# Black-Litterman模型

题中给出了仅给出了比特币和黄金的价格变化以供投资。因此，不可避免地，我们在投资过程中比特币和黄金占有一个配比，这里使用了贝叶斯框架下的Black-Litterman模型。它是从投资的实务出发的一种收益率贝叶斯收缩，并从市场的供需出发，认为投资在整个市场中按其市值的占比体现了当前时常供需关系的均衡状态。

The title gives only the price changes of bitcoin and gold for investment. Therefore, inevitably, we occupy a ration between bitcoin and gold in the investment process, and the Black-Litterman model in a Bayesian framework is used here. It is a Bayesian contraction of returns from the practice of investment and starts from the supply and demand in the market, considering that the investment reflects the current equilibrium of supply and demand from time to time in the overall market in proportion to its market value.

假设我们要在 N 个投资品之间进行资产配置。马科维茨的现代资产配置理论以这些投资品的期望收益率和协方差矩阵作为输入，通过最优化下列目标函数求出最佳的投资组合：

其中 表示投资品的期望收益率向量， 表示投资品的协方差矩阵， 表示投资者的风险厌恶 系数， 则是投资品在投资组合中的配置权重，在这里也就是投资品市值与市场总市值的比值。我们可以通过反推市场均衡状态下各投资品的收益率，把它作为先验：

Suppose we want to allocate assets among N investments. Markowitz's modern asset allocation theory takes as input the expected return and covariance matrix of these investments and finds the optimal portfolio by optimizing the following objective function.

where denotes the vector of expected returns of the investments, denotes the covariance matrix of the investments, denotes the risk aversion coefficient of the investors, and is the allocation weight of the investments in the portfolio, in this case the ratio of the market value of the investments to the total market value. We can use to invert the return of each investment in market equilibrium by taking it as a priori:

# ARIMA——Autoregressive Integrated Moving Average Model

将预测对象随时间推移而形成的数据序列视为一个随机序列，用一定的数学模型来近似描述这个序列。这个模型一旦被识别后就可以从时间序列的过去值及现在值来预测未来值。

ARIMA模型含有三个参数：p,d,q。

p--代表预测模型中采用的时序数据本身的滞后数(lags) ,也叫做AR/Auto-Regressive项

d--代表时序数据需要进行几阶差分化，才是稳定的，也叫Integrated项。

q--代表预测模型中采用的预测误差的滞后数(lags)，也叫做MA/Moving Average项

The data series of the forecast object over time is considered as a random series, and a mathematical model is used to approximate this series. Once the model is identified, it can predict the future values from the past and present values of the time series.

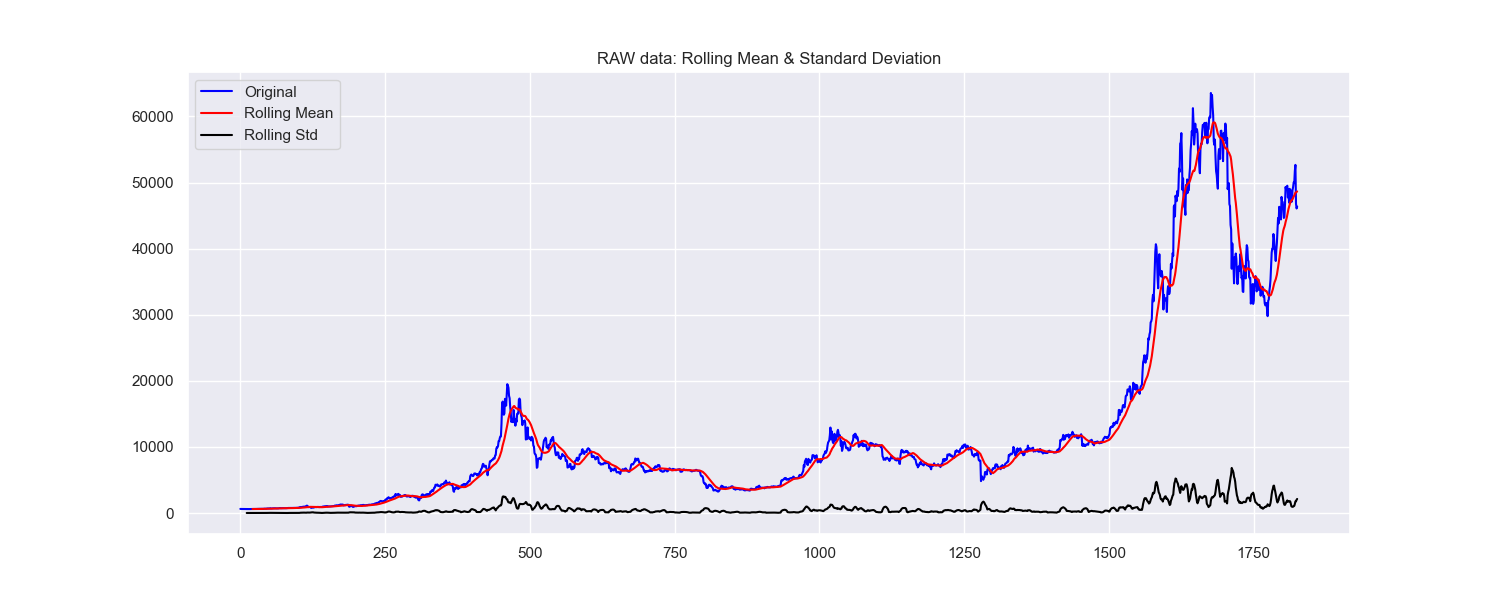
The ARIMA model contains three orders: p,d,q.

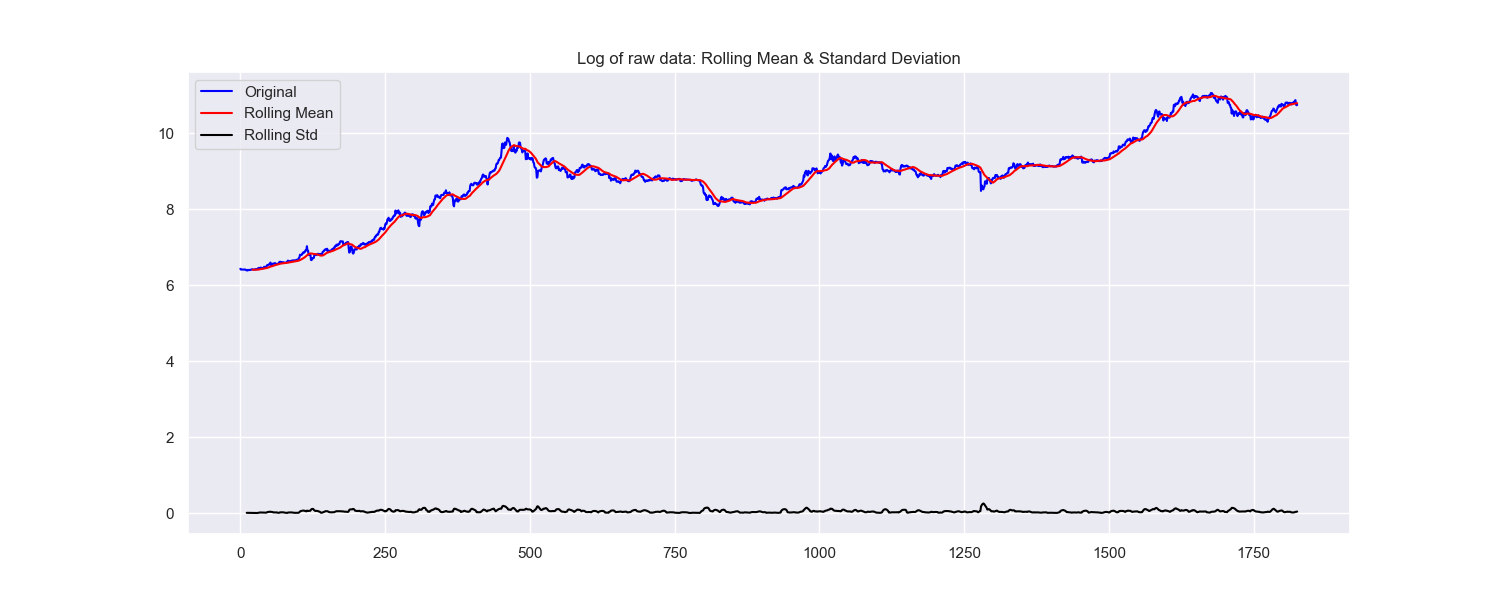
p - represents the number of lags (lags) of the time series data itself used in the forecasting model, also called AR/Auto-Regressive term

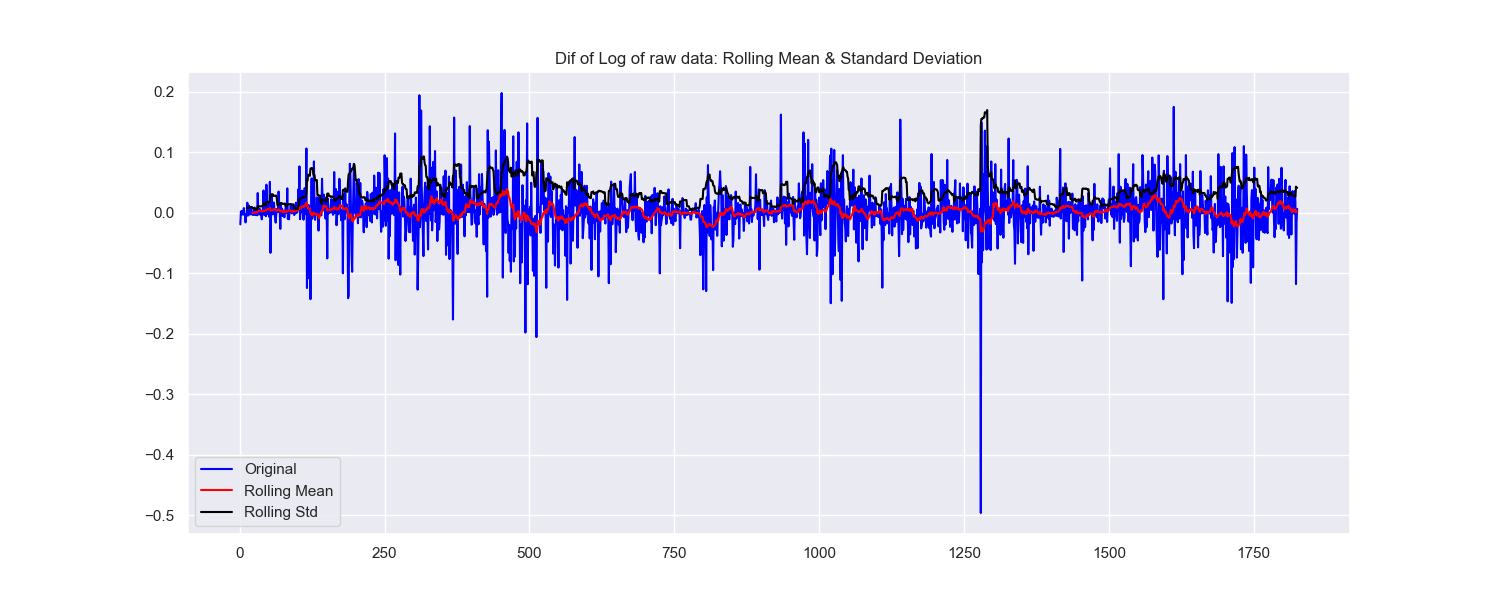
d - represents the number of orders of differencing needed for the time-series data to be stable, also called the Integrated term.

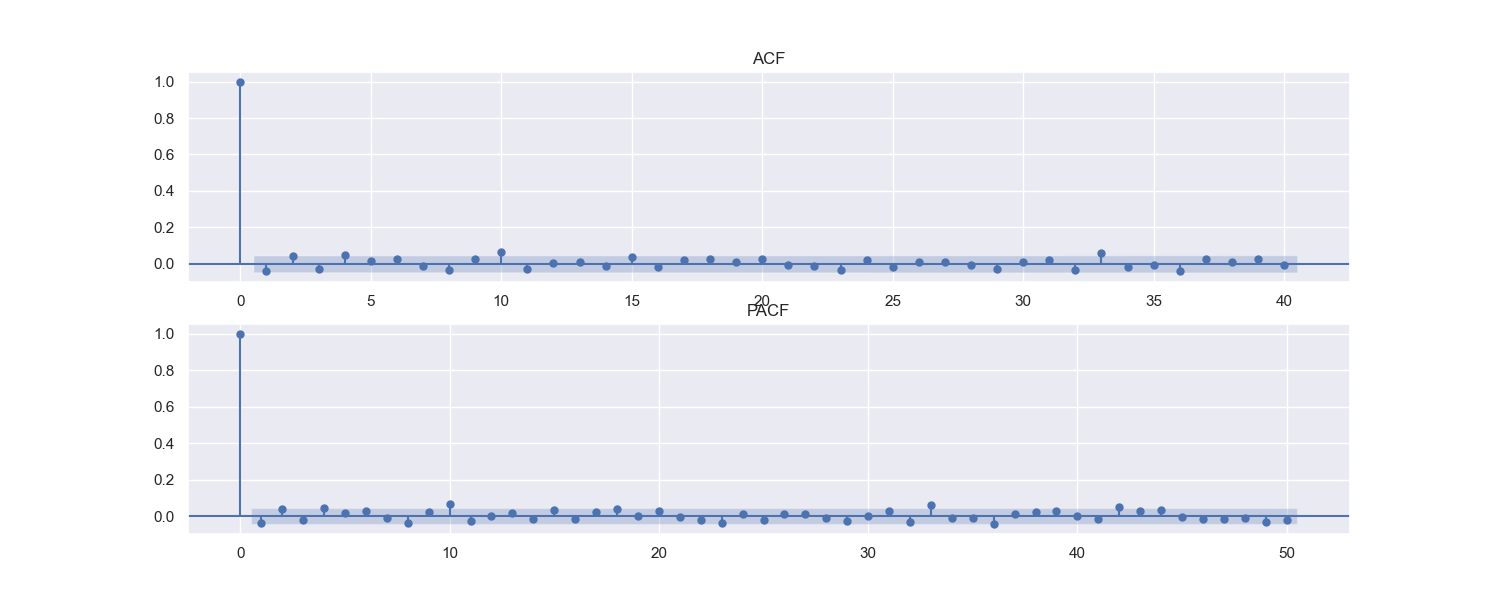
q - represents the number of lags of the prediction error used in the prediction model (lags), also called MA/Moving Average term

只有平稳的数据集才能用ARIMA模型预测，由图xx可以明显看出比特币和黄金的价格呈指数上升并不稳定，通过数据预处理部分的ADF和KPSS也可以佐证这一点。于是通过取对数和差分的方法将其稳定，如图xx和图xx。事实上，比特币价格走势在前半段和后半段有跫然不同的表现——前半段相对缓慢上升，后半段剧烈波动上升。所以如果我们使用单一的ARIMA模型不能适用于整个时间短。在不破坏原始数据的完整前提下，我们需要尝试寻找最佳的ARIMA模型，尽管整段时间的效果不会很契合。根据传统定阶数的方法——ACT和PACF分析如图xx，可以找到相对合适的阶数p和q，并且此操作需要在每次滑动窗口后重新进行。在此情况下，阶数p和q的改变对误差的影响很小【引用论文Bitcoin Price Prediction: An ARIMA Approach】，我们可以固定阶数p和q将ARIMA作为一种结构函数来使用。与此同时，在数据预处理部分，我们对原始数据的协方差很小证明了他的自相关性很弱，间接也说明了阶数的影响不大。

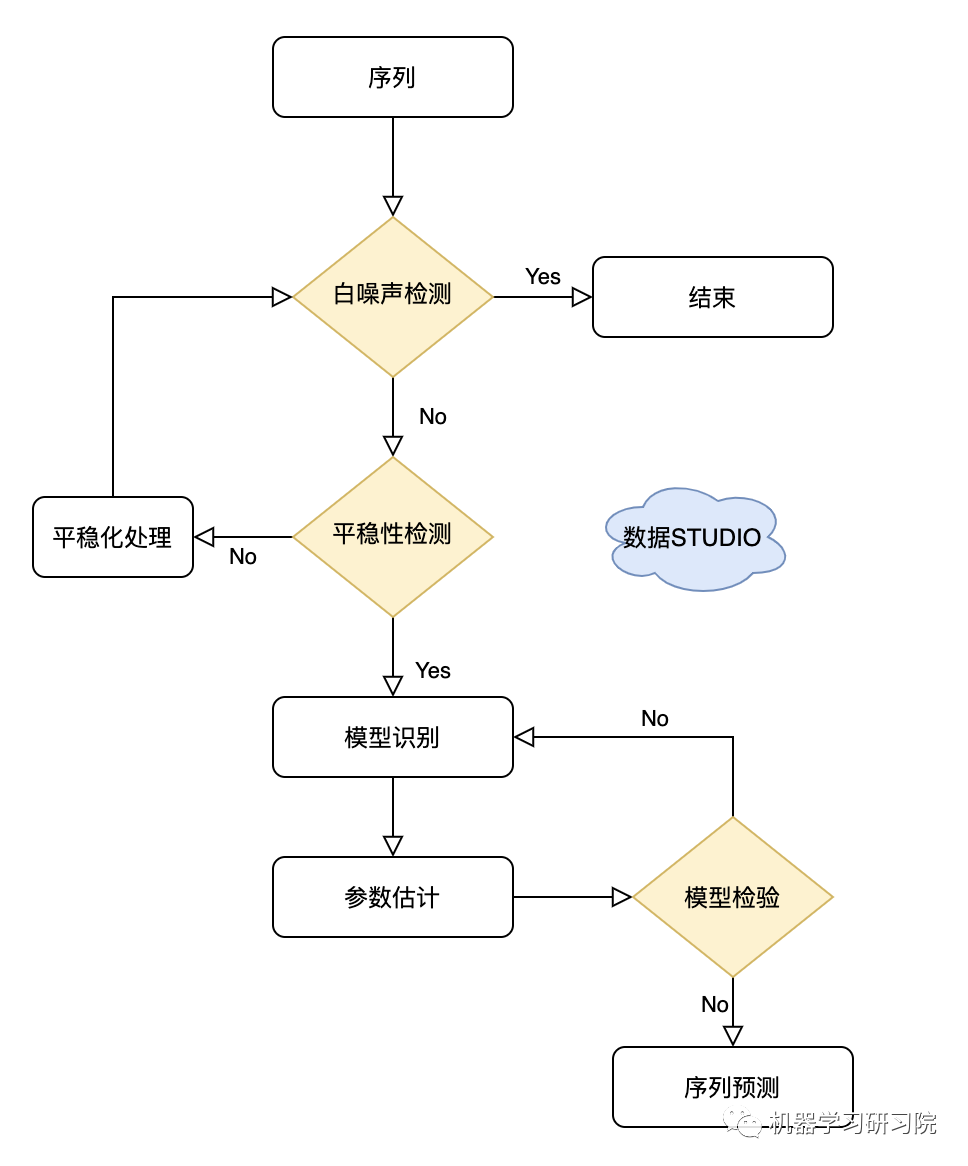


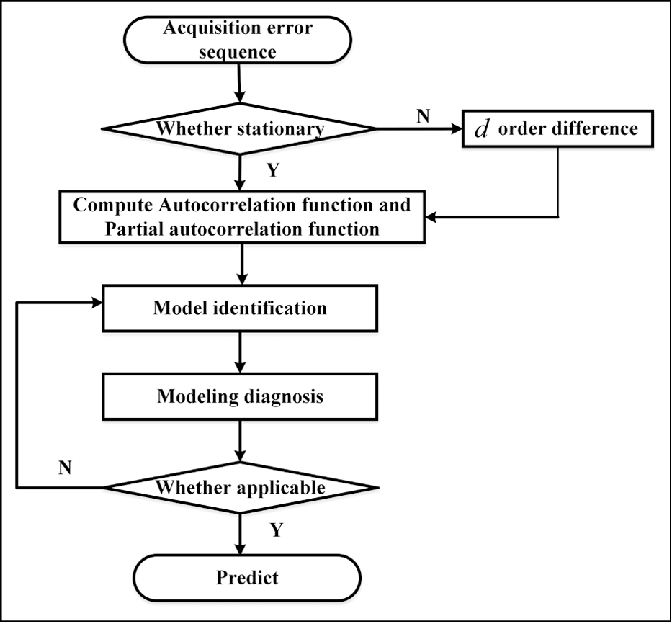






Only a smooth dataset can be predicted with the ARIMA model. It is clear from Figure xx that the exponentially rising prices of bitcoin and gold are not stable, and this is corroborated by the ADF and KPSS in the data preprocessing section. So it is stabilized by taking the logarithm and the difference, as in Figure xx and Figure xx. In fact, the bitcoin price trend has a soundly different performance in the first half and the second half - a relatively slow rise in the first half and a dramatically volatile rise in the second half. So if we use a single ARIMA model cannot be applied to the entire time short. Without destroying the integrity of the original data, we need to try to find the best ARIMA model, although the effect for the whole time period will not fit well. According to the traditional methods of fixing orders - ACT and PACF analysis as in Figure xx - relatively suitable orders p and q can be found, and this operation needs to be repeated after each sliding window. In this case, the change in order p and q has little effect on the error [citing the paper Bitcoin Price Prediction: An ARIMA Approach], and we can fix the order p and q to use ARIMA as a structural function. At the same time, in the data preprocessing section, the small covariance we have for the original data proves that his autocorrelation is weak, which indirectly also shows that the order has little effect.





# CVaR

<https://quantatrisk.com/2016/12/08/conditional-value-at-risk-normal-student-t-var-model-python/?ihc_login_fail=true&ihc_pending_email=true&ihc_success_login=true>

If represents the -day returns then where Keeping that in mind, the Conditional Value-at-Risk, expressed as a percentage of the portfolio value, is given by:

Therefore, in order to derive CVaR for any continuous probability density function of , we need to integrate over till -day VaR (i.e. quantile). Now, one can find that:

is the conditional value-at-risk CVaR in the normal linear VaR model for a random variable over -day horizon where denotes the standard normal density function and the quantile of the standard normal distribution.  
Since maths can be exciting, more excitement comes from a practical example. Suppose that we wish to find 5 -day CVaR for a stock characterised by its annual expected volatility of at . Therefore, the 5-day standard deviation we find as assuming 252 trading days in a calendar year and significance level.

数据预处理图  
