

课程回顾

- 第 1 课 金融数据分析和量化投资、时间序列、统计学的基本概念
- 第 2 课 金融时间序列的基本性质 均值、方差、自相关性、平稳性、随机性
- 第 3 课 平稳时间序列模型 AR、MA 及 ARMA 模型简介
- 第 4 课 确定趋势建模 通过传统回归方法估计非常数均值趋势模型的参数
- 第 5 课 季节模型 针对具有一定循环或周期性的数据
- 第 6 课 非平稳时间序列分析 通过差分平稳化构建的 ARIMA 模型
- 第 7 课 应用实例与模型比较 通过实际的金融数据分析实例来熟悉各个，并比较各个模型的优劣
- 第 8 课 资产收益波动率并非是一个常数——条件异方差模型及应用 讨论用来描述资产收益率的波动率随时间而改变的各种经济计量模型
- 第 9 课 金融时间序列的非线性模型及其应用 介绍非线性模型的检验与各种非线性时间序列的数学模型及其在金融中的应用
- 第 10 课 多元时间序列分析 介绍简单的多元模型和协整的相关知识
- 第 11 课 多元时间序列分析的简化与降维 主成分分析与因子模型
- 第 12 课 动态数据的状态空间模型和 Kalman 滤波的简介
- 第 13 课 马尔可夫链特卡罗方法

大作业

业务背景

今年初至 6 月上旬，伴随指数的快速上涨和 A 股开户数猛增，融资盘这类新生事物也首次大举入市，多数 A 股投资者收获颇丰，然而最近以来风云突变，持续的巨幅调整，给了新老投资者一记闷棍。这是 A 股市场头一遭经历高杠杆下的持续下跌。面对全新的局面，下半年应持怎样的操作思路？我们需要通过客观的数据来观察大盘，能够对整体的时间序列进行深入的分析。

数据来源

使用 quantmod 包抓取上证指数来进行数据分析

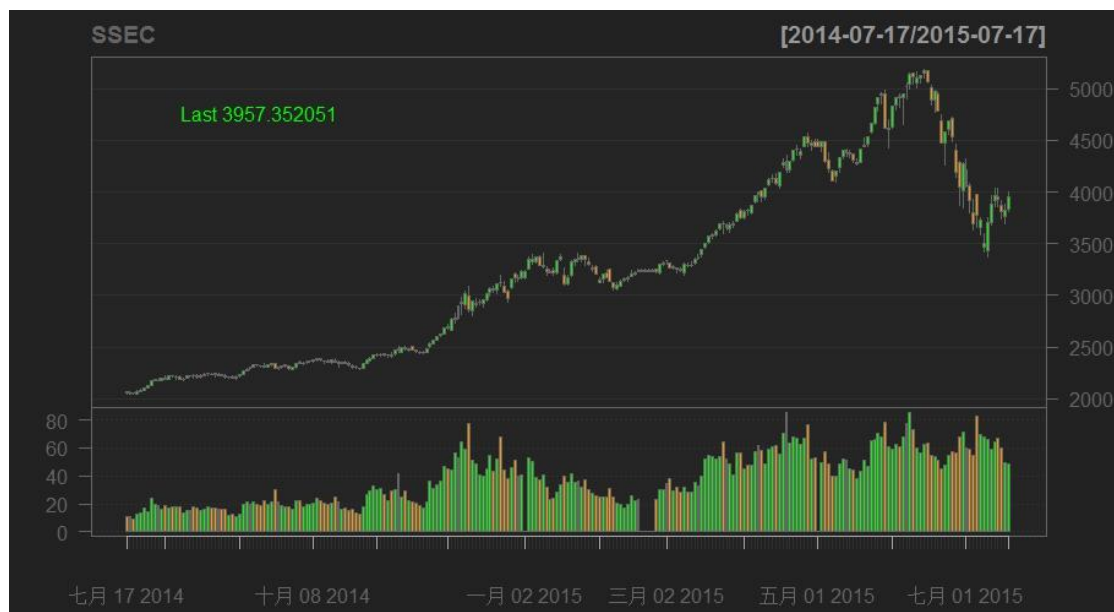
```
library(quantmod)
```

```
getSymbols("^SSEC")
```

```
> head(SSEC)
      SSEC.Open SSEC.High SSEC.Low SSEC.Close SSEC.Volume SSEC.Adjusted
2007-01-04  2728.188  2847.615 2684.818   2715.719     120200    2715.719
2007-01-05  2668.577  2685.804 2617.019   2641.334     106200    2641.334
2007-01-08  2621.068  2708.444 2620.625   2707.199     106800    2707.199
2007-01-09  2711.049  2809.394 2691.360   2807.804     110800    2807.804
2007-01-10  2838.113  2841.741 2770.988   2825.576     111800    2825.576
2007-01-11  2819.367  2841.180 2763.886   2770.110     121600    2770.110
> tail(SSEC)
      SSEC.Open SSEC.High SSEC.Low SSEC.Close SSEC.Volume SSEC.Adjusted
2015-07-13  3918.990  4030.195 3858.637   3970.388     643500    3970.388
2015-07-14  3958.373  4035.435 3855.560   3924.487     670600    3924.487
2015-07-15  3874.968  3914.273 3741.250   3805.703     601300    3805.703
2015-07-16  3758.505  3877.514 3688.442   3823.176     492300    3823.176
2015-07-17  3831.421  3994.477 3814.147   3957.352     481700    3957.352
2015-07-20  3948.421  4021.326 3927.121   3992.110     539100    3992.110
```

选择最近一年数据进行分析，给出蜡烛图

```
chartSeries(SSEC,subset="2014-07-17::2015-07-17")
```



可见，大盘指数在 2015 年 6 月份大幅度跳水，而从 7 月份有开始回升的迹象。

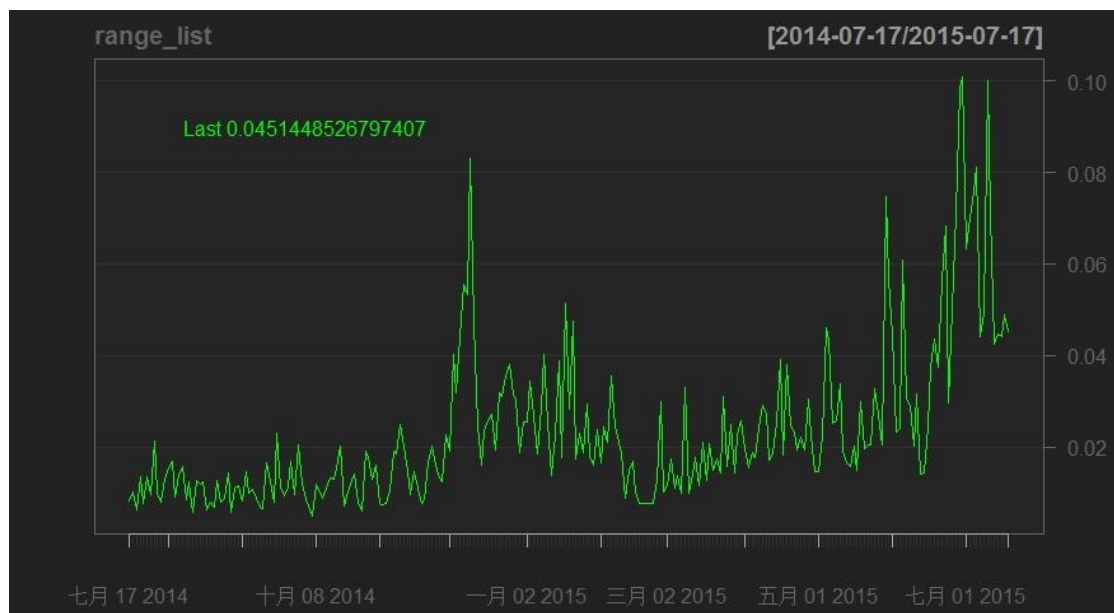
数据建模

我们选择最小值/最大值的指标来衡量波动的稳定性，如下：

1 - 最低价/最高价。这个指标变相的衡量了每天的理想化振动收益。

```
range_list = 1 - Lo(SSEC)/Hi(SSEC)
```

```
chartSeries(range_list,subset="2014-07-17::2015-07-17")
```



summary 一下：

```
> summary(range_list)
      Index      SSEC.Low
Min.   :2007-01-04  Min.   :0.00000
1st Qu.:2009-02-25  1st Qu.:0.01158
Median :2011-04-18  Median :0.01670
Mean    :2011-04-15  Mean    :0.02069
3rd Qu.:2013-06-06  3rd Qu.:0.02556
Max.    :2015-07-20  Max.    :0.10096
```

可见，最大值在 10.096%，而中位数为 1.670%，均值为 2.069%，多空双方厮杀程度相当激烈。

使用另外一个指标衡量真实的变化波动性：

(收盘价-开盘价) / 最高价

$\text{ratio_list} = (\text{Cl}(\text{SSEC}) - \text{Op}(\text{SSEC})) / \text{Hi}(\text{SSEC})$

```
> summary(ratio_list)
      Index      SSEC.Close
Min.   :2007-01-04  Min.   : -0.090839
1st Qu.:2009-02-25  1st Qu.: -0.006193
Median :2011-04-18  Median :  0.001468
Mean    :2011-04-15  Mean    :  0.001095
3rd Qu.:2013-06-06  3rd Qu.:  0.009477
Max.    :2015-07-20  Max.    :  0.075609
```

通过肉眼观察，这个时间序列的稳定性会更高一些，没有使用方差计算。

$\text{ratio_list.numeric} = \text{as.numeric}(\text{ratio_list}\$ \text{SSEC.Close})$

$\text{ar}(\text{ratio_list.numeric}, \text{method} = ' \text{mle} ') \$ \text{order}$ #base on AIC 准则

针对 AR 自回归模型获取定阶为 1。

$\text{library}(\text{fUnitRoots})$

$\text{adfTest}(\text{ratio_list.numeric}, \text{lags} = 1, \text{type} = " \text{c} ")$

使用 ADFTest 做单位根检验：

```
Title:
Augmented Dickey-Fuller Test

Test Results:
PARAMETER:
Lag Order: 1
STATISTIC:
Dickey-Fuller: -33.7784
P VALUE:
0.01

Description:
Tue Jul 21 17:36:38 2015 by user: Administrator

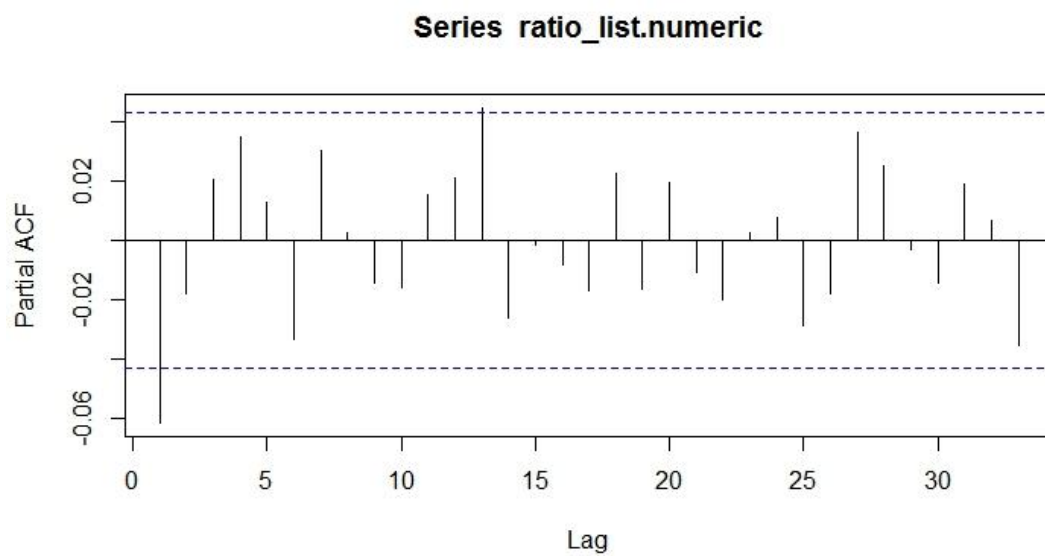
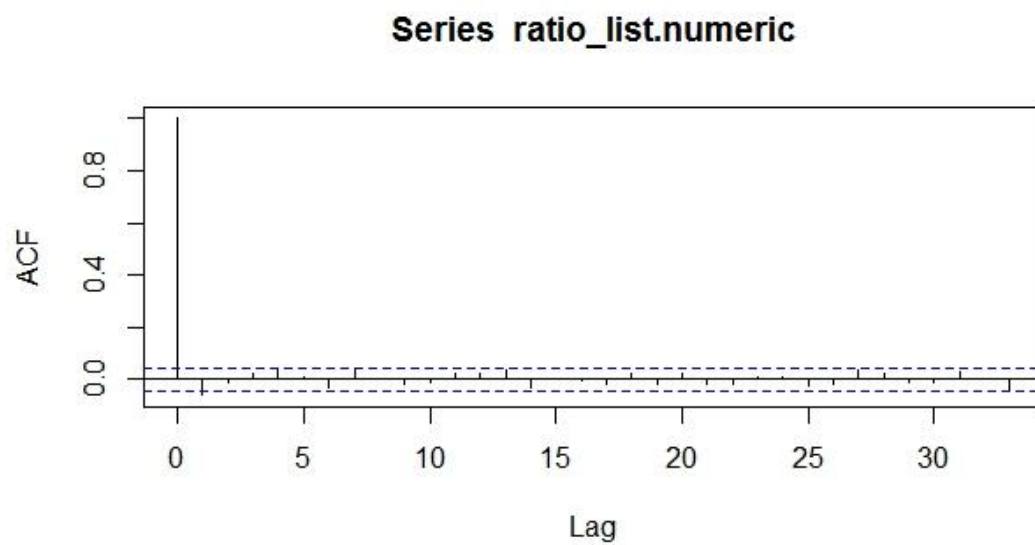
Warning message:
In adfTest(ratio_list.numeric, lags = 1, type = "c") :
  p-value smaller than printed p-value
```

p-value=0.01，很小，可以拒绝 H_0 假设，证明是个平稳的时间序列

$\text{acf}(\text{ratio_list.numeric})$

$\text{pacf}(\text{ratio_list.numeric})$

观察 ACF 图和 PACF 图：



不需差分，是个平稳的时间序列

```
library(TSA)
```

```
eacf(ratio_list.numeric)
```

```
> eacf(ratio_list.numeric)
AR/MA
  0 1 2 3 4 5 6 7 8 9 10 11 12 13
0 x o o o o o o o o o o o o o
1 x o o o o o o o o o o o o o
2 x x o o o o o o o o o o o o
3 x x x o o o o o o o o o o o
4 x x x x o o o o o o o o o o
5 x x x x x o o o o o o o o o
6 x x x x x x o o o o o o o o
7 x x x x o x o o o o o o o o
```

我们根据上述表选择：ARIMA(0,0,1)

```
m2=arima(ratio_list.numeric,order=c(0,0,1))
```

```
m2
```

```
> m2

Call:
arima(x = ratio_list.numeric, order = c(0, 0, 1))

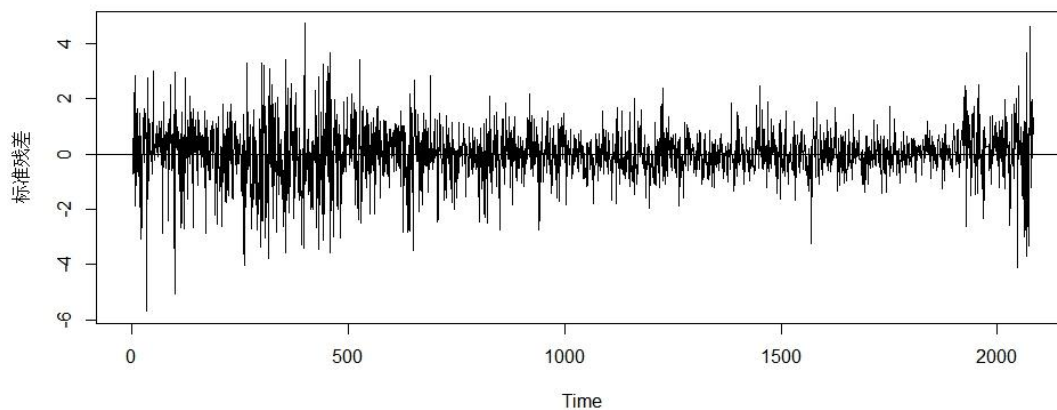
Coefficients:
          ma1      intercept
      -0.0632       0.0011
s.e.   0.0221       0.0003

sigma^2 estimated as 0.0002566:  log likelihood = 5652.81,  aic = -11301.62
```

检查残差是否为白噪声:

```
plot(rstandard(m2),ylab='标准残差',type='l')
```

```
abline(h=0)
```



结论解释

通过上面的分析，我们可以初步判断：未来几天大盘盘中的振动性将会收窄。市场情绪有所缓解，多方短期占优。