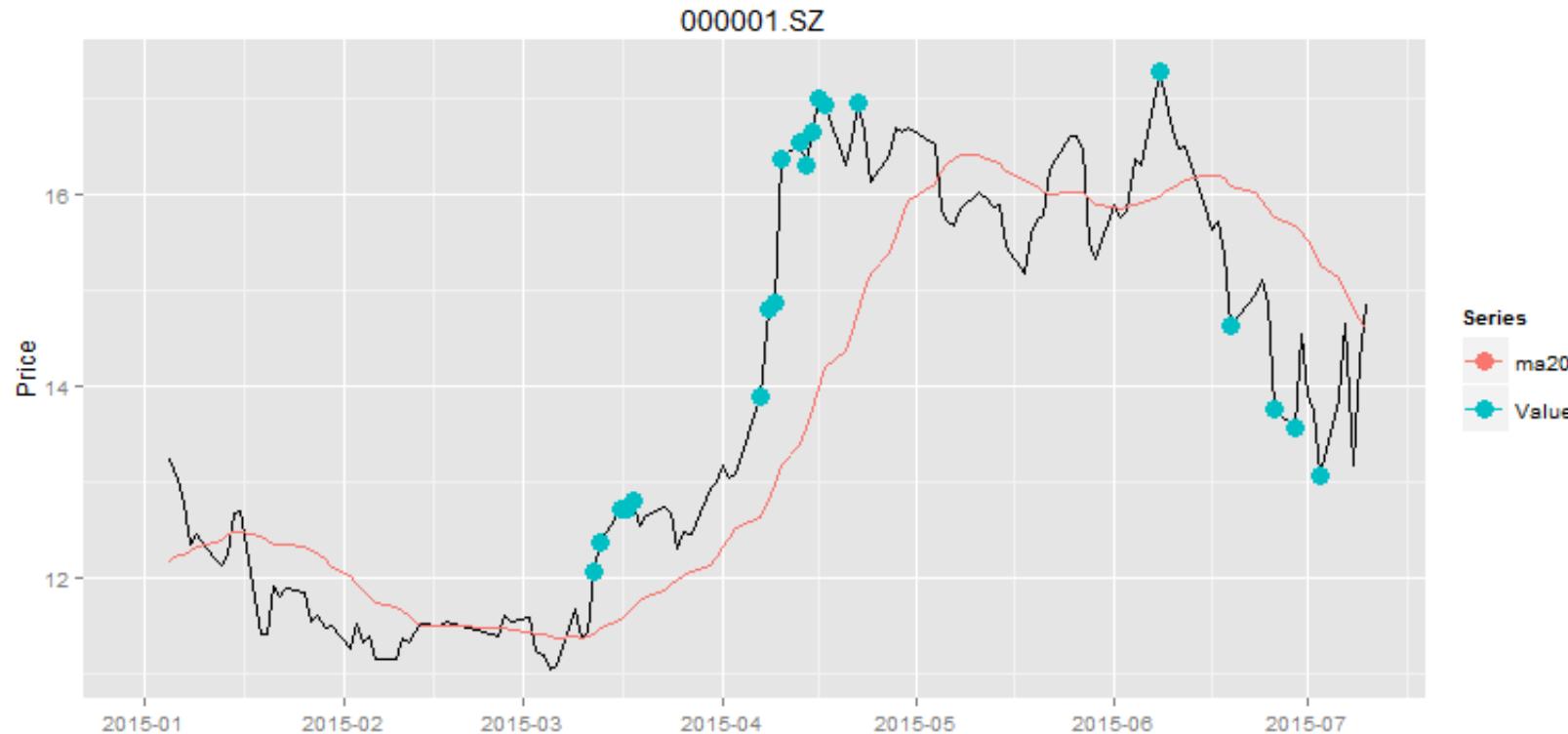
```
# 20日平均线的差值、平均标准差
> ldata20<-getMaSd(ldata,20,sDate,eDate)
> head(ldata20)
  Value      ma20       dif       sd   rate
2015-01-05 13.23673 12.18613 -1.05059293 0.6556366 -1.60
2015-01-06 13.03842 12.23778 -0.80064848 0.6021093 -1.33
2015-01-07 12.79055 12.24810 -0.54244141 0.4754686 -1.14
2015-01-08 12.36089 12.29975 -0.06114343 0.5130410 -0.12
2015-01-09 12.46004 12.33651 -0.12352626 0.5150453 -0.24
2015-01-12 12.20390 12.37163  0.16773131 0.5531618  0.30
```

```
# 60日平均线的差值、平均标准差
> ldata60<-getMaSd(ldata,60,sDate,eDate)
> head(ldata60)
  Value      ma60       dif       sd   rate
2015-01-05 13.23673 10.06939 -3.167340 1.264792 -2.50
2015-01-06 13.03842 10.14678 -2.891644 1.271689 -2.27
2015-01-07 12.79055 10.22087 -2.569677 1.269302 -2.02
2015-01-08 12.36089 10.28752 -2.073368 1.258813 -1.65
2015-01-09 12.46004 10.35527 -2.104766 1.247967 -1.69
2015-01-12 12.20390 10.41821 -1.785691 1.233989 -1.45
```



均值回归模型：建模

- 当差值超过2倍的平均标准差时，我们认为股价出现了偏离，以偏离点做为模型的买入信号。



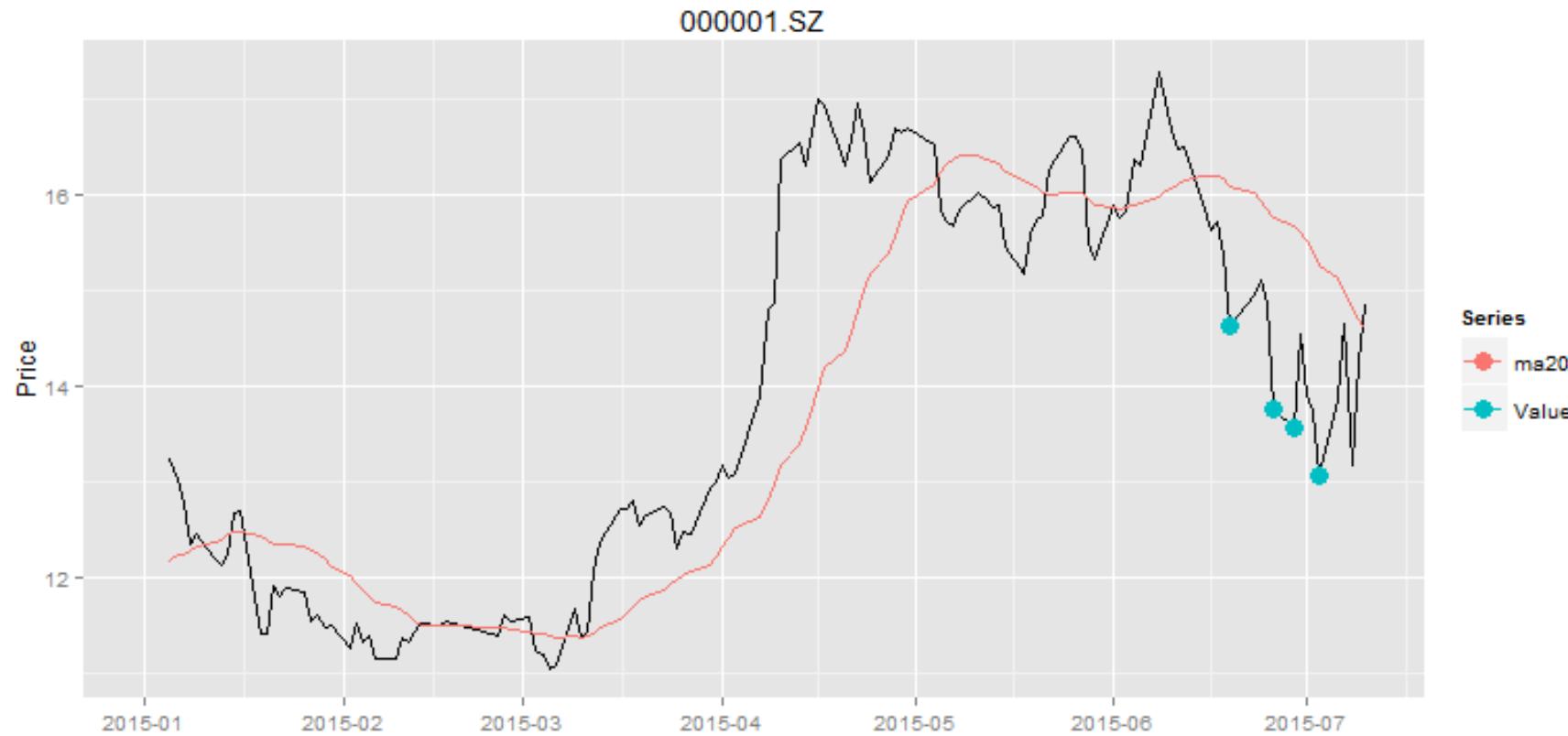


均值回归模型：建模

WOT
2015 | World of Tech
WOT

51CTO 传媒

- 由于股票只能进行单向交易，即**低买高卖**，并不能直接做空，只保留股价低于移动平均线的点，就是我们的买入信号点。



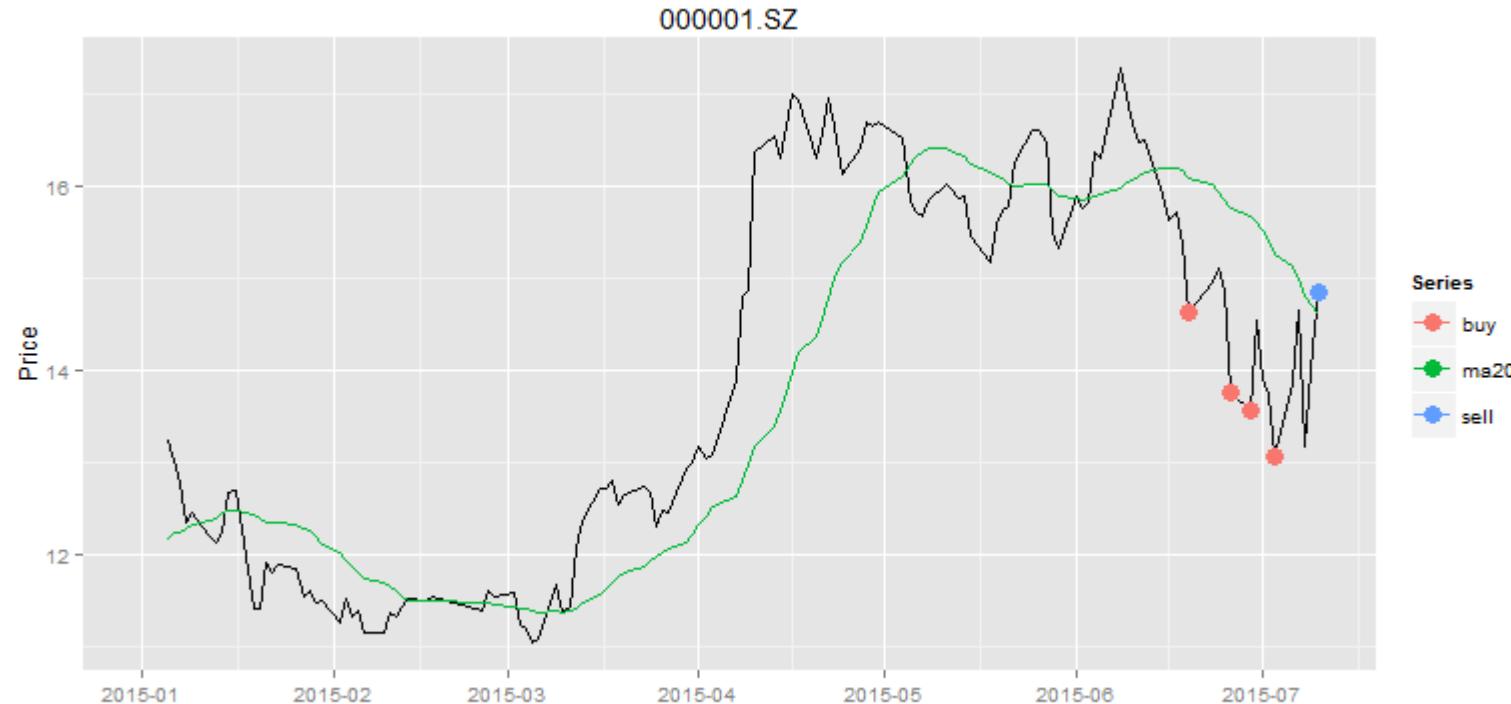


均值回归模型：建模

WOT
2015 | World of Tech
WOT

51CTO 传媒

- 计算卖出的信号点，当买入后，下一个股价与移动平均线的交点就是卖出的信号点。





均值回归模型：建模

- 那么究竟赚了多少钱呢？我们还需要精确的计算出来！
- **模拟交易**：设置交易参数和规则。
- 以10万元人民币为本金。
- 买入信号出现时，以收盘价买入，每次买入价值1万元的股票。如果连续出现买入信号，则一直买入。若现金不足1万元时，则跳过买入信号。
- 卖出信号出现时，以收盘价卖出，一次性平仓信号对应的股票。
- 手续费为0元。



均值回归模型：建模

- 一共发生了5笔交易，其中4笔买入，1笔卖出。
- 最后，资金剩余103275.88元，赚了3275.88元，收益率3.275%。

```
> result$ticks
  Value    ma20      dif       sd   rate op     cash amount     asset      diff
2015-06-19 14.63 16.0965 1.4665 0.6620157 2.22   B 90007.71      683 100000.00      0.00
2015-06-26 13.77 15.7720 2.0020 0.8271793 2.42   B 80010.69     1409 99412.62 -587.38
2015-06-29 13.56 15.6840 2.1240 0.9271735 2.29   B 70016.97     2146 99116.73 -295.89
2015-07-03 13.07 15.2545 2.1845 1.0434926 2.09   B 60018.42     2911 98065.19 -1051.54
2015-07-10 14.86 14.6100 -0.2500 0.7384824 -0.34   S 103275.88      0 103275.88  5210.69
```

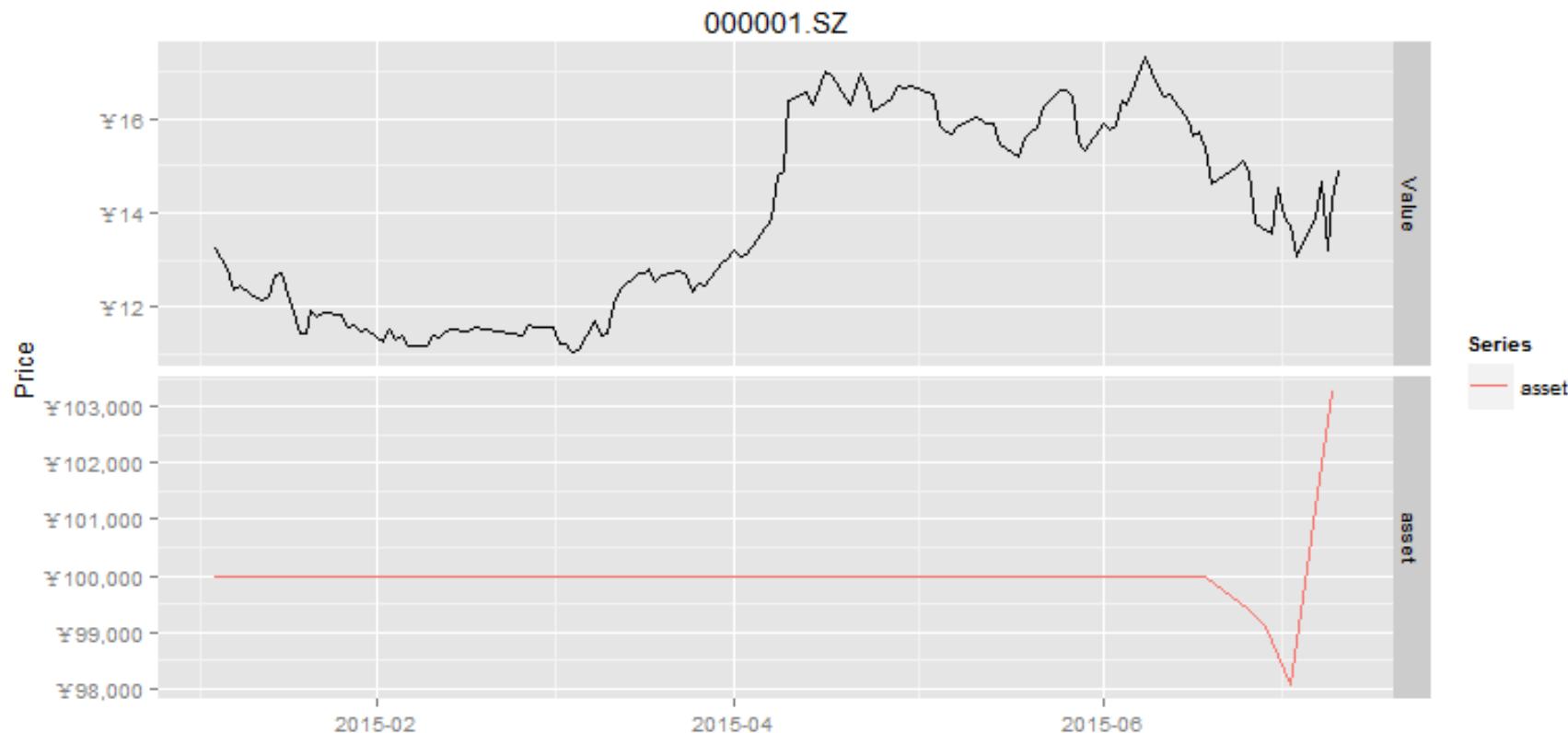


均值回归模型：建模

WOT
2015 | World of Tech
WOT

51CTO 传媒

- 股价和资产净值(asset列)





均值回归模型：建模

WOT
2015 | World of Tech
WOT

51CTO 传媒

- 刚才我们是对一支股票进行了测试，发现是有机会的，那么我再换另外一支股票，看一下是否用同样的效果呢？
- 我们把刚才数据操作的过程，封装到统一的quick函数，就可以快速验证均值回归在其他股票的表现情况了。

```
> quick<-function(title,sDate,eDate) {  
+   stock<-data[[title]]  
+   cdata<-stock[dateArea(sDate,eDate,360)]$Close  
+   ldata<-ma(cdata,c(20))  
+   ldata<-getMaSd(ldata,20,sDate,eDate)  
+   buydata<-buyPoint(ldata,2,1)  
+   selldata<-sellPoint(ldata,buydata)  
+   sdata<-signal(buydata,selldata)  
+   return(trade(sdata))  
+ }
```



均值回归模型：建模

WOT
2015 | World of Tech
WOT

51CTO 传媒

- 我们用乐视网(300104)试一下，看看有没有赚钱的机会！！

```
> title<-"300104.SZ"
> sDate<-as.Date("2015-01-01") #开始日期
> eDate<-as.Date("2015-07-10") #结束日期

> quick(title,sDate,eDate)
$ticks
      Value    ma20     dif      sd rate op    cash amount    asset    diff
2015-06-19 55.04 69.9095 14.8695 5.347756 2.78 B 90037.76    181 1000000.00    0.00
2015-06-23 54.30 68.8075 14.5075 5.477894 2.65 B 80046.56    365 99866.06 -133.94
2015-06-24 56.21 67.8735 11.6635 5.404922 2.16 B 70097.39    542 100563.21   697.15
2015-06-25 51.80 66.8775 15.0775 5.770806 2.61 B 60099.99    735 98172.99 -2390.22
2015-06-26 46.79 65.9830 19.1930 6.580622 2.92 B 50133.72    948 94490.64 -3682.35
2015-06-29 47.05 64.9445 17.8945 7.096230 2.52 B 40159.12   1160 94737.12   246.48
2015-07-07 47.86 58.8150 10.9550 5.401247 2.03 B 30204.24   1368 95676.72   939.60
2015-07-10 57.92 57.3520 -0.5680 5.625309 -0.10 S 109438.80        0 109438.80 13762.08
```

- 我们又赚到了。一共发生了8笔交易，其中7笔买入，1笔卖出。
- 最后，资金剩余109438.80元，赚了9438.80元，收益率9.43%。

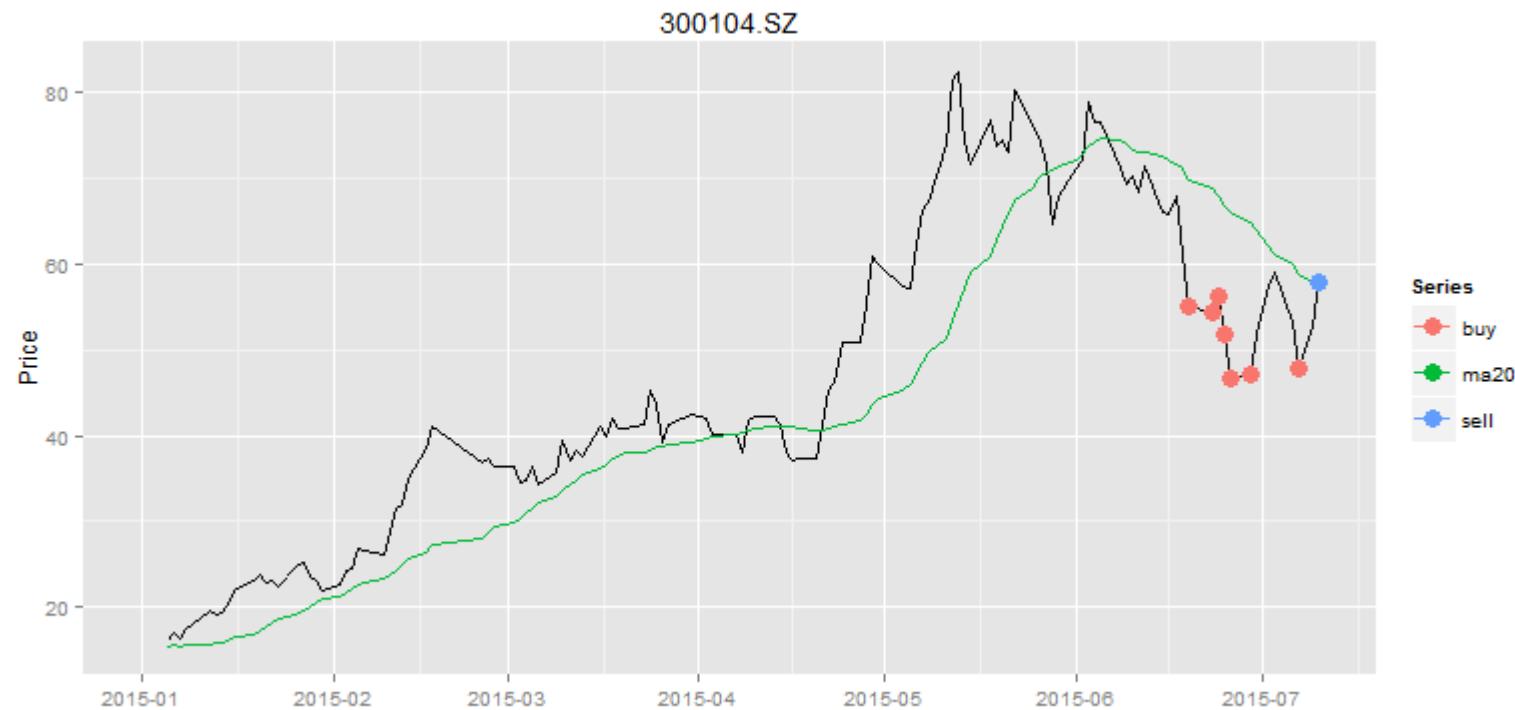


均值回归模型：建模

WOT
2015 | World of Tech
WOT

51CTO 传媒

- 乐视网(300104)的交易信号。



- 在恐慌的6月份，当别人都被套牢30%以上的情况下，我们还有9%正收益，那么应该是多么舒心的一件事情啊！！



量化选股

WOT
2015 | World of Tech
WOT

51CTO 传媒

3. 量化选股



量化选股

- 如果我们利用模型对**全市场的股票进行扫描**，会产生更多的交易信号，找到更多的投资机会，这样我们就能如何能获得更大的收益。
- 那么，接下来我们就根据均值回归的理论进行量化选股。
- 根据我们之前的经验，当股价与平均标准差的偏离越大，有可能带来的收益就越大。那么通过量化的手段，在整个的市场2700多支股票中，把**每天偏离最大股票**的找出来进行交易，就可以有效地分配我们的资金，进行更有效的投资。
- 我们要试一下，市场是否是和我们的思路是一致的？



量化选股

全市场共有**2782**支股票，
有**2697**支股票是符合条件的，
有**85**支股票由于数据样本不足被排除。

```
> sDate<-as.Date("2015-01-01")          # 开始日期
> eDate<-as.Date("2015-07-10")          # 结束日期

# 计算差值、平均值和平均标准差
> data0<-lapply(data,function(stock){})    # 代码省略

# 去掉空数据
> data0<-data0[!sapply(data0, is.null)]

# 全市场股票
> length(data)
[1] 2782

# 有效的股票
> length(data0)
[1] 2697

# 查看第1支股票
> head(data0[[1]])
  Value      ma20       dif       sd rate
2015-01-05 13.23673 12.18613 -1.05059293 0.6556366 -1.60
2015-01-06 13.03842 12.23778 -0.80064848 0.6021093 -1.33
2015-01-07 12.79055 12.24810 -0.54244141 0.4754686 -1.14
2015-01-08 12.36089 12.29975 -0.06114343 0.5130410 -0.12
2015-01-09 12.46004 12.33651 -0.12352626 0.5150453 -0.24
2015-01-12 12.20390 12.37163  0.16773131 0.5531618  0.30
```



量化选股

- 对2697支股票进行筛选，找到符合要求的买入信号点。有1819支股票，在这半年中产生过买入信号。
- 每支股票产生的买入信号的时间和频率都是不同，这样我们就可以把钱分散投资到不同的股票上，同时分散风险。
- 如果交易信号同一天出现在多支的股票上，而我们资金有限，又想让收益最大化，那么我们可以选择**偏离值最大的**股票进行交易。

```
# 查看有买入信号的股票
> length(buys)
[1] 1819

# 查看买入信号
> head(buys)
$`000001.SZ`
  Value    ma20     dif      sd rate
2015-06-19 14.63 16.0965 1.4665 0.6620157 2.22
2015-06-26 13.77 15.7720 2.0020 0.8271793 2.42
2015-06-29 13.56 15.6840 2.1240 0.9271735 2.29
2015-07-03 13.07 15.2545 2.1845 1.0434926 2.09

$`000002.SZ`
  Value    ma20     dif      sd rate
2015-03-05 11.90 12.568 0.668 0.2644101 2.53
2015-03-06 11.94 12.509 0.569 0.2674732 2.13

$`000004.SZ`
  Value    ma20     dif      sd rate
2015-01-05 15.69 17.7210 2.0310 0.7395717 2.75
2015-07-06 26.03 39.1540 13.1240 6.3898795 2.05
2015-07-07 23.43 38.2025 14.7725 6.9421723 2.13
2015-07-08 22.22 37.2635 15.0435 7.4287088 2.03

$`000005.SZ`
  Value    ma20     dif      sd rate
2015-07-06  6.02 10.9600 4.9400 2.381665 2.07
2015-07-07  5.42 10.5655 5.1455 2.333008 2.21

$`000006.SZ`
  Value    ma20     dif      sd rate
2015-01-19 5.829283 6.519462 0.6901792 0.26929 2.56
```



量化选股

- 每天选一支进行交易，我们选出81个买入信号点。

```
# 查看买入信号
> nrow(buydatas)
[1] 81

# 查看买入信号细节
> head(buydatas)
  .id      date    Value     ma20      dif      sd rate
1 002551.SZ 2015-01-05 16.573846 19.565446 2.9916000 0.74591596 4.01
2 002450.SZ 2015-01-06 18.548809 19.766636 1.2178275 0.34008453 3.58
3 300143.SZ 2015-01-07 11.480000 12.603000 1.1230000 0.32028018 3.51
4 300335.SZ 2015-01-08 12.113677 13.139601 1.0259238 0.21760484 4.71
5 300335.SZ 2015-01-09 12.243288 13.043888
6 300335.SZ 2015-01-12 11.994036 12.941694 # 卖出信号去重
```

```
> selldatas<-unique(selldatas)
> nrow(selldatas)
[1] 33
```

```
# 查看卖出信号
> head(selldatas)
```

	Value	ma20	dif	sd	rate	.id	date	op
2015-01-12	19.232308	18.848908	-0.38340000	0.9051374	-0.42	002551.SZ	2015-01-12	S
2015-01-08	19.814257	19.729006	-0.08525126	0.3782955	-0.23	002450.SZ	2015-01-08	S
2015-01-28	11.210000	11.019500	-0.19050000	0.7781848	-0.24	300143.SZ	2015-01-28	S
2015-01-21	13.190448	12.899321	-0.29112706	0.3871871	-0.75	300335.SZ	2015-01-21	S
2015-01-213	7.140000	6.989500	-0.15050000	0.2007652	-0.75	002505.SZ	2015-01-21	S
2015-01-22	5.561561	5.490668	-0.07089242	0.2127939	-0.33	600077.SH	2015-01-22	S

有了买入信号，继续找到卖出信号。



量化选股

- 合并买入信号和卖出信号，并计算收益。

```
# 查看合并的信号
> head(sdatas)
  .id      date Value    ma20      dif      sd rate op
36  000002.SZ 2015-03-05 11.90 12.56800  0.668000 0.26441011  2.53  B
100 000002.SZ 2015-03-16 12.49 12.38050 -0.109500 0.23702768 -0.46  S
58  000553.SZ 2015-05-06 14.35 15.50882  1.158824 0.38429912  3.02  B
110 000553.SZ 2015-05-21 16.57 15.18903 -1.380972 0.55647152 -2.48  S
26  000725.SZ 2015-02-09  2.80  3.11400  0.314000 0.07934585  3.96  B
94  000725.SZ 2015-02-16  3.09  3.06500 -0.025000 0.08182388 -0.31  S
```

```
# 查看信号的股票
> names(results)
[1] "000002.SZ" "000553.SZ" "000725.SZ" "000786.SZ" "000826.SZ" "002240.SZ" "002450.SZ"
[8] "002496.SZ" "002505.SZ" "002544.SZ" "002551.SZ" "002646.SZ" "002652.SZ" "300143.SZ"
[15] "300335.SZ" "300359.SZ" "300380.SZ" "300397.SZ" "300439.SZ" "300440.SZ" "300444.SZ"
[22] "600030.SH" "600038.SH" "600077.SH" "600168.SH" "600199.SH" "600213.SH" "600375.SH"
[29] "600490.SH" "600536.SH" "600656.SH" "600733.SH" "600890.SH" "601179.SH" "601186.SH"
[36] "601628.SH" "601633.SH" "601939.SH" "603019.SH"
```



量化选股

- 我们查看万科A(000002)的股票。
- 通过偏离值最大的规则，一共有2笔交易，赚了**495元**。

```
> results[['000002.SZ']]$ticks
  date Value    ma20      dif       sd   rate op    cash amount    asset diff
36 2015-03-05 11.90 12.5680  0.6680 0.2644101 2.53  B 90004.0     840 100000.0   0.0
100 2015-03-16 12.49 12.3805 -0.1095 0.2370277 -0.46  S 100495.6      0 100495.6 495.6
```

- 如要我们没有进行算法优化，只交易万科A，那么会发生3笔交易，我们可以赚**955.95元**。

```
> quick('000002.SZ',sDate,eDate)$ticks
  Value    ma20      dif       sd   rate op    cash amount    asset diff
2015-03-05 11.90 12.5680  0.6680 0.2644101 2.53  B 90004.00     840 100000.0  0.00
2015-03-06 11.94 12.5090  0.5690 0.2674732 2.13  B 80010.22    1677 100033.6 33.60
2015-03-16 12.49 12.3805 -0.1095 0.2370277 -0.46  S 100955.95      0 100955.9 922.35
```



量化选股 – 模型优化



- 如果从交易的角度讲，上面的模型还不能算完成。
- 对模型参数的优化：
 - 用10日均线代替20日均线，
 - 用3倍标准差偏移代替2倍标准差偏移
 - 对样本进行正态分布的检验
 - 结合其他趋势类模型共同产生信号
 -
- 大家可以 [况客](#) 的网站上，继续跟踪这个策略的发展，也会发现更多不一样的策略。



总结

- 本文从 **均值回归**的理论的介绍开始，到**市场特征**检验，再到**数学公式**，**R语言建模**，**历史数据回测**，最后找到**投资机会**，是一套完整的从理论到实践的学习方法。虽然困难重重，但做为有理想的极客，我们是有能力来克服这些困难的。
- 本文同时用到了计算机、金融、数学、统计等多学科知识的结合，我认为这是技术复合人才未来的发展方向。如果说过去10年是房地产的黄金10年，那么未来的10年将是金融的黄金10年。当我们IT人掌握了足够的金融知识，一定会有能力去金融市场抢钱的。
- **程序员，加油！！**



关于作者

WOT
2015 | World of Tech
WOT

51CTO 传媒

- 张丹, [况客科技\(北京\)有限公司](#), 创始人/CTO
 - Blog: <http://blog.fens.me>
 - 视频: <http://onbook.me>
 - Email: bsspirit@gmail.com
-
- 《R的极客理想》系列图书作者





关于况客



况客
Qutke.com

建立**理财**投资者与**策略**提供者之间的桥梁！



扫二维码添加况客微信

高薪诚聘：

- 产品经理
- 金融量化研究员
- 高级测试工程师

投递简历至：**hr@qutke.com**



Thank You!