

# 中南财经政法大学

## 本科课程论文



论文题目：基于 ARIMA 及 LSTM 模型的石油期货价格分析及预测

院系名称：统计与数学学院

专业名称：应用统计 1701

作者姓名：曹奕涵

作者学号：201721090024

2019~2020 第二学期

使用 L<sup>A</sup>T<sub>E</sub>X 撰写于 2020 年 6 月 24 日

## 摘 要

本文主要研究了石油期货价格变化的问题。在石油期货价格极其不稳定的“后疫情”时代，合理的预测石油期货价格的长短期走势，变得非常重要。在中国银行“原油宝”事件之后，石油期货也在国内金融市场上掀起轩然大波。本文主要通过构造组合模型 ARIMA-LSTM，并与长短期的 ARIMA 模型、LSTM 进行对比，选择能够最优地对近 31 天的石油期货价格进行预测的模型。结果表明，单独的 LSTM 在长期以及短期拟合趋势上的结果最好，所有模型在短期预测的精度都不够理想。其主要原因是 ARIMA-LSTM 适用于呈现了较明显非线性趋势的时间序列数据，而石油期货价格数据，在一阶差分后即平稳，其拟合的 ARIMA 模型的残差也是随机序列，故难以再从中提取更多信息。在之后的改进中，应当进一步结合前沿时间序列模型，或者加入更多的有关变量，构建多变量模型。

**关键字：** ARIMA   LSTM   组合时间序列模型

# 目录

- 1 研究背景及方法 ..... 2
  - 1.1 研究意义 ..... 2
  - 1.2 研究方法 ..... 3
  - 1.3 研究流程 ..... 3
- 2 模型原理 ..... 4
  - 2.1 ARIMA 模型原理 ..... 4
  - 2.2 LSTM 模型原理 ..... 5
- 3 实证分析 ..... 8
  - 3.1 描述性统计分析 ..... 8
  - 3.2 单变量独立模型的建立 ..... 9
    - 3.2.1 长期 ARIMA 模型的拟合 ..... 9
    - 3.2.2 短期 ARIMA 模型的拟合 ..... 11
  - 3.3 LSTM 模型的训练 ..... 13
  - 3.4 组合模型 ARIMA-LSTM 的建立 ..... 14
- 4 结果及结论 ..... 16
  - 4.1 实证分析结果 ..... 16
  - 4.2 结论 ..... 16
- A 附录 ..... 18

# 1 研究背景及方法

## 1.1 研究意义

原油期货为一种重要的石油期货，其主要受原油价格与原有需求的影响。新冠疫情的影响之下，全球原油需求骤减。路上车辆变少，原油下游产业的工厂全部停工。为了稳定原油价格，就需要减少供应量，然而，在国际形势面前，石油大国却选择了进一步增产石油。在沙特宣布增产的第二天，美股熔断，金融市场崩溃，金融机构砍掉头寸，再次导致了美股的再次熔断。

新冠疫情下，原油价格暴跌，出乎了所有投资者的意料。然而，这却给了一些投机者抄底的希望，国内的部分散户想要在油价暴跌期间购买，并在反弹期间获利。他们通过购买中行的“原油宝”产品，去赌原油价格的变化趋势。中国银行的这一产品，与原油期货直接挂钩，而中行则主要通过对冲的手段，保证散户的多头空头能够相互赔付，从中赚取手续费。当中行代购的原油期货到期时，中行将有两种选择，一种是交割，也就是去取原油；另一种则是以交易价跟交易所交易。新冠期间，原油储存地早已被预购一空，这也就是说，所有的原油期货都只能在最后交易日和交易所交易，这样，才有足够的地方存储原油。于是，中行在 4 月 20 日晚 10 点，停止了原油宝的交易系统，并对持有原油下达了 TAS 指令，在凌晨 2:30，以当时结算价格成交。然而，原油期货在 2:08 即突破正值，走向了负数。在 2:30 时，以-37 美元每桶的价格成交，导致所有在原油宝购买多头的人，不仅仅亏光了所有本金，还背上了几十倍的负债。

上面的原油宝的案例，说明原油期货与经济形势，常常是相互挂钩的。原油期货价格的变化，不仅仅影响到个体投资者，同时也会对整个金融市场、经济形势产生巨大的影响。原油价格由 0.01 变为-0.01，在数学上可能只是 0.02 的变化，然而，对于市场来说，确实轩然大波。当原油价格变为负数，说明原油储量已经突破极限，储存原油的成本已经超过了将原油送到工厂进行加工。从历史上来看，几乎每一次的经济革命、金融剧变都有着原油期货价格的身影。2003-2004 年期间，世界经济增长，原油价格上升了 70% 以上；2007 年，美元走强游资，使得油价变化极大，急升骤降，而这次的疫情，使得原油期货价格历史上首次跌破负值，更是给了投资者们一个更大的“惊吓”。

以上案例均说明，研究原油期货的价格是非常有意义的。一个优良的指导模型，应当能够合理地预测原油期货价格的走势，能够丰富市场信息，带动金融市场发展。本文主要使用了 ARIMA 及 LSTM 模型，比较了两种模型的特性以及预测效果，并根据预测效果，对石油市场进行进一步的预测，提出合理建议。

## 1.2 研究方法

本文从美国能源信息部（Energy Information Administration, EIA）官方网站上收集了自 1986 年 1 月 2 日以来，OK WTI Crude Oil Future 产品上市之后的每个交易日的期货价格。据初步观察结果，该数据集仅有 2020 年 4 月 20 日这一天的价格为负值，这一天之后，价格均反弹回正值。本文将对数据进行描述性统计，对原油期货价格变化进行分析，并进一步结合单变量 LSTM 预测、单变量 ARIMA 预测、联合模型 ARIMA-LSTM 预测来判断模型的拟合效果，并给出进一步的建议。

## 1.3 研究流程

研究流程如下所示。

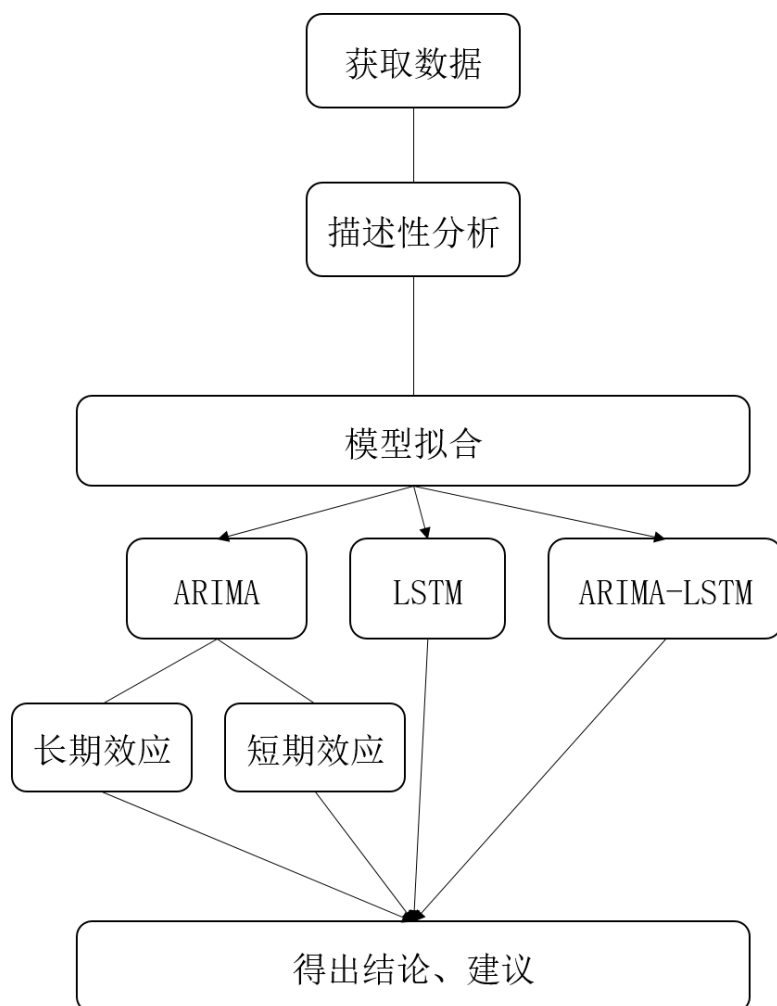


图 1.1: 研究流程图

## 2 模型原理

### 2.1 ARIMA 模型原理

ARIMA 模型是一种用于分析非平稳时间序列的经典模型，其基于的是 ARMA 模型。ARMA 模型是一种经典的平稳时间序列模型，如果说线性回归模型尝试通过某些解释性变量和随机项，去分析一个变量，那么 ARMA 模型就是时间序列中的线性回归模型。在 ARMA 模型中，时间序列的往期值，就是解释变量（AR），时间序列的噪声（MA），就是随机项。

ARMA(p,q) 模型由自回归（AR）模型以及移动平均（MA）模型两部分组成。当  $q=0$  时，该模型为 AR(p) 模型，当  $p=0$  时，该模型为 MA(q) 模型。ARMA 模型的基本结构如下：

$$\begin{aligned}\phi(B)X_t &= \theta(B)Z_t \\ \phi(z) &= 1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p \\ \theta(z) &= 1 + \theta_1 z + \cdots + \theta_q z^q \\ X_t &= \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}\end{aligned}\tag{2.1}$$

然而，并不是2.1所有的解都符合 ARMA 模型。ARMA 模型的一个重点，就是根据模型去做预测，如果解无法满足 ARMA 模型的前提条件，也就是时间序列是平稳的，那么预测结果将难以解释。一个 ARMA 过程平稳的条件为，其作为 AR 跟 MA 过程都应该是平稳的。

AR 过程的平稳性来自于 AR 过程的前后因果关系<sup>[3]</sup>。一个具有因果性的时间序列  $X_t$ ，其现值应当取决于过去值跟过去随机项。从数学性质上来说，就是  $\phi(z)$  的根全部在单位圆外。对于一个 AR(1) 过程：

$$X_t = \phi X_{t-1} + Z_t$$

其对应

$$\phi(z) = 1 - \phi z$$

上式存在单位根  $\phi^{-1}$ 。当且仅当该根不为 1 时，时间序列平稳。平稳性要求，对  $X_t = \Phi(B)Z_t = \sum_j \Phi_j Z_{t-j}$ ，对于  $\Phi = \frac{\theta}{\phi}$ ， $\phi_j = 0$ ，对每个  $j < 0$  成立。

那么，从复分析的角度来看，研究洛朗级数  $\Phi(z) = \sum_{j=-\infty}^{\infty} \Phi_j z^j$ ，这个洛朗级数应当在单位圆  $\{z \in C : |z| \leq 1\}$  的邻域上解析。那么，既然洛朗展开  $\sum_{j=-\infty}^{\infty} \Phi_j z^j$  在单位圆内解析，这就说明， $z$  的负幂次方不出现。所以， $\phi$  的根应当在单位圆外，因为只有这时， $\Phi = \theta/\phi$  在单位圆上是解析的。

MA 过程的平稳性来自其可逆性。可逆性的定义为，对于  $Z_t = \pi(B)X_t = \sum_j \pi_j X_{t-j}$ ， $\pi = \frac{\phi}{\theta}$

而言,  $\pi_j = 0$ , 对每个  $j < 0$  都成立。与其对应的数学性质为,  $\pi(z)$  在单位圆上解析, 或者  $\theta$  所有的根都在单位圆外。

而 ARIMA 模型, 就是基于 ARMA 模型的原理, 将原来不平稳的数据, 变成平稳的, 再拟合。将其变为平稳的方法, 就是做差分。这样, 就能对非平稳的时间序列, 也使用 ARMA 模型进行分析。

## 2.2 LSTM 模型原理

LSTM 是一种自然语言处理中经常会使用到的模型, 其为递归神经网络 (RNN) 的一种衍生模型。我们首先从 RNN 开始研究。RNN 是一类具有短期记忆能力的神经网络<sup>[1]</sup>, 通过随时间反向传播算法来进行参数的求解, 执行学习任务。然而, 随时间反向传播在处理长序列问题时, 常常存在梯度爆炸和消失的问题, 没有长期记忆性。这就好比我们的记忆能力, 在处理序列数据时, 我们的大脑除了短期记忆外, 还会记住数据变化的长期趋势, 但 RNN 却无法处理这个问题。正是基于这一点, 人们在 RNN 中引入了“门控”这一设置, 形成了能够筛选长短期信息的 LSTM 神经网络。

RNN 的工作原理, 是通过带自反馈机制的神经元, 处理时间序列数据。假设现在有一个序列  $x_{1:T} = \{x_1, x_2, \dots, x_T\}$ , 那么数据通过输入层, 进入隐藏层, 会进入一个延迟器, 该延迟器能够输出一个活性值  $h_t$ :

$$h_t = f(h_{t-1}, x_t), h_0 = 0 \quad (2.2)$$

活性值将作为隐藏层的输出, 再进入输出层。

通过 BPTT (随时间反向传播)<sup>[4]</sup> 算法能够对模型进行参数学习。其将 RNN 看作是一个展开的多层前馈网络, 每一层对应每一个时刻, 这样, RNN 的梯度计算可以之间使用前馈网络中的 BP 算法, 因为, 在展开的网络中, 这些层的参数都是共享的。

假设给定一个训练样本  $(x, y)$ , 其中  $x_{1:T} = (x_1, \dots, x_T)$  对应  $y_{1:T} = (y_1, \dots, y_T)$ 。定义损失函数  $L_t = L(y_t, g(h_t))$ 。其中,  $g(h_t)$  为输出。整个序列的损失函数关于参数, 设为  $U$  的梯度为:

$$\frac{\partial L}{\partial U} = \sum_{t=1}^T \frac{\partial L_t}{\partial U} \quad (2.3)$$

现在计算偏导  $\frac{\partial L_t}{\partial U}$ 。参数  $U$  和隐藏层在每个时刻的净输入  $z_k = Uh_{k-1} + Wx_k + b$  有关, 故这里使用链式法则, 有:

$$\frac{\partial L_t}{\partial u_{ij}} = \sum_{k=1}^t \frac{\partial z_k}{\partial u_{ij}} \frac{\partial L_t}{\partial z_k} \quad (2.4)$$

如果我们保持  $z_k$  中  $h_{k-1}$  不变, 对其求偏导 (需要知道, 因为当前隐藏层输入  $z_k$  只跟当前值

$x_k$  与前期活性值  $h_{k-1}$  有关, 所以其他维度上, 都是 0), 计算得到:

$$\frac{\partial z_k}{\partial u_{ij}} = [0, \dots, [h_{k-1}]_j, \dots, 0] \quad (2.5)$$

再求后面一个微分的值:

$$\frac{\partial L_t}{\partial z_k} = \frac{\partial h_k}{\partial z_k} \frac{\partial z_{k+1}}{\partial h_k} \frac{\partial L_t}{\partial z_{k+1}} = \text{diag}(f'(z_k)) U^T \frac{\partial L_t}{\partial z_{k+1}} \quad (2.6)$$

令 2.6 为  $\delta_{t,k}$ , 那么就可以得到参数梯度。

这就揭示了一个问题, 如果我们把 2.6 展开的话, 我们应该能够得到如下形式一个式子:

$$\delta_{t,k} = \gamma^{t-k} \delta_{t,t}. \quad (2.7)$$

那么, 如果  $\gamma > 1$ , 当  $t - k \rightarrow \infty$  时, 梯度就会爆炸。当  $\gamma < 1$ , 梯度就会消失, 所以这实际上是一个非常不稳定的模型, 简单的神经网络难以建设这种长程依赖关系。我们需要对其进行改善, 于是, 就加入了“门控机制”, 门控机制能够在模型中有选择的加入新信息, 并且遗忘之前的信息, 这类网络被称为 Gated RNN (GRNN)。通过这种方法, 能够控制信息的累积速度。

本文中主要使用的是 LSTM, 长短期记忆网络<sup>[2]</sup>。LSTM 在 RNN 的基础上, 引入了一个新的内部状态  $c_t$ , 进行线性的递归信息传递, 同时非线性地导出信息给外部状态  $h_t$ :

$$\begin{aligned} c_t &= f_t c_{t-1} + i_t \hat{c}_t \\ h_t &= o_t \times \tanh(c_t) \end{aligned} \quad (2.8)$$

其中,  $f_t$ 、 $i_t$ 、 $o_t$  为三个门, 来控制信息的传递。门是一个二值变量  $[0, 1]$ , 当为 0 时关闭, 1 时打开。LSTM 中, 门的取值在  $(0, 1)$  之间, 以一定的比例允许信息通过。

上面的三个门分别定义为:

1. 遗忘门:  $f_t$  控制上一个时刻的内部状态,  $c_{t-1}$ , 需要遗忘的信息;
2. 输入门:  $i_t$  控制当前时刻的候选状态  $\hat{c}_t$  需要保存多少信息;
3. 输出门:  $o_t$  控制当前内部状态  $c_t$  有多少信息需要输出给外部状态。

这些值均为  $\sigma(z_t)$ 。

在递归神经网络中, 隐状态  $h$  存储了所有的历史信息, 在简单递归网络中,  $h$  的每个时刻都会更新一次。在 LSTM 中, 记忆单元  $c$  可以在某个时刻捕捉到关键信息, 并且将关键信息保存一定的时间间隔, 且这个时间间隔总要长于短期记忆  $h$ 。这就减少了每次计算需要存



储的短期信息量，当时间序列变长，或者文本变长时，LSTM 能够筛选出长期序列的趋势，避免梯度消失、梯度爆炸的问题。

## 3 实证分析

### 3.1 描述性统计分析

对数据进行简单的探索性分析。首先对期货价格数据作时间序列图如下:

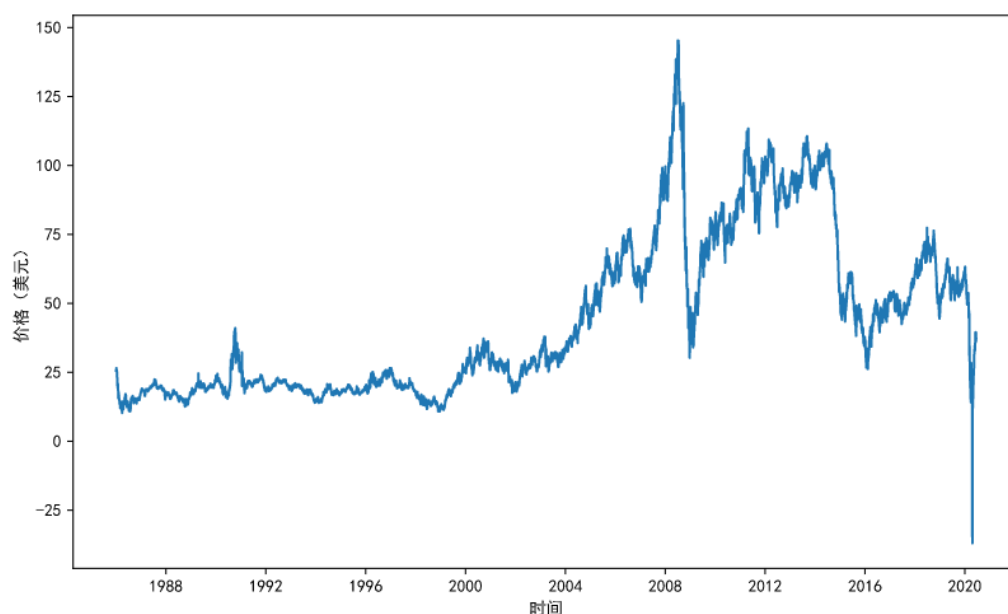


图 3.1: 石油期货价格（每加仑）1986-2020

本数据集收集了自 1986 年 1 月 2 日开始的 OK WTC Crude Oil Future 的交易日价格，到 2020 年 5 月 20 日为止。5 月 20 日之后的数据将作为测试数据，在模型结果与建议中进一步分析。根据上图3.1可以看出，石油期货价格变化幅度较大，并没有明显的显性趋势，说明需要进行进一步的差分处理。总体趋势是上升的，但并不稳定，在 2008 年、2015 年、2020 年出现了较大的变化。对其做箱线图如下：

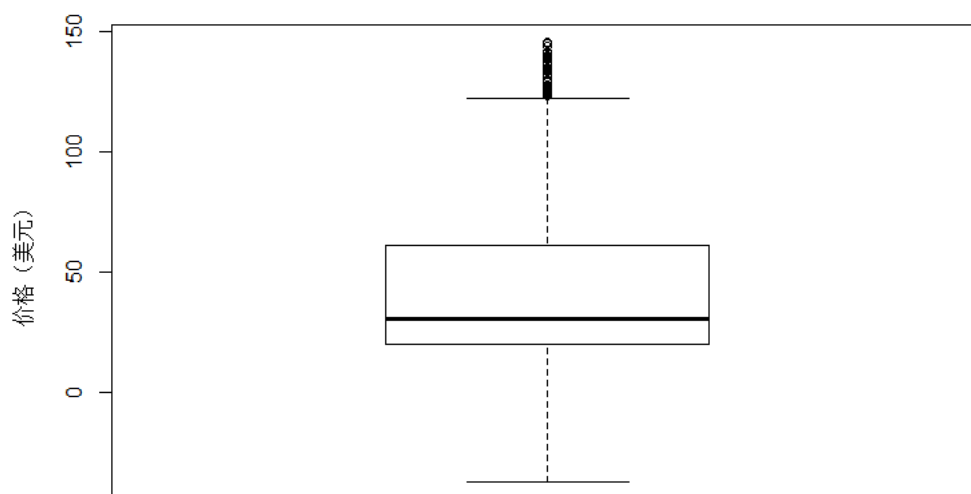


图 3.2: 石油期货价格（每加仑）1986-2020 箱线图

根据图3.2，平均价格为 30 美元左右，存在较多的高价格的离群值。对数据进行均值、方差计算：

表 3.1: 描述性统计结果表

样本量	标准差	均值	中位数	最小值	最大值
8652	29.13	44.1	31.12	-36.98	145.31

平均价格为 44.1 美元，标准差较高，说明数据的总体变化程度较大，中位数为 31.12 美元，说明整个数据的价格水平位于 30 美元左右。在正常状态下，每加仑的石油期货价格为 30 美元。

## 3.2 单变量独立模型的建立

### 3.2.1 长期 ARIMA 模型的拟合

长期 ARIMA 模型选择了数据集中的所有数据，用这些数据一起来进行模型拟合与分析。

对数据进行 ADF 单位根检验，检查序列平稳性，结果显示，滞后期数为 20 期，p 值为 0.5656，接受原假设，认为序列不平稳。此时，需要对数据进行差分处理，使其成为平稳的时间序列。做一阶差分后，对数据重新作图如下：

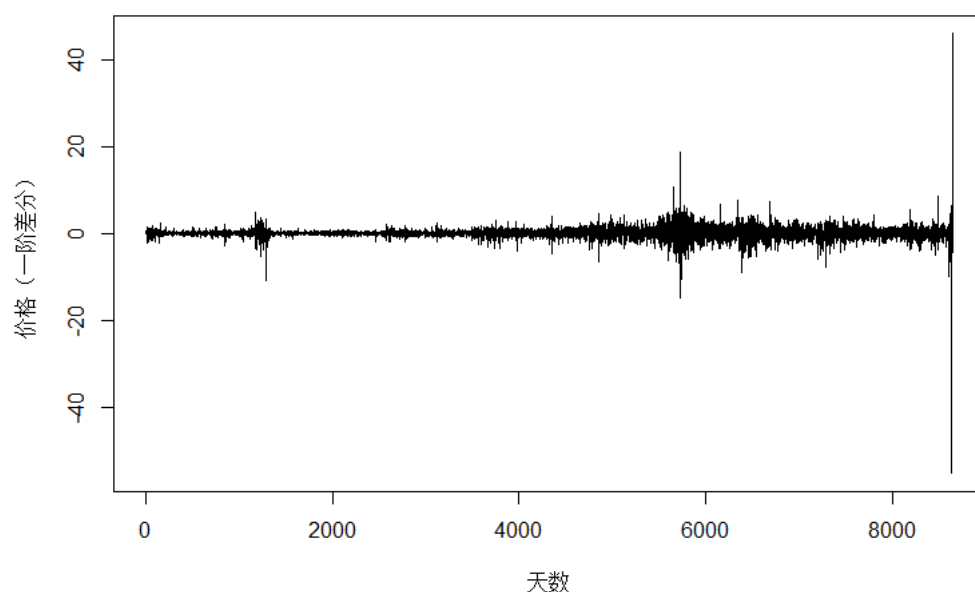


图 3.3: 一阶差分后石油期货时间序列图

如图所示, 序列前部分已经趋近平稳, 后部分仍然呈现不平稳的趋势。对差分后的数据进行 ADF 单位根检验。根据检验结果, 滞后期为 21 期,  $p$  值小于 0.01, 为平稳的时间序列。故原数据适合进行  $ARIMA(p,1,q)$  模型的拟合。

计算一阶差分后的时间序列的 ACF 以及 PACF, 并作图, 以确定 ARIMA 模型的阶数。作图如下:

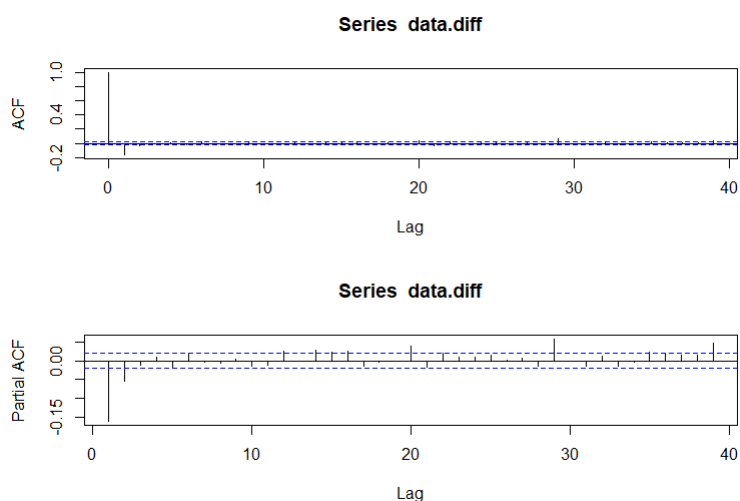


图 3.4: ACF 与 PACF 图 (一阶差分)

根据图3.4, ACF 于 2 阶截尾, PACF 拖尾, 故拟合  $ARIMA(0,1,2)$ , 计算得到 AIC 值为

30139.15。模型的估计值如下：

表 3.2: ARIMA(0,1,2) 模型拟合结果

变量	估计值	标准误
ma1	-0.1723	-0.0259
ma2	0.0108	0.0107

根据上述模型，对接下来 30 天的值进行预测，预测结果如图3.7所示。

### 3.2.2 短期 ARIMA 模型的拟合

考虑到数据信息的有效性，选择近 1000 天的数据进行 ARIMA 模型的拟合。首先对近期的 1000 条数据进行 ADF 单位根检验，检验结果显示，滞后期为 6，p 值为 0.9562，接受原假设，认为序列不平稳。对其做一阶差分后，作图如下：

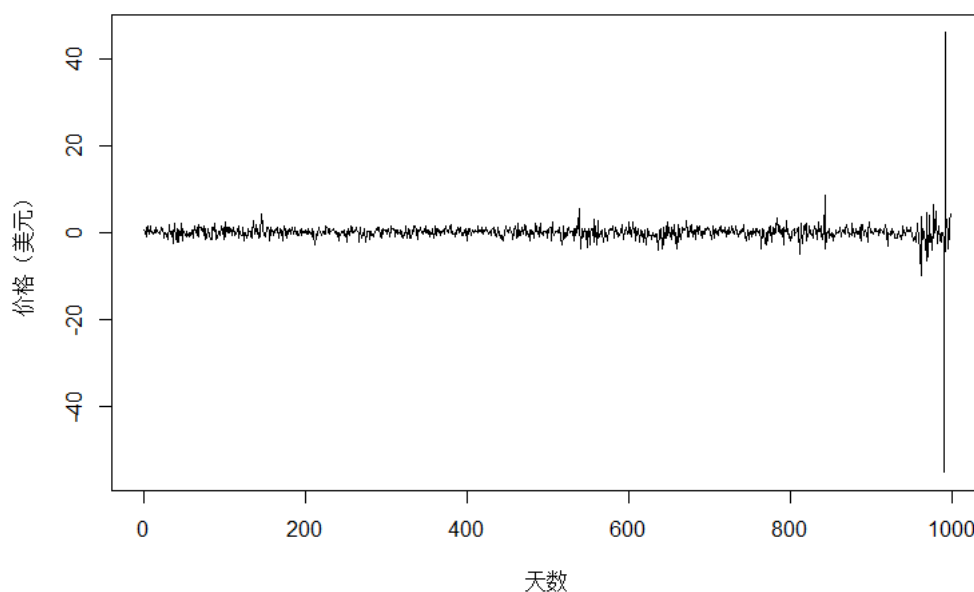


图 3.5: 每加仑石油期货价格（一阶差分后）

除了某些离群点，数据以及基本接近平稳。对数据进行 ADF 单位根检验，p 值小于 0.01，认为已经为平稳序列。计算差分后数据的 ACF 以及 PACF，作图如下：

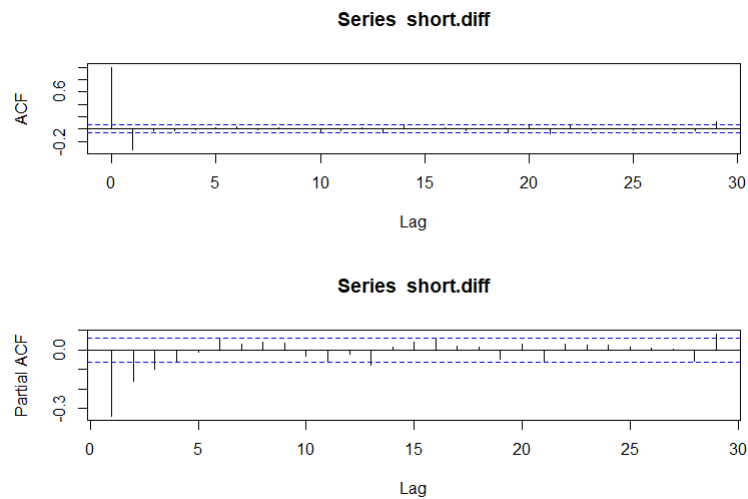


图 3.6: PACF 以及 ACF 值

根据上图，确定模型阶数为  $ARIMA(0,1,1)$ 。移动平均项  $ma1$  的估计值为 -0.4352，标准误差为 0.0294，AIC 值为 4593.08。

对后 30 期进行预测，结果如下：

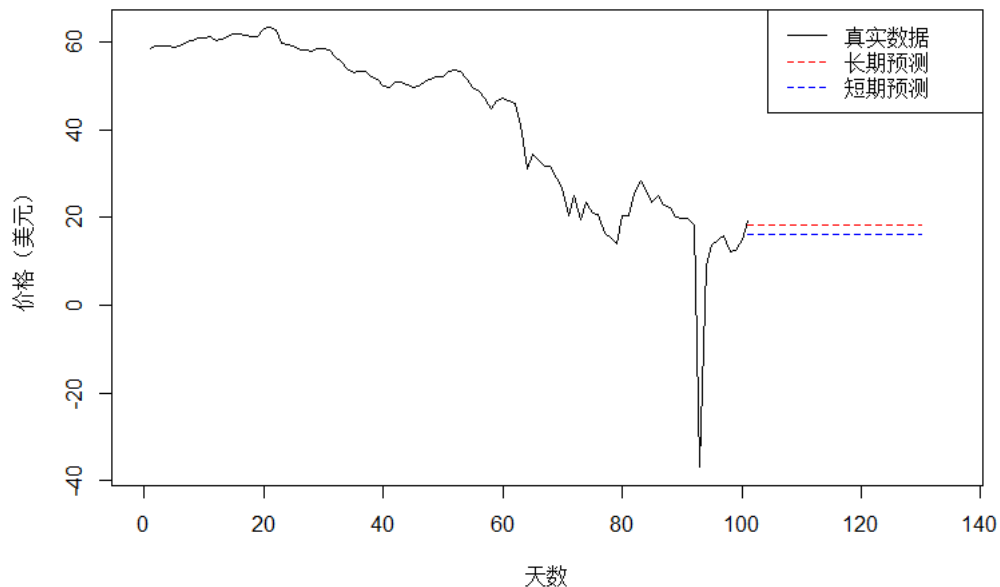


图 3.7: 石油期货价格（每加仑）30 天 ARIMA 预测值

对上面的 ARIMA 水平建模模型进行进一步的探索，对残差序列进行纯随机性检验，检查是否需要进一步拟合 GARCH 模型。检验结果  $p$  值均大于 0.995，说明并不需要再进行 GARCH 模型的拟合。对残差序列及残差平方项作图如下：

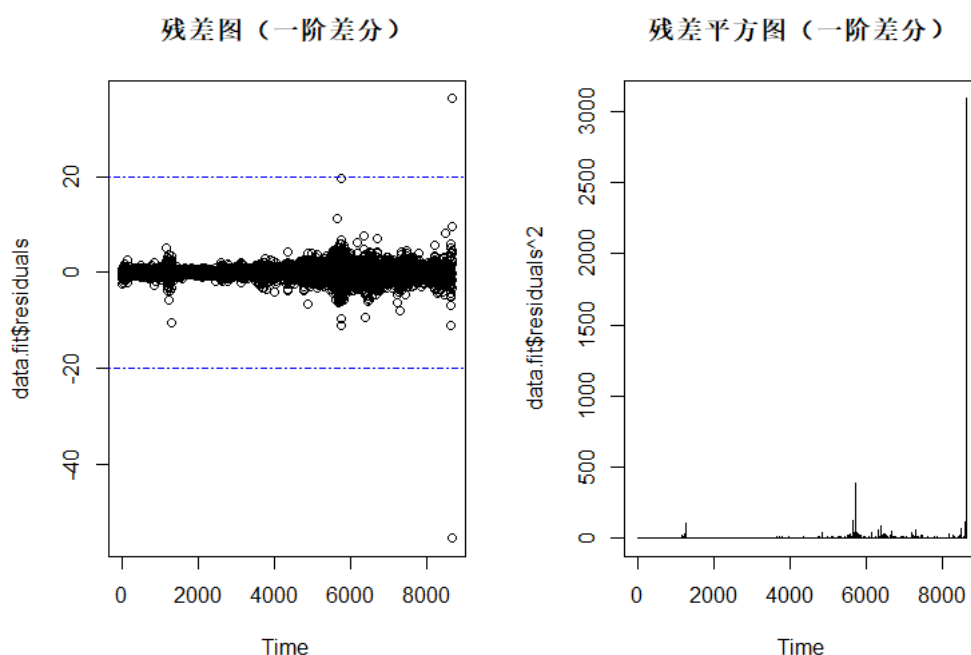


图 3.8: 长期 ARIMA 模型残差值

### 3.3 LSTM 模型的训练

对单变量的时间序列进行 LSTM 拟合, 需要假设单变量时间序列的现值受多少期滞后值的影响。本文中设置滞后期为 30 期, 即现值主要受前近一个月的值的影响。在训练时, 首先分离前期数据与后 300 期数据, 将 300 期数据作为待测数据对其进行预测, 以便于查看模型的拟合情况。

设置隐含层为 1 层全连接层, 同样也可以设置两层的全连接层, 没有太大影响。鉴于训练时间与硬件设备上的不足, 本文只能将训练次数限制在 10 次, 学习率设置在 0.0001。优化器选择 AdamOptimizer, 损失函数选择 MSE 损失。对模型进行训练, 模型收敛速度较快, 在第 9 次时损失达到最小。对后面的 300 期进行预测, 结果如下:

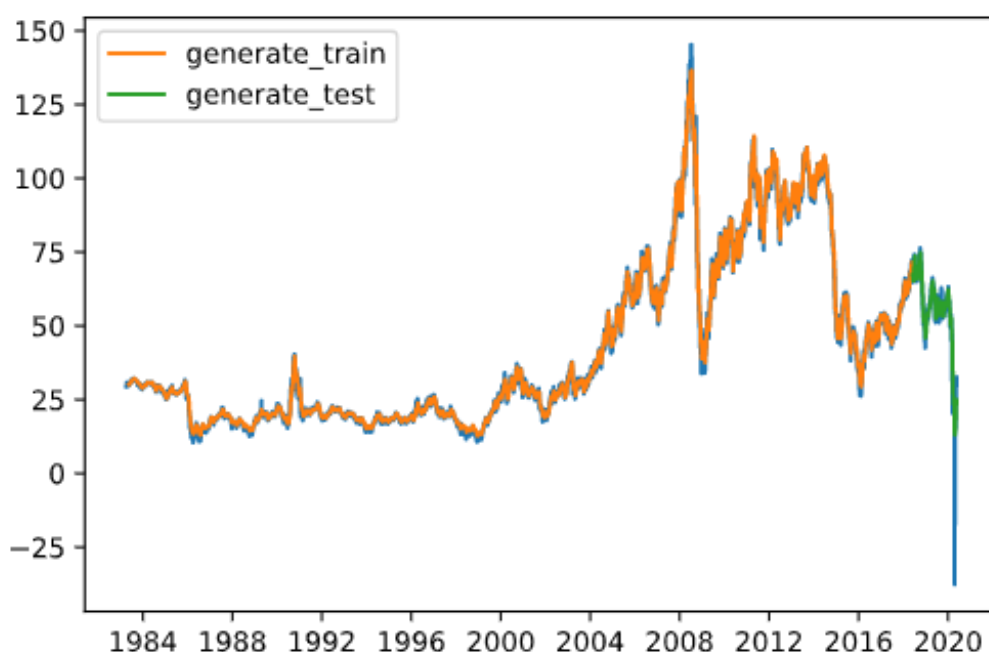


图 3.9: LSTM 模型的拟合效果

图中，蓝色为原数据，橙色为用于训练的数据集，绿色为用于测试的数据集，可以看到，拟合的效果非常好，精确预测了未来数据的发展态势。对其预测的 31 期结果进行平方和计算，结果为 6525.786。

### 3.4 组合模型 ARIMA-LSTM 的建立

正如模型原理中所提到的，ARMA 模型是时间序列模型中的线性模型，其拟合的主要为数据集的线性趋势，而很多时候，模型中还存在着非线性趋势。这种非线性趋势就需要使用其他方法来对数据进行拟合。本文建立了 ARIMA-LSTM 模型，主要通过 ARIMA 模型来拟合数据中的线性趋势，再对模型残差，使用 LSTM 拟合数据中的非线性趋势。

组合模型的建立，分成两部分进行检验。模型 1 为根据上文中的长期 ARIMA 模型 (3.2.1)，对 1000 期进行拟合预测，记为  $L_t$ ，为线性项。根据真实值，可以得到  $Y_t - L_t$ ，记为  $\Sigma_t$ ，为非线性项。使用 LSTM，保持上面的参数设置（一层全连接，Adam 优化，20 次循环训练，MSE 损失），对 1000 项的  $\Sigma_t$  进行拟合，预测 31 期，记为  $\Sigma_{1:31}$ 。再根据 3.2.1，再次预测 31 期，记为  $L_{1:31}$ 。那么， $\hat{Y}_{1:31} = L_{1:31} + \Sigma_{1:31}$ 。

对拟合的时间序列，计算拟合值，做出图像如下：



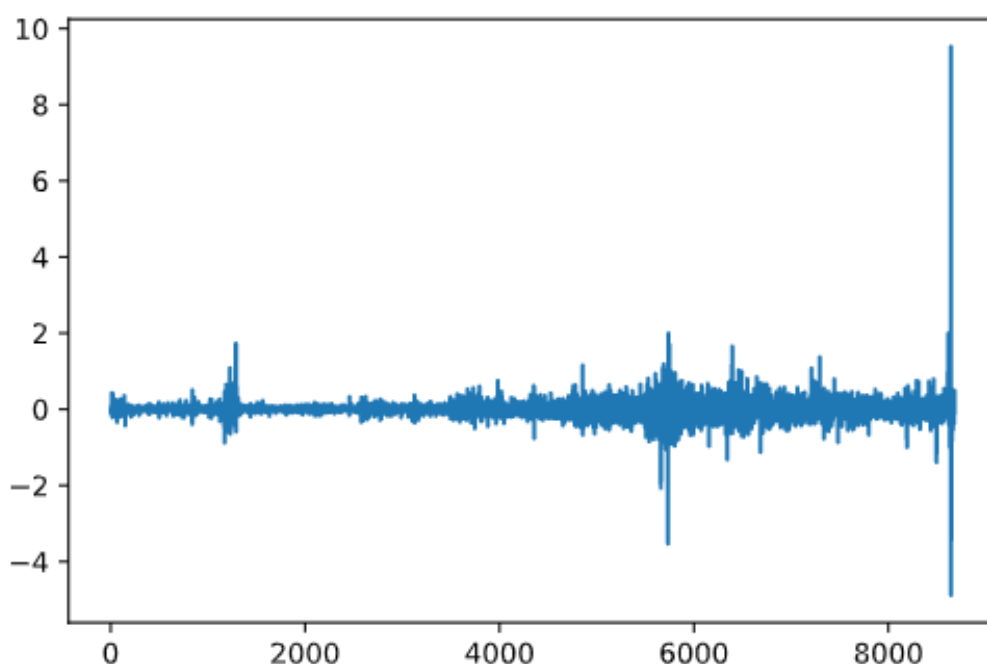


图 3.10: ARIMA(0,1,2) 拟合值

上图为对差分后序列的拟合情况，实际上已经拟合的较好，此时的残差序列中不一定仍含有非线性信息。实际上，对于这样的数据集而言，并不一定需要继续对残差进行 LSTM 的拟合。对于含有明显的非线性趋势的模型，可以进一步使用该方法，在模型的拟合过程中，也没有很明显的收敛趋势，说明该序列的随机性很强。对残差进行 LSTM 的拟合，并预测 31 期，与预测值相加，计算得到长期 ARIMA 的总误差平方为 6739.684，长期组合 ARIMA 的总误差平方为 6831.304。短期 ARIMA 的总误差平方为 7873.18，组合模型的总误差平方为 7977.044。总体来说，模型拟合的结果并不够好，这主要是因为序列中并不存在着明显的非线性趋势。

## 4 结果及结论

### 4.1 实证分析结果

本文一共拟合了 4 种主要的模型，对应的 31 期预测的误差平方和如下：

表 4.1: 模型预测误差平方和 (31 期)

模型	误差
长期 ARIMA(0,1,2)	6739.684
短期 ARIMA(0,1,1)	7873.18
长期组合模型 (ARIMA(0,1,2) + LSTM)	6831.304
短期组合模型 (ARIMA(0,1,1) + LSTM)	7977.044
LSTM 模型	6525.785

根据 31 期的预测误差结果，认为单独的 LSTM 模型预测效果最佳。

### 4.2 结论

本文主要通过 ARIMA 模型与 LSTM 模型的结合分析，对石油期货价格的变化进行了预测分析。根据分析的结果，单独的 LSTM 模型预测效果最佳。实际上，这个结果并不一定是一个稳定的结果。从数据量上来看，预测 31 期的值较少。并且，ARIMA 模型并不具有长期预测的能力，相比之下，LSTM 的长期趋势预测更为精确。从短期趋势上来看，长期的 ARIMA 模型获取的信息更为全面，虽然模型比短期的 ARIMA 模型稍微复杂一些，但是精度却比短期的 ARIMA 模型高很多。

本文的主要研究对象为组合的 ARIMA-LSTM 模型，但是实际上，ARIMA-LSTM 在石油期货数据上的表现能力并不够好。这主要是因为 ARIMA 适用于线性趋势的分析，而石油期货的数据具有线性趋势的基本特征，排除线性趋势后，除个别异常点外，残差序列呈随机游走状态，这说明数据并没有明显的非线性趋势。对于呈现了非线性趋势的模型而言，ARIMA-LSTM 的表现更好，因为去掉线性趋势后，LSTM 能够拟合出优秀的非线性变化。

当然，这并不意味着模型没有可以改进的地方。对于单独的 LSTM 模型，可以增加网络的训练次数、隐藏层网络的层数。通过引入注意力 (Attention) 机制，可以合理地对变量（前 30 天的时间序列值）进行赋权，从而进一步改进模型。一般来说，增加了 Attention 机制的网络，其预测的精度往往更加准确。

## 参考文献

- [1] Jeffrey Elman. Finding structure in time, pages 289–312. 02 2020.
- [2] Klaus Greff, Rupesh Srivastava, Jan Koutník, Bas Steunebrink, and Jürgen Schmidhuber. Lstm: A search space odyssey. IEEE transactions on neural networks and learning systems, 28, 03 2015.
- [3] A.W. van der Vaart. Time series lecture notes. pages 117–125, 2010.
- [4] Paul Werbos. Backpropagation through time: what it does and how to do it. Proceedings of the IEEE, 78:1550 – 1560, 11 1990.

## 附录 A 附录

代码文件与原始数据，详见‘代码’、‘数据’文件夹。