时间序列分析报告

——对近十年上证综指交易数据的分析

代一

June 21, 2017

班级： 基科 5 2

院系： 理学院物理系

学号： 2015012204

目录

I 简介 **2**

II 数据分析 **3**

1 数据来源及简介 **3**

2 初步处理 **3**

3 平稳性检验及差分 **4**

4 ARMA(p,q) Model **6**

5 ARCH Model **8**
- 5.1 建立均值方程 9
- 5.2 ARCH 效应的检验 10
- 5.3 ARCH 模型的建立 12
- 5.4 ARCH 模型的检验和预测 14

6 GARCH Model **18**
- 6.1 GARCH 模型的建立 18
- 6.2 GARCH 模型的预测和检验 20

7 总结 **22**

III 附录 **24**

8 参考文献 **24**

9 补充说明 **24**

I 简介

摘要：本文主要利用 python 的统计和时间序列分析工具，对上证综指近十年的日交易数据进行了简要分析。通过对 ARMA 模型，ARCH 模型，GARCH 模型的尝试和比较，最终选择使用 GARCH 模型对波动率进行预测。得出的结论是金融市场具有较大的波动性，同时体现出波动性聚集的特点，使用异方差模型能够对此进行较好的刻画和预测，实际使用中对交易标的的选择具有重要的意义。

关键词：波动率 ARMA ARCH GARCH

II 数据分析

1 数据来源及简介

我们的数据来自于清华大学图书馆所购买的国泰安金融数据库，为上证综合指数近十年的日交易数据。开始日期为 2007-6-11，截止日期为 2017-5-23，股指代号 000001，数据包括每个交易日的开盘价，收盘价，以及涨跌幅。

上证综合指数是我国股票交易所最早发布的指数，它涵盖了上海证券交易所挂牌上市的全部的股票，是以发行量为权重的加权综合股价指数，具有较大的代表性，能够科学的表征上海证券市场乃至全国市场的波动情况。同时由于股票指数数据较易获得，数据也具有较高程度的规范性，因此我们最终选择上证综合指数作为我们分析的数据。

具体获得方法为：

- 由清华内网进入国泰安金融数据库 (http://cn.gtadata.com/)
- 进入数据中心-> 股票市场系列-> 市场指数
- 左侧选择交易数据-> 指数日行情文件
- 右侧依照上述内容设置相应的指数，日期和需要的数据
- 提交查询，并下载所得数据

2 初步处理

我们选择上证综指的每个交易日的开盘价 (Open) 进行分析，首先对其时间序列图像进行绘制，进行初步的观察的分析：

Figure.1 Open-Time

我们可以初步目测其并不平稳。同时发现图像在一些区间中存在着明显的趋势，这种趋势反映出市场行情的变化，但整体而言，每一段短时间内都没有大的波动。

3 平稳性检验及差分

我们选择检验平稳性的方法为 D-F 单位根检验，这是我们常用的一种平稳性检验方法，原假设 H0：序列为非平稳的，备择假设 H1: 序列是平稳的。检验结果为：

```
Results of Dickey-Fuller Test:
Test Statistic                   -2.377830
p-value                           0.148092
#Lags Used                       13.000000
Number of Observations Used    2409.000000
Critical Value (5%)              -2.862741
```

```
Critical Value (1%)              -3.433067
Critical Value (10%)             -2.567409
```

我们发现 p-value 较大，同时 Test Statistic 大于 Critical Value (10%)，因此我们不能推翻原假设，原假设成立。序列为非平稳的。

我们需要进行差分来使得序列平稳，进行一阶差分，得到一阶差分序列 (diff1)，其图像为：

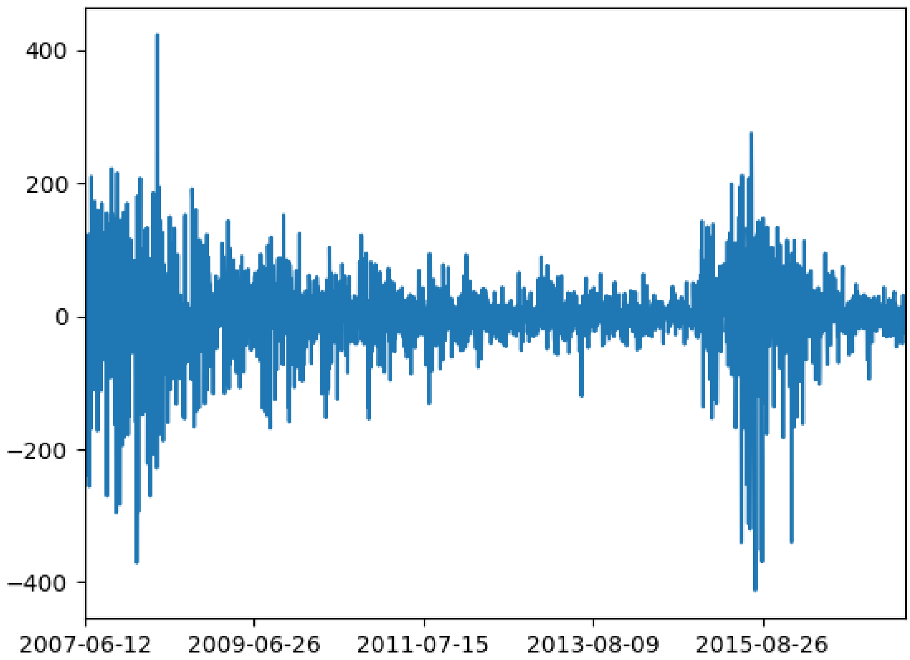

Figure.2 diff1-Time

从图像中可以看出序列基本已经趋向平稳，趋势已经消失。

同样进行单位根检验，结果为：

```
Results of Dickey-Fuller Test:
Test Statistic                  -22.553096
p-value                           0.000000
#Lags Used                        3.000000
Number of Observations Used    2418.000000
Critical Value (5%)              -2.862736
```

```
Critical Value (1%)            -3.433057
Critical Value (10%)           -2.567407
```

我们发现 p-value 极小，已经小于计算机精度，自然小于显著性水平，同时 Test Statistic 小于任何一个 Critical Value， 原假设不成立，选择备择假设，即序列平稳，置信概率大于 99%。注意到检验的统计量远远小于 Critieria，这很大一部分原因是我们使用的数据较多，使得整体更加平稳。

4 ARMA(p,q) Model

我们首先考虑 ARMA 模型，一方面，序列已经稳定，符合 ARMA 模型稳定的要求，另一方面，ARMA 模型较为基础，也有助于其他可能模型的分析。

首先对 ACF 和 PACF 进行计算和绘制，来初步判断 ARMA 模型的阶数，图像为：

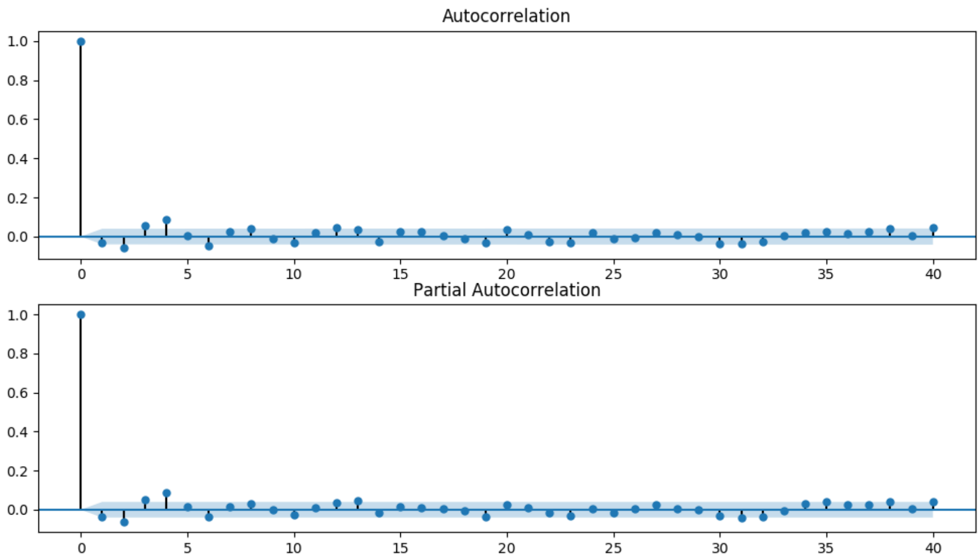

Figure.3 ACF and PACF

从图像中我们观察到，ACF 和 PACF 均没有发生拖尾，但截尾阶数较高，为了防止过拟合的情况，我们选择较为小的截尾阶数。我们发现 ACF 和 PACF 均在 4 阶处截尾。

事实上，ACF 和 PACF 在这里的表现均不算良好，性质与白噪声十分相似。

当然，为了进一步准确进行判断，我们需要 AIC 和 BIC 信息法来对模型进行选择，不妨即将 p,q 上限定为 4，对 AIC，结果为：

```
Out[17]:
{'aic':            0              1              2              3              4
 0  26779.415840  26778.297906  26773.937331  26766.416437  26754.359347
 1  26778.657218  26778.927365  26773.795437  26759.866172  26756.338218
 2  26771.826761  26772.179725  26742.724990  26744.388607  26746.336695
 3  26767.874637  26759.110787  26744.385594  26741.080688  26743.442346
 4  26752.085671  26753.681547  26746.289217  26742.922861  26744.684760,
 'aic_min_order': (3, 3)}
```

Figure.4 AIC

对 BIC，结果为：

```
Out[18]:
{'bic':            0              1              2              3              4
 0  26791.000538  26795.674953  26797.106727  26795.378181  26789.113441
 1  26796.034265  26802.096761  26802.757181  26794.620265  26796.884661
 2  26794.996157  26801.141470  26777.479083  26784.935050  26792.675486
 3  26796.836381  26793.864881  26784.932037  26787.419480  26795.573486
 4  26786.839765  26794.227989  26792.628009  26795.054001  26802.608249,
 'bic_min_order': (2, 2)}
```

Figure.5 BIC

我们发现两者的判断结果并不相同，我们最终选择 BIC 判断的结果，因为 BIC 有着较高的阶数惩罚，而 AIC 则较低，容易过估计。因此我们选择 ARMA(2,2) 模型进行拟合。

我们使用 python 中模块函数 statsmodels.tsa.arima_models.ARMA 进行拟合，在运行过程中遇到问题：

python 报错为：

```
ValueError: The computed initial AR coefficients are not stationary
You should induce stationarity, choose a different model order, or you can
pass your own start_params.
```

经过检查和分析，发现原因有以下两点：

- diff1 时间序列的值过大，使得波动的绝对值较大，函数运行时发生错误。
- ARMA(2,2) 模型仍然不能很好的描述该序列，序列存在着较大的残差，同时可能存在着异方差使得拟合所得结果难以平稳进行。

综上，我们放弃 ARIMA 模型，该模型的拟合结果可能并不准确，也较难以拟合。我们根据 diff1 图像中也能看出波动率聚集的现象，ARCH 模型也许是更适用的。我们选择 ARCH 模型作为下一个试验模型。考虑到 diff1 的值均较大，

我们选择另一个与其性质基本等效的常用的金融量——涨跌幅 (Ratio) 作为我们的分析对象。涨跌幅实质上上也是一阶差分的结果，同时往往更受到人们的关心，更加重要。

5 ARCH Model

在传统计量经济学模型中，干扰项的方差被假设为常数。但是许多经济时间序列呈现出波动的集聚性，在这种情况下假设方差为常数是不恰当的。

ARCH 模型将当前一切可利用信息作为条件，并采用某种自回归形式来刻划方差的变异，对于一个时间序列而言，在不同时刻可利用的信息不同，而相应的条件方差也不同，利用 ARCH 模型，可以刻划出随时间而变异的条件方差。

我们首先刻画出涨跌幅随时间的变化：

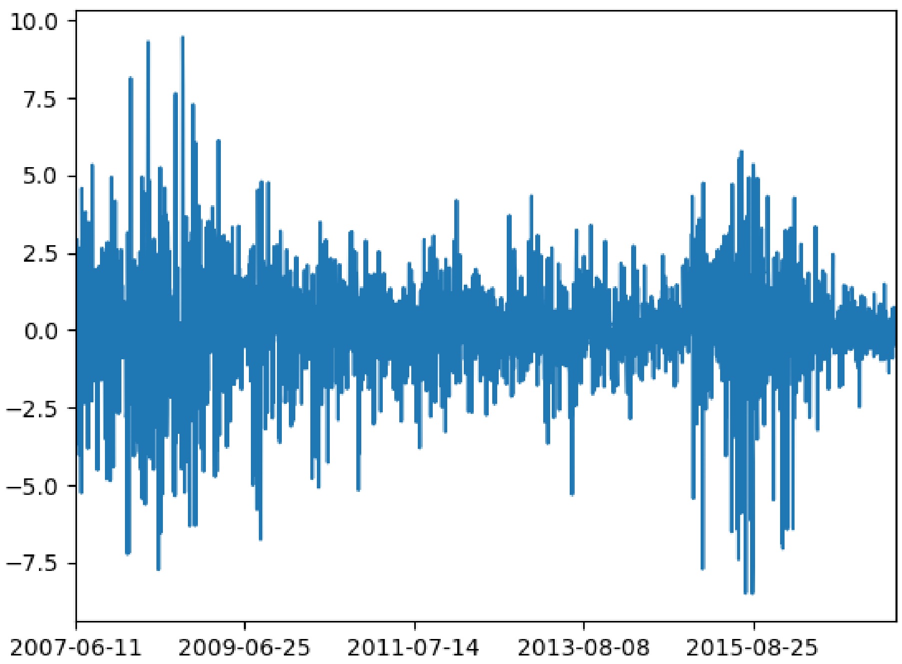

Figure.6 Ratio-Time

波动率的聚集性从 Ratio 的图像中也较为明显，两端的波动显然较大，因此 ARCH 模型很可能适用于涨跌幅 Ratio。

ARCH 模型建立的具体步骤为：

1. 通过检验数据的序列相关性建立一个均值方程，如有必要，对收益率序列建立一个模型（如 ARMA 模型）来消除线形依赖
2. 对均值方程的残差进行 ARCH 效应检验
3. 如果具有 ARCH 效应，则建立波动率模型
4. 检验拟合的模型，并对其进行可能的改进

5.1 建立均值方程

建立均值方程，这里即建立一个简单的 ARMA 模型或者 ARIMA 模型。我们这里建立 AR 模型作为我们的均值方程，原因有如下几点：

- AR 模型由于其自相关的性质，使得拟合时计算量较小，较为方便；
- AR 模型作为均值方程，与 ARMA 和 ARIMA 相比并没有明显的弊端，因为 ARCH 模型的异方差中的方差的拟合可以对此进行一部分补偿和修正；
- python 中 arch 库的均值方程仅仅支持常数和 AR 模型。

类似之前的步骤，首先检验平稳性，是否需要进行差分。原假设 H0：序列为非平稳的，备择假设 H1: 序列是平稳的。检验结果为：

```
Results of Dickey-Fuller Test:
Test Statistic                -1.164089e+01
p-value                        2.148394e-21
#Lags Used                     1.400000e+01
Number of Observations Used    2.408000e+03
Critical Value (5%)           -2.862741e+00
Critical Value (1%)           -3.433069e+00
Critical Value (10%)          -2.567409e+00
```

p-value 小于显著性水平，Test Statistic 小于 Critical Value (1%)，拒绝原假设，选择备择假设，序列是平稳的，置信程度大于 99%。

只需使用 AR 模型，因此我们绘制 PACF 图像来直接判断阶数：

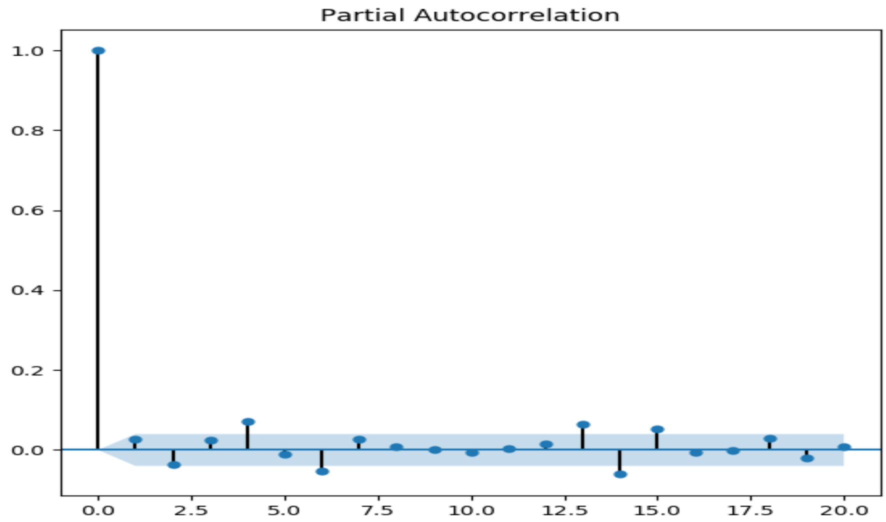

Figure.7 PACF

同样的，并没有较为严格的截断阶数，我们选择较为小的 4，作为我们的截断阶数。

最终选择 AR(4) 模型作为我们的均值方程。并对其进行拟合，同样使用 statsmodels.tsa.arima_model.ARMA 函数进行，本次顺利拟合，结果将用于后续操作。

5.2 ARCH 效应的检验

我们会用混成检验 (Ljung-Box 检验) 来检验序列残差的平方的相关性，来判断是否具有 ARCH 效应。残差直接通过实际结果以及通过 AR(4) 模型计算所得结果做差得出。

残差和残差的平方图像为：

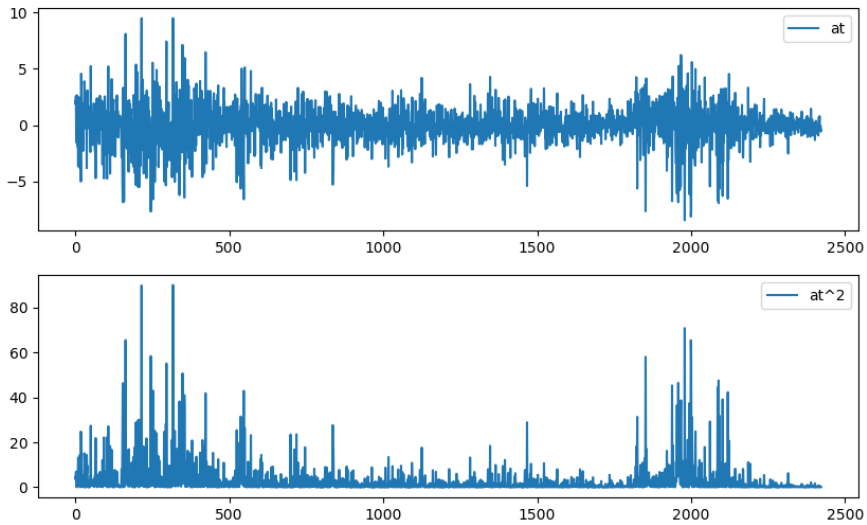

Figure.8 Residual (at) and its square

我们根据图像进一步确认了波动率聚集的现象。

对 at2 序列，即残差的平方的序列进行混成检验：原假设 H0: 序列没有相关性，备择假设 H1：序列具有相关性。我们检验 25 个自相关系数和 p-value。结果为：

Out[46]:

lag	AC	Q	P-value
1.0	0.210372	107.322182	3.783755e-25
2.0	0.230412	236.118084	5.340840e-52
3.0	0.258510	398.308946	5.145739e-86
4.0	0.243083	541.778081	6.148598e-116
5.0	0.248459	691.725516	3.022139e-147
6.0	0.156545	751.276080	5.167655e-159
7.0	0.208206	856.660497	1.093814e-180
8.0	0.190960	945.346486	9.309151e-199
9.0	0.189200	1032.441099	1.739771e-216
10.0	0.233956	1165.670269	3.654892e-244
11.0	0.161324	1229.043878	8.897359e-257
12.0	0.178410	1306.584138	1.899095e-272
13.0	0.152739	1363.439132	1.153664e-283
14.0	0.157863	1424.198051	1.002079e-295
15.0	0.194935	1516.883225	0.000000e+00
16.0	0.205126	1619.554969	0.000000e+00
17.0	0.191578	1709.150042	0.000000e+00
18.0	0.140841	1757.592689	0.000000e+00
19.0	0.227850	1884.431053	0.000000e+00
20.0	0.231556	2015.483984	0.000000e+00
21.0	0.179031	2093.857411	0.000000e+00
22.0	0.168266	2163.118215	0.000000e+00
23.0	0.111670	2193.635702	0.000000e+00
24.0	0.149153	2248.101294	0.000000e+00
25.0	0.168085	2317.299957	0.000000e+00

Figure.9 Test of Residual

p-value 小于显著性水平 0.05，我们拒绝原假设，即认为序列具有相关性。AC 也显著不为零。因此序列具有 ARCH 效应。

5.3 ARCH 模型的建立

在完全确定 ARCH 模型之后，我们需要确立 ARCH 模型的阶数，这可以通过残差的平方 at2 序列的偏自相关函数 PACF 来确定：

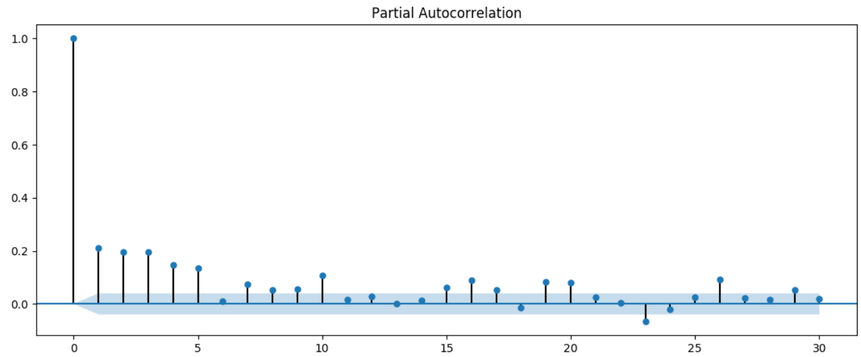

Figure.10 PACF of Residual's square

这则出现了明显的拖尾现象，是预期的结果。当然，也并不是十分理想，我们发现截尾的阶数可能很大，至少在我们所列出的 lags 中，并无法十分确定截尾阶数。为了计算的简便，我们选择截尾阶数为 5 。

接下来即对 ARCH(5) 模型进行拟合，均值方程为 AR(4) 模型，拟合的结果为：

```
                         AR - ARCH Model Results
==================================================================================
Dep. Variable:                            y   R-squared:                  -0.004
Mean Model:                              AR   Adj. R-squared:             -0.006
Vol Model:                             ARCH   Log-Likelihood:           -4468.66
Distribution:                        Normal   AIC:                       8959.32
Method:                  Maximum Likelihood   BIC:                       9022.98
                                              No. Observations:             2409
Date:                      Wed, Jun 21 2017   Df Residuals:                 2398
Time:                              09:40:16   Df Model:                       11
                                    Mean Model
==================================================================================
                coef    std err          t      P>|t|       95.0% Conf. Int.
----------------------------------------------------------------------------------
Const         0.0472  3.026e-02      1.560      0.119   [-1.210e-02,  0.107]
y[1]          0.0235  2.353e-02      0.998      0.318  [-2.264e-02,6.959e-02]
y[2]         -0.0478  2.753e-02     -1.737  8.233e-02     [ -0.102,6.130e-03]
y[3]     -3.2439e-03  3.142e-02     -0.103      0.918  [-6.483e-02,5.835e-02]
```

```
y[4]        -0.0315  3.306e-02   -0.952    0.341  [-9.627e-02,3.334e-02]
                              Volatility Model
==============================================================================
                 coef    std err          t      P>|t|       95.0% Conf. Int.
------------------------------------------------------------------------------
omega          0.9449      0.105      8.974  2.870e-19      [  0.739,  1.151]
alpha[1]       0.0655  3.650e-02      1.795  7.267e-02  [-6.023e-03,  0.137]
alpha[2]       0.1813  4.382e-02      4.137  3.524e-05   [9.538e-02,  0.267]
alpha[3]       0.2113  5.048e-02      4.185  2.846e-05      [  0.112,  0.310]
alpha[4]       0.1742  4.643e-02      3.752  1.754e-04   [8.321e-02,  0.265]
alpha[5]       0.1036  3.025e-02      3.424  6.163e-04   [4.429e-02,  0.163]
==============================================================================
```

我们发现一些拟合的系数并不稳定 ，对于均值方程和波动方程，一些系数甚至并没有显著地非零，其 95% 置信区间中包含零，同时方程的 R-squared 的绝对值很小，这让我们对拟合的结果产生了一些怀疑。

经过更多的信息搜集得知，这种结果在金融数据中也是比较正常的，因为市场环境往往比较恶劣，噪声的影响较为严重，使得我们拟合时信号噪声比很低，难以准确的拟合出理想的结果。

因此我们的 ARCH 模型的最终表达式为：

$$y_t = Const + \Sigma_{i=1}^{4} y[i] * y_{t-i} + \epsilon_t$$

$$\epsilon_t = \sigma_t * z_t. \quad z_t \sim i.i.d.(0,1)$$

$$\sigma_t^2 = omega + \Sigma_{i=1}^{5} alpha[i] * \epsilon_{t-i}^2$$

其中 y[i] 与 alpha[i] 取相应行的 coef 列对应的值。

5.4 ARCH 模型的检验和预测

对于整体的拟合情况，ARCH 模型的目的往往是反映出波动性的变化，尤其是波动率的聚集情况，拟合结果因此以波动率作为表示：

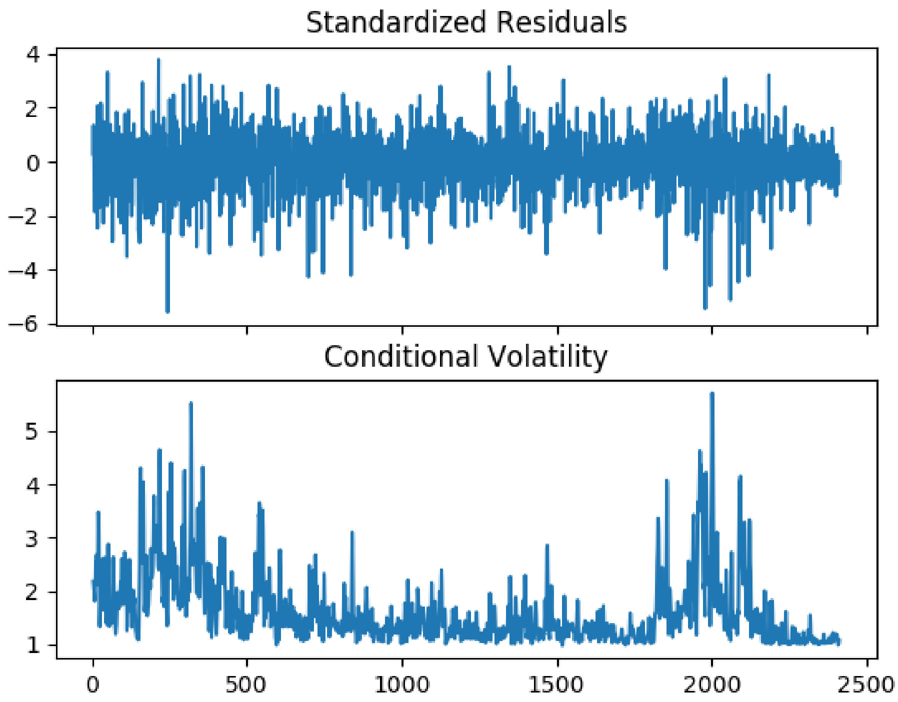

Figure. 11 Result of ARCH

其中第一张图是标准化的残差，这近似是一个平稳序列，说明模型在一定程度上是正确的。第二张图为条件异方差序列。

将其与我们估计的序列的波动进行对比，为：

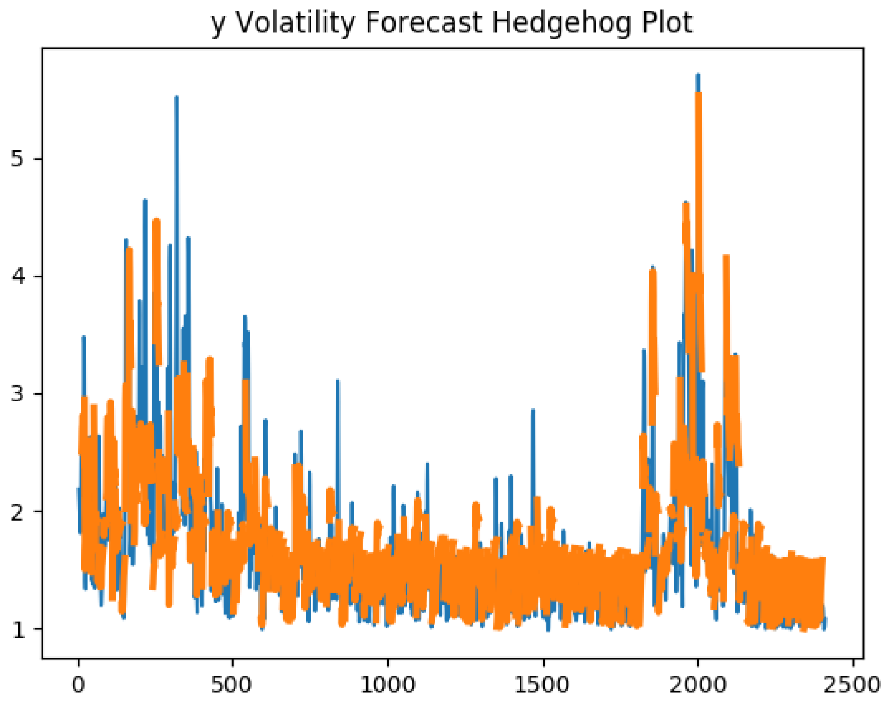

Figure. 12 Comparison of ARCH and Reality

其中橘黄色的为我们的 ARCH 模型估计的 σ_t^2，蓝色的为实际序列的 σ_t^2，我们发现两者的重合率十分之高，在末端波动较大的地方，我们的估计也很好的表现了波动率的聚集。在初始波动率较大的地方，有些估计值甚至比实际值要高，这可能是受到了后面的影响。但整体而言，对波动率的估计是较为准确的。

我们依据前面的数据，对最后十个数据进行预测（最后十个数据并未参加模型的训练），预测结果为：

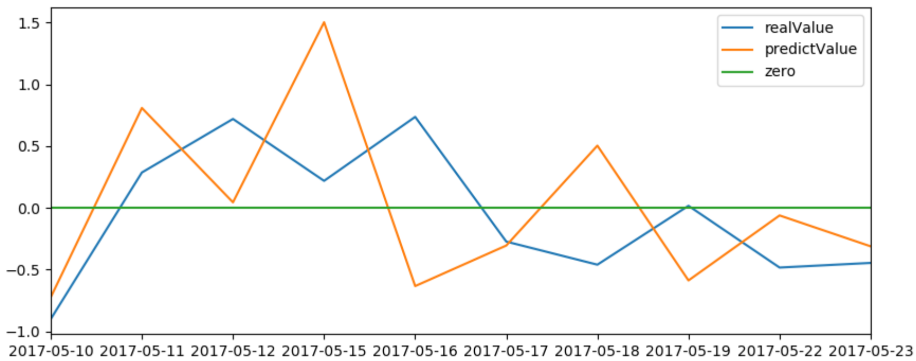

Figure. 13 Forecasting of ARCH

我们发现在实际的预测中表现的差强人意，我们的预测值往往与实际值相差较大，仅根据涨跌的正负情况来看，我们只有着 60% 的正确率。主要解释为两个方面：

- 对于 ARCH 模型，其主要性质为异方差，因此模型的好坏主要取决于方差的拟合情况，仅仅通过预测值来否定一个模型是较为武断的；
- 在金融领域中，60% 预测的正确率事实上已经相当之高，因为金融市场的噪声十分大。而在长期来看，我们依据该模型进行投资，依然有着较高的收益的期望。

关键在于对波动率的预测，下面我们对最后十个波动率进行预测：

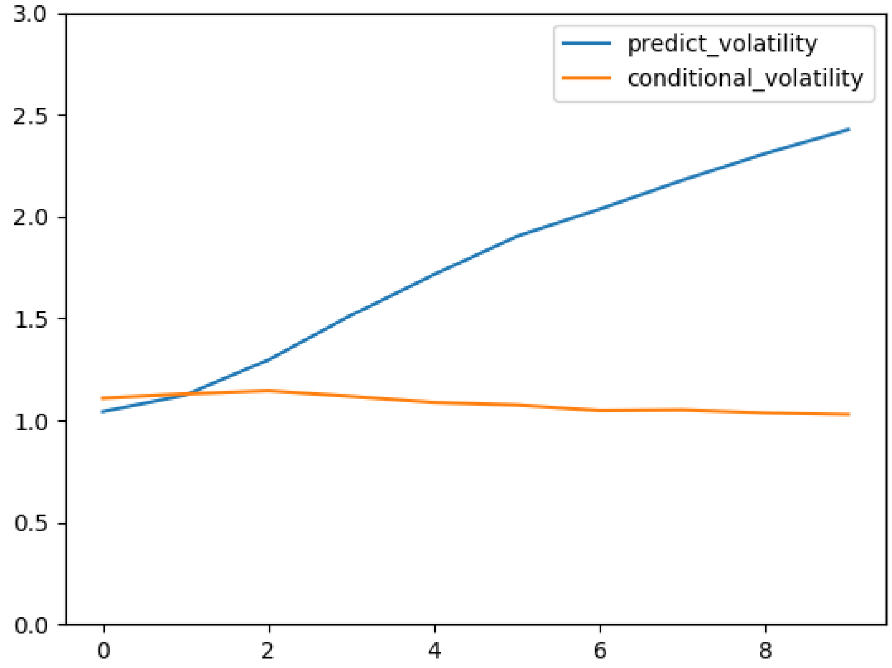

Figure. 14 Volatility Forecasting of ARCH and Reality

预测结果比实际要大许多，仅仅在量级上相同，总体上并不算太坏，可以作为参考。

综上，ARCH 模型能够不错的刻画出波动率的变化情况，同时表现出了较为明显的波动率聚集。相对于简单的 ARMA 而言，总体上来说是一种较为优秀的模型。但在预测波动率时并不能得到令人满意的结果，我们试图转向其他模型。

6 GARCH Model

为了更加充分的刻画收益的波动，同时更好的预测波动率的变化，我们往往需要更多的参数。此时 GARCH 模型可能更加符合我们的需求，接下来我们试使用 GARCH 模型来对序列进行浅析。

6.1 GARCH 模型的建立

类似于 ARCH 模型中的分析，我们使用 AR(4) 模型作为我们的均值方程，同时 ARCH 的性质既以得到检验，此处不再重复。与 ARCH 模型的建立过程也

类似，但此时 GARCH(m,s) 模型的定阶较为困难，人们往往使用低阶的 GARCH 模型，如 GARCH(1,1)，GARCH(1,2) 等。在金融领域，GARCH(1,1) 模型也常用于涨跌幅的拟合，我们此处使用 GARCH(1,1) 模型。

GARCH(1,1) 模型的拟合结果为：

```
                           AR - GARCH Model Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.005
Mean Model:                        AR   Adj. R-squared:                  0.003
Vol Model:                      GARCH   Log-Likelihood:               -4329.05
Distribution:                  Normal   AIC:                           8674.09
Method:            Maximum Likelihood   BIC:                           8720.39
                                        No. Observations:                 2409
Date:                Wed, Jun 21 2017   Df Residuals:                     2401
Time:                        10:43:37   Df Model:                            8
                                  Mean Model
==============================================================================
                 coef    std err          t      P>|t|       95.0% Conf. Int.
------------------------------------------------------------------------------
Const          0.0162  2.597e-02      0.622      0.534 [-3.475e-02,6.707e-02]
y[1]           0.0214  2.123e-02      1.007      0.314 [-2.024e-02,6.300e-02]
y[2]          -0.0181  2.189e-02     -0.825      0.409 [-6.097e-02,2.485e-02]
y[3]           0.0162  2.199e-02      0.738      0.461 [-2.688e-02,5.933e-02]
y[4]           0.0252  2.216e-02      1.137      0.256 [-1.824e-02,6.864e-02]
                              Volatility Model
==============================================================================
                 coef    std err          t      P>|t|       95.0% Conf. Int.
------------------------------------------------------------------------------
omega      6.3852e-03  4.515e-03      1.414      0.157 [-2.464e-03,1.523e-02]
alpha[1]       0.0494  1.022e-02      4.836  1.325e-06  [2.940e-02,6.947e-02]
beta[1]        0.9489  1.040e-02     91.281      0.000      [  0.929,  0.969]
==============================================================================
```

仅从标准差和 R-squared 来看，拟合的结果似乎与 ARCH 在同一量级，即并没有明显的改善。但从 beta[1] 的值来看，我们可以发现其显著地不为 0，即

具有较高的显著性，我们可以据此认为 GARCH 模型在此具有更多的意义，相对于 ARCH 模型，在波动率拟合的结果很可能更好。

我们得到的 GARCH(1,1) 模型表达式为：

$$y_t = Const + \Sigma_{i=1}^4 y[i] * y_{t-i} + \epsilon_t$$

$$\epsilon_t = \sigma_t * z_t. \quad z_t \sim i.i.d.(0,1)$$

$$\sigma_t^2 = omega + alpha[1] * \epsilon_{t-1}^2 + beta[1] * \sigma_{t-1}^2$$

其中 y[i], alpha[1], beta[1] 取相应行的 coef 列对应的值。

6.2 GARCH 模型的预测和检验

对于整体的拟合情况，GARCH 模型的目的同样是反映出波动性的变化和波动率的聚集情况，拟合结果因此以波动率作为表示：

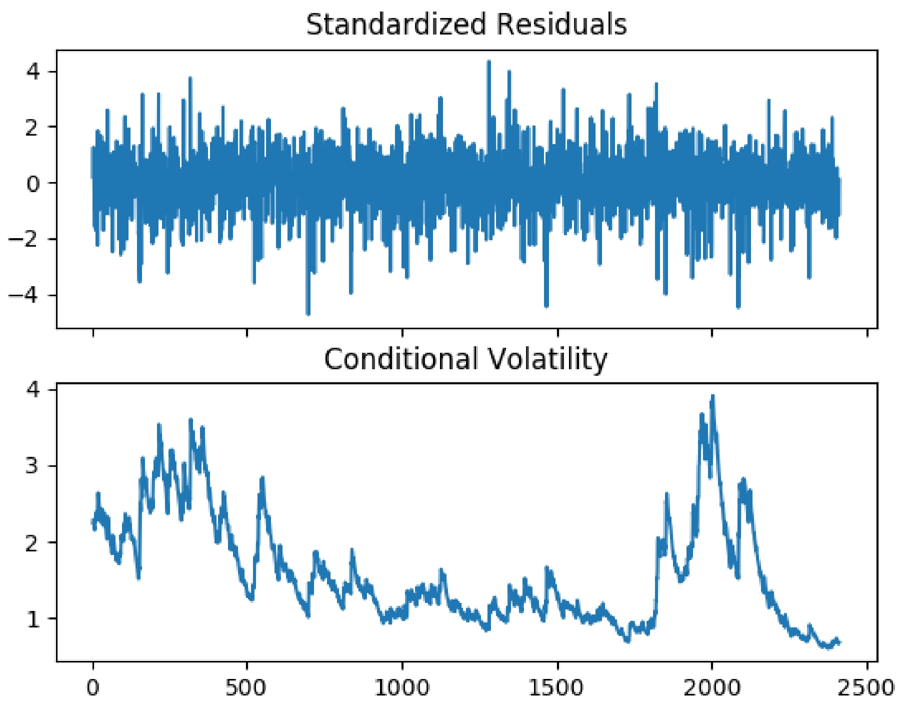

Figure. 15 Result of GARCH

图像的意义与 ARCH 对应的图像相同。

与估计序列的波动进行对比，为：

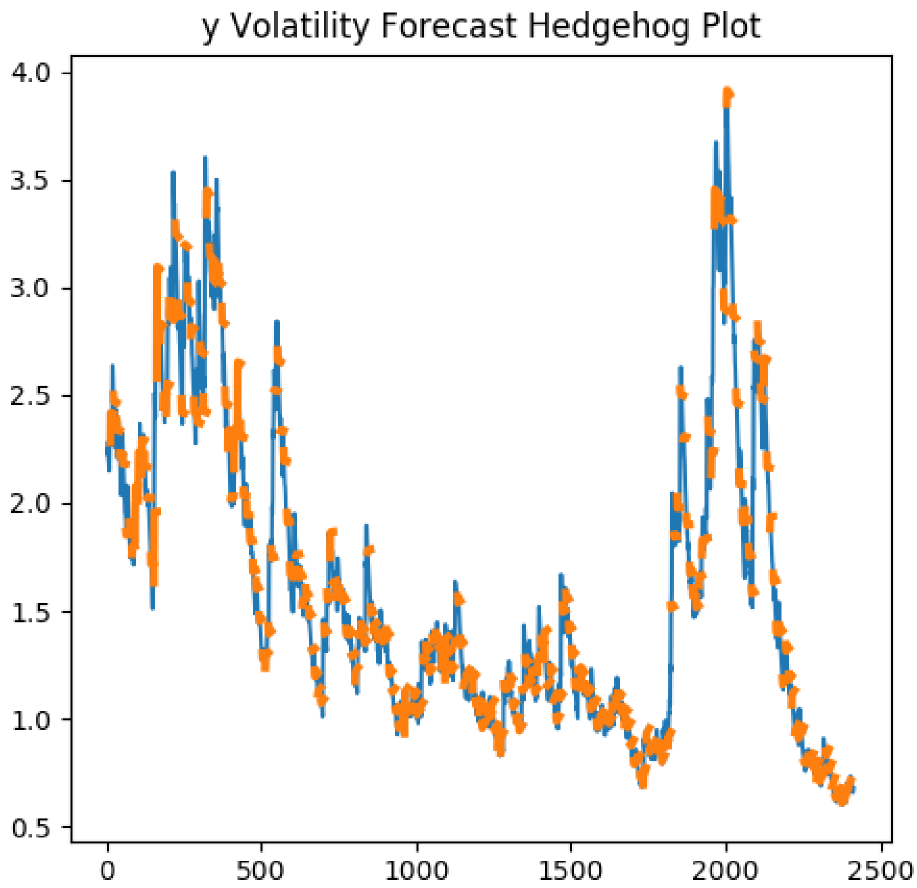

Figure. 16 Comparison of GARCH and Reality

其中橘黄色的为我们的 GARCH 模型估计的 σ_t^2，蓝色的为实际序列的 σ_t^2。我们发现基于数据可能较多的原因，一些结果暂时无法在绘图中显现。仅根据图像来看，两者的重合率相当之高，尤其是在波动率较大的地方，GARCH 模型的波动率与实际的波动率几乎相同，相对于 ARCH 模型，对波动率的预测显然更加良好。

我们不再对实际数据进行预测，仅对波动率进行预测，对最后十个波动率的预测结果结果为：

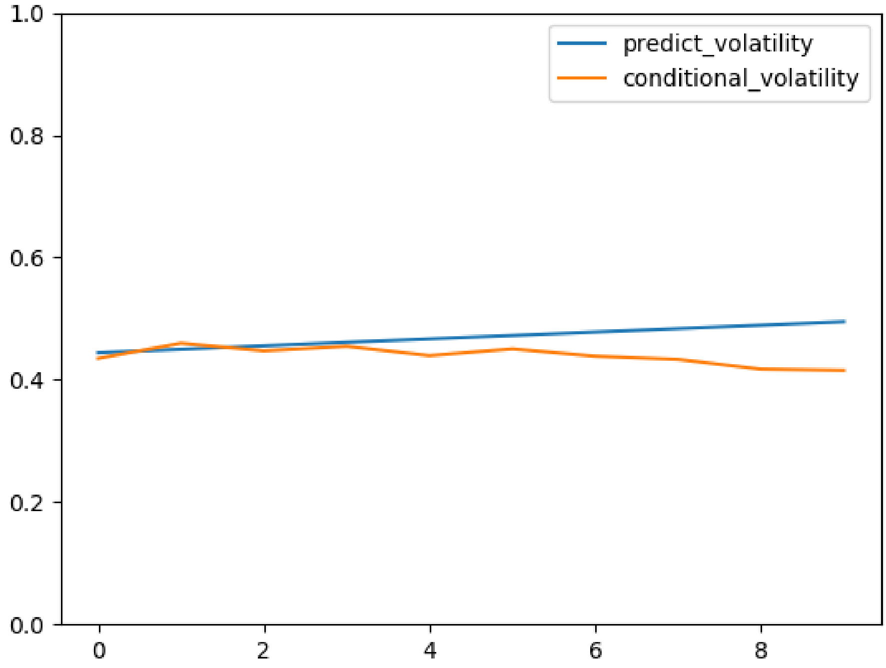

Figure. 17 Forecasting of GARCH and Reality

我们发现波动率的预测结果好的令人难以置信，当然，在拟合时较为显著的 beta 某种意义上已经预示了这一点，即 GARCH(1,1) 模型对波动率的预测和估计均十分良好，相对于 ARCH 模型，在最后十个波动上的预测优秀许多。

综上，在对波动率的预测方面，GARCH 模型总体来说是最为适合的，预测结果也十分令人满意。

7 总结

本次实验中，我们首先使用 ARMA 模型对开盘价 Open 进行了尝试，进行了一阶差分之后稳定，但在接下来的过程中遇到了问题，转向其他模型，并分析与 diff1 类似的涨跌幅 Ratio。在使用 ARCH 模型分析的过程中，我们着重于对股指波动率的估计和预测，ARCH 模型能够较好的体现出波动率聚集的现象，但在预测时表现并不十分良好。我们最终使用 GARCH 模型，较好的表现了波动率的变化情况，同时对波动率进行了十分准确的预测。

分析的主要意义在于，对波动率的估计可以应用到各个指数，股票中去，选择波动率较大的股票，可能存在着较多的套利机会，选择波动率较小的股票，资

产的风险就能够得到限制，同时可能会收到稳健的收益。

本次的主要工作在于：

- 对不同模型进行了尝试，并进行了比较；
- 建立并检验了 ARCH 和 GARCH 模型；
- 重点放在波动率的估计和预测方面，对此进行了模型的选择和改进。

不足之处在于：

- 基于现有工具的原因，暂时并没有使用更加高级的均值方程如 ARMA 模型，仅仅使用了 AR 模型。
- 图形的绘制较为粗糙，同时一些图像显示中出现问题，给分析带来一定的不便；
- 经过对相关专业人士的咨询，发现笔者所使用的数据过多，使得不同数据之间的相互影响较大。对于股指数据，近两年的已经足够，更多年份事实上并没有太多意义，它们之间也可能并没有太多的相互关系。

对数据可以进行的后续深入分析为：

- 选取更加高级的模型对价格或者涨跌幅进行更准确的估计和预测，而不是仅仅对波动率进行预测。
- 选择一些其他序列，如其他市场的波动情况，某个行业的一些参数，作为解释变量，来使用 VAR 模型，对指数的变动进行原因的分析和预测。

最后，对老师一学期的教导和助教的辛勤工作致以诚挚的谢意！

代一

二零一七年六月

III

附录

8 参考文献

[1] 《金融时间序列分析》第 2 版 Ruey S.Tsay 著王辉、潘家柱译
[2] Python Data Analysis. 2014 年十月第一版， Ivan Idris 著
[3] Python arch 4.1 https://pypi.python.org/pypi/arch
[4]Python Statesmodels http://www.statsmodels.org/stable/index.html

9 补充说明

除所引用参考文献之外，一些函数，程序同样参考了其他网络来源，主要有 StackOverflow, python wiki 等。