Team Control Number

201986768

Problem Chosen

C

2019

ShuWei Cup

Summary Sheet

Summary

This paper mainly studies the pork price fluctuations in China.

Part I: A. Firstly, identify the relevant variables according to the equilibrium theory of supply and demand, and then select the main influencing factors whose correlation coefficient with pork price is higher than 0.85. Next, to verify the reliability of these factors, the pork price is decomposed in time series to obtain trend series, season series and residual series, and these selected variables are used to fit the trend series by using multiple linear regression methods. Finally, we conclude that the main influencing factors of the pork price are: corn price, soybean meal price, pork storage, the previous price of pork the supply amount of beef and mutton. **B**. Use hypothesis testing to analyze the data, and it turns out that the factors influencing the recent fluctuations of pork prices are the outbreak of the African swine fever, poor transportation of pork and Sino-US trade friction, not the common fluctuations.

Part II: A. Firstly, a nonlinear, unbalanced cobweb model is established to predict demand at when pork is expensive. Secondly, given the cost factors such as cross-provincial transportation, we introduce a transport model based on the graph-related knowledge of the shortest path problem to make the model more practical. Finally, considering the constraints of demand and regional factors, a linear planning model with the lowest cost is established to predict the monthly pork farming plan. **B**. The total demand gap of the whole country can be calculated based on the monthly demand gap of the provinces, and after comprehensive analysis of historical data and the dependence of importing pork from other countries, a reasonable pork import proposal is given.

Part III: A. Firstly, the demand for pork in each province is recalculated because the time zones should be considered when the supply and demand model is established under the condition of the pig cycle. Secondly, a new linear planning model concluding the time factor is established to predict the future demand gap of pork of different time zones in different province. **B**. In the light of the actual background, a reasonable pork storage strategy should deal with the serious imbalance between supply and demand in a short time, to arrange pork storage points rationally, and to provide policy subsidies.

Key word: Fluctuations of Pork Prices, Time Series, Cobweb Model, LP Model

Content

1. Introduction...3
    1.1 Background...3
    1.2 Work...3
2. Problem analysis...3
    2.1 Data analysis...3
    2.2 Analysis of question one...4
    2.3 Analysis of question two...5
    2.4 Analysis of question three...5
3. Symbol and Assumptions...6
    3. 1 Symbol Description...6
    3.2 Fundamental assumptions...6
4. Model...7
    4.1 Non-stationary Time Series Model...7
    4.2 Statistical hypothesis testing...7
    4.3 Nonlinear non-equilibrium cobweb model...8
    4.4 Floyd algorithm...8
    4.5 Linear Programming...9
5. Test the Models...9
    5.1 Question one...9
    5.2 Question two...14
    5.3 Question three...19
6. Sensitivity Analysis...21
7.Strengths and Weakness...22
    7.1 Strengths...22
    7.2 Weakness...22
8.Conclusion...22
References...23
Appendix...24

1. Introduction

1.1 Background

Pork farming is a fundamental industry in China, whose stability affects the interests of all participants in the industrial chain such as fodder producers, farmers, operators and consumers. As the demand for pork increases, pork farming, which is related to China's political stability, economic development and the welfare of resident, has become the core industrial of China's meat production and consumption [1].

As the world's largest producer and consumer of pork, China's pig farming industry has its own characteristics. Its production and price variation are more volatile, because of its dependence on policies and it is easily influenced by external factors. Especially in the past six months, the price of pork in China has shown a three-fold growth trend. Therefore, it is of great significance to study the influencing factors of pork prices, explore the strategy to deal with the price fluctuations, and find out the optimal supply and storage strategy.

1.2 Work

Based on the above background, this article will address the following issues:

(1) Identify the main influencing factors of pork price fluctuation. On the one hand, we use the data from the past few years to analyze the common influencing factors of pork price fluctuations in the long run. On the other hand, in order to figure out whether the recent fluctuations in pork prices are highly correlated with the common fluctuations in pork prices, we use the data from the most recent year to analyze it. If not, we will find out the main factors influencing the recent pork price.

(2) Research on the fallback of pork prices. Considering that pork farming usually has a certain periodicity and some regions are impossible for pig farming, we need to propose a reasonable farming plan to achieve a fallback of pork price in the short term when the price of pork is high. Under the premise that pork farming cannot be completed in a short time, we should develop a reasonable procurement plan for other countries and ensure that domestic pork prices are relatively stable.

(3) The optimal supply and storage strategy of pork. When pork demand is relatively stable, propose the optimal farming plan for different time zones in different regions. Otherwise, when there is a peak demand for pork in a certain area, an effective pork storage strategy is proposed to ensure the stability of pork prices.

2. Problem analysis

2.1 Data analysis

This paper uses five types of data:
Crop data such as soybean meal price, corn price, fodder price, pig grain ratio…
Livestock data such as total cattle and sheep, chicken price, pork production, pork stock, pork price…

Population data such as urban population, rural population, CPI, rural CPI…
Import data such as pork export volume …
Geographic such as provincial capital distance …
The sources of these data are: National Bureau of Statistics[2], Pig price system[3], China Animal Husbandry[4], Bric Agricultural Data Terminal[5],
Agricultural Information Network [6], National Agricultural Products Cost-benefit Data Collection[7], China Business Intelligence Network[8]

2.2 Analysis of question one

2.2.1 Part A

Analyze the common influencing factors of pork price fluctuations by using data in the past years.

Firstly, we select alternative variables based on economic theory. Given that the equilibrium price of commodities is the result of demand and supply in the market, we assume that the factors affecting pork price mainly include three aspects--supply, demand and other external factors. Among them, the supply factors include the supply amount of pork, the previous price of pork, and the price of raw materials for fodder, etc. The demand factors include the consumption amount for pork, the income of residents, the population both in urban and rural area and the price of substitute, etc. The external factors include economic growth, macroeconomic regulation and control, public opinion, etc.

Secondly, we conduct a correlation analysis on the variables to identify the main influencing factors, and five variables were gained from it, including the supply factors--corn price, soybean meal price, pork storage, the previous price of pork and the demand factor--the supply amount of beef and mutton.

Finally, to verify the reliability of these factors, the pork price is decomposed in time series to obtain trend series, season series and residual series, and these selected variables are used to fit the trend series by using multiple linear regression methods. The results show that the model has a good effect:$R^2 = 0.94$, $MSE = 1.42$..

2.2.2 Part B

To determine whether the recent fluctuations in pork prices are highly correlated with the common fluctuations in pork prices, we use the statistical hypothesis testing to test whether they are within a normal range.

First, we propose a null hypothesis H_0: the recent fluctuations in pork prices are highly correlated with the common fluctuations in pork prices.

Second, we propose the method for hypothesis testing. According to the time series model established in Q1 Part A, the residual sequence obeys the Normal distribution, then the high correlation means that the difference between the pork price and the estimated value in recent years should within a certain probability interval $[\mu - 3\sigma, \mu + 3\sigma]$. If the residual exceeds this reasonable range, it means that the recent fluctuations in pork prices are not highly correlated with the common fluctuations in pork prices.

Then we conduct an empirical analysis. We collect data from 2016 to 2019 to predict the recent pork price based on the model established in Q1 Part A, and compare it with the actual values. It turns out that the residual from 2018.3 to 2018.5 and from 2019.3 to 2019.9 exceeds the 3σ range, indicating that the recent fluctuations in pork prices do not match the common fluctuations, that is, they are not highly correlated.

Finally, in order to determine the main influencing factors of the recent price of pork, we review the relevant information[9] and find out the reasons for the recent impact on pork prices, such as environmental demolition, global outbreaks (African swine fever), and Sino-US trade friction.

2.3 Analysis of question two

2.3.1 Part A

We need to propose a pork farming plan when pork prices are high. The reason for the high price of pork is that the pork is in short supply. Therefore, in order to achieve a fallback in pork prices in the short term, we should propose a reasonable planting plan to bring the market to a balance between supply and demand.

First, given the price is high and it is in not a state of equilibrium, we establish a nonlinear unbalanced cobweb model to predict the demand at this time.

Secondly, during the pork farming programs, in addition to the rational use of local aquaculture resources, the resources of other province should also be included in the consideration. So we introduced a transportation model to measure costs and make the model as realistic as possible. That is, when considering transporting pork from remote areas, the transportation cost needs to be within a reasonable range and as small as possible. In order to obtain the minimum transportation cost, this paper uses the relevant knowledge of graph theory of the shortest path problem under the premise of satisfying the supply of pork.

Finally, based on the demand and regional factors, a linear programming model with the lowest total cost is established, and the monthly pig production in each province and the recharge plan between the provinces are calculated.

2.3.2 Part B

Based on the monthly demand gap of each province, we calculate the national monthly total demand gap, and comprehensively consider the international market share of imported pork, then analyze China's import plan.

2.4 Analysis of question three

2.4.1 Part A

The research background of this question is to consider the impact of the pig cycle on supply under the condition of stable demand, and then propose the best farming plan for time zones in different regions.

Firstly, we introduce the time dimension, divide the year into four quarters under the condition of considering the pig cycle, and establish a supply and demand model to recalculate the pork demand of each province in different quarters.

Secondly, a new linear planning model concluding the time factor is established to predict the future demand gap of pork of different time zones in different province. We should notice that the total amount of pig breeding in each adjacent quarter of each province should be limited to no more than the local breeding limit.

Finally, calculate the demand gap for the pigs in different time zones in each province and the inter-provincial replenishment plan.

2.4.2 Part B

Based on the actual background, we put forward a pork storage strategy: timely responding to serious imbalances in supply and demand, rational arrangement of pork storage points, and policy subsidies.

3. Symbol and Assumptions

3. 1 Symbol Description

Symbol	Meaning
Y	pork price
X_1	corn price
X_2	soybean meal price
X_3	pork storage
X_4	the supply amount of beef and mutton
X_5	the previous price of pork
D_t	the demand of pork
S_t	the supply of pork
P_t	the price of pork
$\omega_{i,j}$	the distance between province i and province j
δ	tolerance parameter
m_i	the maximum quantity of pigs in province i
s_i	the increment in demand in province i
T_b	pig breeding cycle

3.2 Fundamental assumptions

1. The pig farmers are rational people, they will make relevant decisions based entirely on the models provided in this paper.

2. When the supply and demand do not reach equilibrium, the pork price also cannot reach the equilibrium.

3. The transportation cost is proportional to the length of the transportation route and the number of pigs transported.

4. The international market price is stable and will not be affected by the supply and demand condition in China.

5. Do not consider the impact of “prohibition of inter-provincial transportation” and other related policies.

4. Model

4.1 Non-stationary Time Series Model

The Non-stationary Time Series is: When the random disturbances are removed from the data, the remainder can be represented by a deterministic time function. That is, if Y represents a certain time series, the sequence can be decomposed into the following parts:

$$Y = f(T, C, S, I)$$

Among them,

T: Trend, the change of the time series in the long-term

C: Circle, periodic changes over one year

S: Season, periodic changes within one year

I: Immediate, the influence of unpredictable accidental factors on the time series

A time series may include all or part of the above four parts. According to this question, since the deviation of the trend increases with time, we use the multiplication model (combine T and C variables):

$$x_t = T_t \cdot S_t \cdot I_t$$

Considering that the long-term trend of the sequence exhibits linear features, we use a linear model to fit, and the model can be written as:

$$T = \beta_0 + \beta_1 X_1 + \cdots + \beta_5 X_5 + \varepsilon$$

Among them, ε is random fluctuations.

Then determine the coefficients of a and b according to the principle of least squares:

$$Q = \sum_{i=1}^{n} e^2 = \sum_{i=1}^{n} (Y_i - \widehat{Y_i})^2 = \sum_{i=1}^{n} [Y_i - (\widehat{\beta_0} + \widehat{\beta_1} X_{i1} + \widehat{\beta_1} X_{i2} + \cdots + \widehat{\beta_k} X_{ik})]^2$$

4.2 Statistical hypothesis testing

Hypothesis testing is a statistical inference method used to determine whether the differences between samples and samples, between samples and population are caused by sampling errors or essential differences. Significance test is the most commonly used method and the most basic form of statistical inference in hypothesis test. Its basic principle is to make a hypothesis on the characteristics of the population firstly, and then make a deduction on whether the hypothesis should be rejected or accepted through the statistical inference of sampling research. The hypothesis test method chosen in this paper is Z test.

4.3 Nonlinear non-equilibrium cobweb model

The traditional cobweb model has two basic assumptions: 1 the supply function and the demand function are linear functions; 2 the supply and demand equilibrium for each period, on the basis of which the algebraic expression is:

$$\begin{cases} D_t = a + bp_t \\ S_t = a' + b'p_{t-1} \\ D_t = S_t \end{cases}$$

"Equilibrium of supply and demand in each period" is the most ideal and simplest theoretical assumption for the convenience of research. Based on the purpose of expressing the relationship between pork supply and demand in the real pork market, this paper explores the expression of pork supply and demand relationship function. The price change directly reflects the change in supply and demand. The current price is affected by the difference in supply and demand in the current market, and is also affected by the price of the previous period. This effect has the direct impact of the consumer market on the current price expectation brought by the previous period price. There is also the indirect impact of the previous period price affecting the current supply, which in turn affects the current balance of supply and demand.

Lingl.D. extended the traditional cobweb model, established a nonlinear non-equilibrium cobweb model of the following form

$$\begin{cases} D_t = -ap_t^2 + bp_t + c \\ S_t = \alpha - \dfrac{\beta}{p_{t-1}} \\ p_t = p_{t-1} + \gamma(D_t - S_t) \\ a, b, c, \alpha, \beta, \gamma > 0 \end{cases}$$

According to the data obtained, the demand for pork is matched with the current pork price using the polynomial to obtain the demand function, and then the function relationship between the previous price and the supply is determined and the supply function is obtained. Finally, the pork price cobweb model is established.

4.4 Floyd algorithm

Floyd algorithm, also known as the interpolation method, is an algorithm that uses the idea of dynamic programming to find the shortest path between multi-source points in a given weighted graph.

Above all, the weighted adjacency matrix A is taken as the initial value of distance matrix D, that is, $D^{(0)} = A$.

$D^k = (d_{ij}{}^{(k)})_{m\times n}$, element $d_{ij}^{(k)} = \min\{d_{ij}^{(k-1)}, d_{ik}^{(k-1)} + d_{kj}^{(k-1)}\}$ is the shortest path length from the path $v_1, v_2, \cdots, v_k$.

So, D^n is the distance matrix. In the iterative process of finding the distance matrix, the matrix R of subsequent nodes can also be found out, so that the shortest path between any two points can be found out from R.

4.5 Linear Programming

Linear Programming (LP) is an important branch of mathematical programming in Operations Research. In solving practical problems, it is necessary to reduce the problem to a linear programming mathematical model. The key and difficulty lies in choosing the appropriate decision variables to establish the appropriate model.

The objective function and constraints of linear programming are all linear functions, and the constraints are recorded as S.T. (subject to). The objective function can be the maximum value or the minimum value, and the inequality sign of constraint condition can be less than sign or greater than sign.

The mathematical standard form of general linear programming problem is

$$max \quad z = \sum_{j=1}^{n} c_j x_j$$

$$s.t. \begin{cases} \sum_{j=1}^{n} a_{ij} x_j = b_j & i = 1,2, \dots, m \\ x_j \geq 0 & j = 1,2, \dots, m \end{cases}$$

5. Test the Models

5.1 Question one

5.1.1 Part A Establish a time series model of pork prices

Step 1: Variables are selected according to economic theory

According to the relevant theory of economics, the price of commodities is the result of the combined effects of demand and supply on the market. When supply and demand are balanced, the price is in equilibrium. Therefore, we propose that the factors affecting pork prices mainly include three aspects, supply factors, demand factors and other external factors.

Among them, the supply factors include the supply amount of pork, the previous price of pork, and the price of raw materials for fodder, etc. The demand factors include the consumption amount for pork, the income of residents, the population both in urban and rural area and the price of substitute, etc. The external factors include economic growth, macroeconomic regulation and control, public opinion, etc. (As shown below)

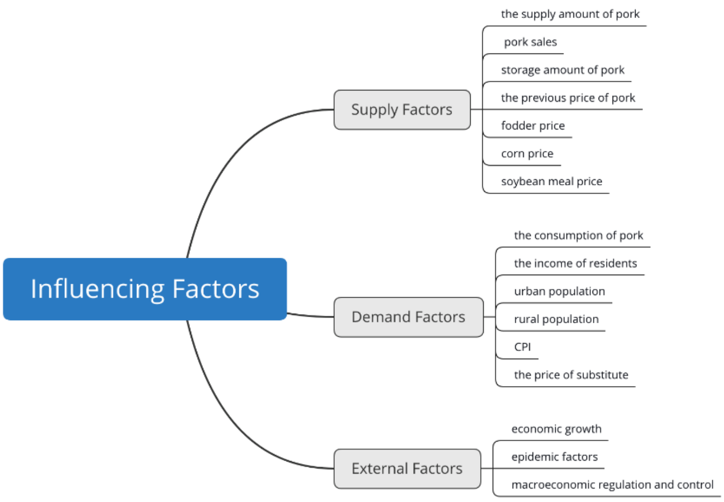

Fig 1. Factors Affecting Pork Prices

In order to initially select the main influencing factors, we conducted a correlation analysis on the variables, and the results are shown in the following table:

TABLE 1. CORRELATION COEFFICIENT OF INFLUENCING FACTORS

	Pork price	*Soybean price*	*Corn price*	*Feed price*	*Pork price/grain price*	*Total beef and mutton*	*Chicken price*	*Pork output*	*Pork storage*	*Urban population*	*Rural population*	*Urban and rural population ratio*	*City CPI*	*Rural CPI*
Pork price	1.00	0.87	0.91	0.92	0.77	0.88	0.96	0.79	0.66	0.86	-0.87	0.87	0.64	0.64
Soybean price	0.87	1.00	0.82	0.91	0.65	0.90	0.85	0.66	0.70	0.86	-0.86	0.85	0.48	0.46
Corn price	0.91	0.82	1.00	0.97	0.46	0.92	0.96	0.89	0.80	0.94	-0.95	0.95	0.39	0.39
Feed price	0.92	0.91	0.97	1.00	0.53	0.97	0.96	0.85	0.84	0.95	-0.96	0.96	0.40	0.39
Pork price/grain price	0.77	0.65	0.46	0.53	1.00	0.50	0.62	0.34	0.21	0.44	-0.44	0.44	0.79	0.78
Total beef and mutton	0.88	0.90	0.92	0.97	0.50	1.00	0.92	0.79	0.88	0.92	-0.92	0.93	0.29	0.28
Chicken price	0.96	0.85	0.96	0.96	0.62	0.92	1.00	0.87	0.78	0.92	-0.93	0.94	0.46	0.46
Pork output	0.79	0.66	0.89	0.85	0.34	0.79	0.87	1.00	0.74	0.91	-0.92	0.92	0.21	0.22
Pork storage	0.66	0.70	0.80	0.84	0.21	0.88	0.78	0.74	1.00	0.79	-0.81	0.82	-0.08	-0.09
Urban population	0.86	0.86	0.94	0.95	0.44	0.92	0.92	0.91	0.79	1.00	-1.00	1.00	0.32	0.31
Rural population	-0.87	-0.86	-0.95	-0.96	-0.44	-0.92	-0.93	-0.92	-0.81	-1.00	1.00	-1.00	-0.32	-0.31
Urban and rural population ratio	0.87	0.85	0.95	0.96	0.44	0.93	0.94	0.92	0.82	1.00	-1.00	1.00	0.31	0.30
City CPI	0.64	0.48	0.39	0.40	0.79	0.29	0.46	0.21	-0.08	0.32	-0.32	0.31	1.00	1.00
Rural CPI	0.64	0.46	0.39	0.39	0.78	0.28	0.46	0.22	-0.09	0.31	-0.31	0.30	1.00	1.00

Considering the problems of multicollinearity, heteroscedasticity and other problems, we finally selected five indicators with a correlation coefficient of pork price Y greater than 0.85 as the independent variables of the model:

Supply factors: corn price X_1, soybean meal price X_2, pork storage X_3, the previous price of pork X_5.

Demand factor: the supply amount of beef and mutton X_4 (replacement output).

Step 2: Time series decomposition

According to the data [11], we know that there is a certain period of pork price. Therefore, we use the 2000-2011 monthly data to identify its cycle and find that pork prices have a traditional cyclical change: in general, pork prices begin to decline in January each year, and increase after April. After that, pork prices will show short-term fluctuations from September.

Fig 2. The Traditional Periodic Change of Pork Price over the Years

We select 12 months as the cycle length to decompose the pork price in time series to obtain the trend series, season series and residual series. The decomposition results are shown in the following figure:

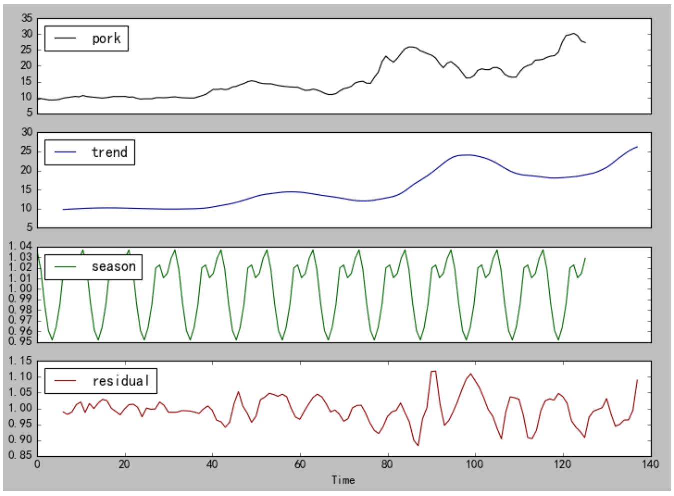

Fig 3. Result of Pork Price Series Decomposition

Using the least squares estimation method to fit the long-term trend of the sequence(70% of the data is selected as the training set, and 30% of the data is used as the test set), we can find that:

The trend series shows that pork prices generally show an increasing trend.

The season series is the extracted pig cycle. Observing the cycle change, the pork prices will start to decline in January each year, and will start to rise in April. From September, pork prices will fluctuate briefly, which is correspond to the reality.

The residual series is the residual between the predicted price and the real price.

Step3: Trend fitting, determining coefficients

Using the least squares method, get the regression equation of Trend:

$$\text{Trend} = 17.3909 + 8.9395X_1 + 0.8642\text{X}_2 - 6.0132\text{X}_3 + 23.37\text{X}_4 + 0.4333X_5$$

According to the evaluation results of the model, as shown:

$$R^2 = 0.939981$$

$$MSE = 1.420708$$

$$RMSE = 1.191934$$

The effect coefficients for the Season series is calculated

TABLE 2. SEASON EFFECT COEFFICIENT

Month	Jan.	Feb.	Mar.	Apr.	May.	June.
Periodic coefficient	1.0368	1.0189	0.9859	0.9606	0.9517	0.9635
Month	July.	Aug.	Sept.	Oct.	Nov.	Dec.
Periodic coefficient	0.9855	1.0199	1.0226	1.0106	1.0145	1.028

Step 4: Evaluating model

Bring the data of Trend sequence, Season sequence and residual sequence into the formula $X_t = T_t \cdot S_t \cdot T_t$ to obtain the final result.

The figure below is the fitting effect of the least squares estimation method, combined with the fitted image of Season's seasonal effect:

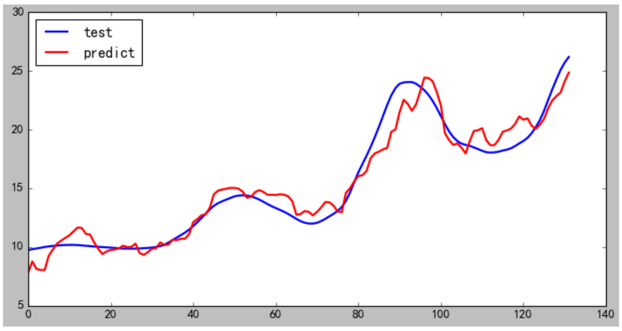

Fig 4. The Fitting Effect of the LSE Method

It can be seen that the increase in pork prices is highly correlated with the price of corn and soybean meal, the sales of pork, the total consumption of beef and mutton, and the previous price of pork, which can be used as the main factors to predict the price of pork as follows:

The price of corn and soybean meal per catty increased by 1 yuan, which will increase the price of pork per catty by about 8.94 and 0.86 yuan respectively.

Every additional 100 million pigs in storage will lead to a drop in the price per catty of pork by about 6.01 yuan. And for every 100 million tons of beef and sheep produced, the price of pork will increase by 23.37 yuan.

The previous price of pork can also have a certain impact on the price of this year. When the previous price of pork rose by 1 yuan, the price of pork at this time rose by 0.43 yuan.

5.1.2 Part B

Step1: Put forward the null hypothesis

To analyze whether the recent fluctuations of pork price are highly correlated with the common fluctuations, it is essentially to determine whether the fluctuation of pork prices are within the normal range.

According to the model fitted with historical data in Part Ⅰ-A, the ratio of the predicted pork price to the real pork price is appropriately obeyed the normal distribution with a mean of 1.0 and a variance of 0.042. For every moment, the prediction bias of the model will have a great probability of falling around 1.0.

Then we propose the null hypothesis (H_0): the distribution of prediction bias is the same as the common bias.

For a given prediction bias, first standardize it using formula:

$$e_0^* = \frac{e_0 - m}{s}$$

Then calculate the probability:

$$p = P(|e| \geq |e_0^*|) = 2 \times \int_{e_0^*}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2} dx$$

If the p-value calculated at a moment is less than 0.01, we can reject the null hypothesis (H0), which leads to the conclusion that the prediction bias is abnormal and, that is, the recent fluctuations are not highly correlated with the common fluctuations in pork prices. The figure below shows the prediction of pork prices in recent years.

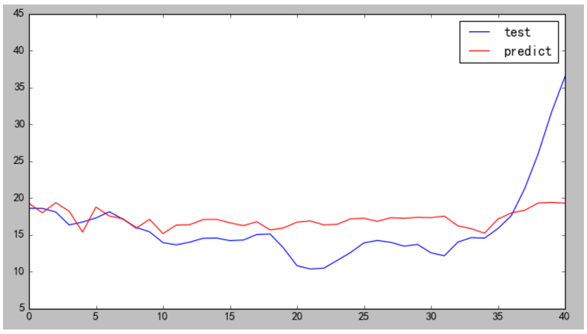

Fig 5. Forecast Results from June 2016 to November 2019

Step 2: Analyze the main influencing factors of recent pork price volatility

The price of September-November 2019 is higher than the range of 3σ, which is due to:

i. Outbreak of African swine fever

In August 2018, the outbreak of the African swine fever epidemic had a significant impact on the domestic pig breeding industry and the pork market. Due to the pig breeding cycle, the farmers did not dare or unable to fill the bar, which weakened the domestic medium and long-term supply capacity. As a result, domestic pork supply has shown a large decline since 2019, and the price has risen considerably.

ii. Poor transportation in the pork production and marketing area

China's pork production areas and consumption areas are relatively clear. The main mode of pork circulation is the transportation of live pigs. The poor transportation of pork production and sales areas limits the ability to adjust the gap between different regions of the country.

Once the circulation of pork products between the above regions is blocked, it will lead an important impact for market supply. After the outbreak, in order to prevent the spread of the fever, the Ministry of Agriculture and Rural Affairs temporarily closed all the live pig trading markets in the province. The pigs in each production area were slaughtered on the spot, transported to the sales area through cold chain logistics, and the “tune up pig” was changed to “tune the meat” to break the original logistics balance.

iii. Sino-US trade friction inhibits pork import regulation expectations

In order to counteract the United States' tariff on imported goods from China, the Ministry of Commerce imposed a 25% tariff on pork and products imported from the United States in March and August 2018 respectively. Since the implementation of counter-measures on pork and products imported from the United States, the tax rate has reached 62%. This has limited the market expectation of pork import regulation to a certain extent. Although this has not had a major impact on domestic pork supply in 2018, it will increase market expectations of insufficient domestic supply in the context of a sharp drop in domestic pork supply in 2019.

5.2 Question two

5.2.1 Part A

In the analysis of the first question, we considered how to find the main factors affecting the price fluctuations of pork in the case of balance between supply and demand. However, after the recent predictions did not match the changes in the indicators, we found through investigation that it is due to the imbalance between supply and demand.
Therefore, in question two, we need to establish a nonlinear non-equilibrium model to predict the demand when the pork price is high in order to propose the corresponding framing plan. We consider the inter-regional transportation as an optimization link in the planting plan, and establish a transportation model to solve the problem that some areas cannot be transported and the use of remote areas to reduce costs.
Step1： Establish a nonlinear non-equilibrium cobweb model to predict demand
According to the nonlinear non-equilibrium cobweb model we proposed in 4.3 and the real data, the following form has been established:

$$\begin{cases} D_t = -6.843 \times 10^{-5} p_t^{\,2} + 2.414 p_t + 26616.021 \\ S_t = 50628.798 - \dfrac{6.288 \times 10^7}{p_{t-1}} \\ p_t = p_{t-1} + \gamma\left(D_t - S_t\right) \\ \gamma > 0 \end{cases}$$

Step2: Establish a transportation model to measure costs

Based on the supply and demand model (cobweb model), it may happen that some regions need transportation replenishment in other regions. At this time, a transportation model needs to be established to consider the cost and geographical limitation.

We use the shortest path model in the graph theory to solve the problem. The mapping process is:

We use the provincial capital of each province as the apex of the map, connect an edge between the provincial capitals of the neighboring provinces, and calculate the distance between the provincial capitals as the edge weight through the map, and then obtain the adjacency matrix of the graph. Then use the Floyd algorithm to find the shortest path on any of the two provinces. And use its shortest path as an indicator to measure its transportation costs.

Fig 6. The Graph Structure of Chinese Map

Step3: Establish a linear programming model to predict the demand gap

When the local pork supply cannot meet the demand, we must consider raising pigs in some remote provinces and then transporting them. Considering the geographical location (freight), the number of pig farm facilities (the upper limit of the number of acceptable pig breeding), the proportion of rural population, climatic conditions (whether it can be farmed), religious beliefs, etc., we need to develop a reasonable pork farming plan for each province in the country.

We use a linear programming model to determine the optimal plan:

Constraints: For each province, the total amount of increased pig breeding should not exceed the local breeding limit, and the total amount of pigs received from other provinces should be close to the difference between supply and demand.

$$\begin{cases} \sum_k x_{i,k} \le m_i \\ \sum_k x_{k,j} \le (1+\delta) * s_j \\ \sum_k x_{k,j} \ge (1-\delta) * s_j \\ x_i \ge 0 \end{cases}$$

Objective function: Under the premise of satisfying the above constraints, the total freight rate is the lowest.

$$\min z = \sum_i \sum_j \omega_{i,j} * x_{i,j}$$

Step4: Make a plan

Using the non-linear cobweb model to give the pork price supply and demand relationship, calculate the difference between the supply and demand of pork in each province and city in October 2019, and use the transportation model to calculate the length of the shortest path from province i to province j. Then, according to the historical supply, the largest pig breeding amount in each province is obtained and substituted into the linear programming model. The results are as follows:

Among them, the positive external recharge amount indicates outward transportation, and the negative number indicates foreign import. The amount of locally planned aquaculture is the total amount after taking into account local actual demand and external recharge, in units of 10,000 tons.

TABLE 3 MONTHLY PORK FARMING PLAN OF EACH PROVINCE

Province	*Local actual demand*	*External supply*	*Local planned breeding quantity*
Anhui	5.230	0.373	5.602
Beijing	1.092	-0.748	0.345
Chongqing	4.436	0.044	4.480
Fujian	4.324	0.043	4.367
Gansu	1.167	-0.012	1.156
Guangdong	14.837	-3.860	10.977
Guangxi	5.964	2.092	8.055
Guizhou	3.892	-0.039	3.853
Hainan	0.559	1.674	2.234
Hebei	4.138	0.832	4.970
Henan	4.986	-0.050	4.936
Heilongjiang	2.543	-0.025	2.517
Hubei	5.062	-0.051	5.011
Hunan	7.527	-0.075	7.452
Jilin	1.921	-0.019	1.902
Jiangsu	6.790	4.885	11.675
Jiangxi	4.344	-0.043	4.300
Liaoning	3.901	-0.039	3.862
Inner Mongolia	1.775	-0.018	1.757
Ningxia	0.162	-0.002	0.160
Qinghai	0.000	0.000	0.000
Shandong	5.653	-0.057	5.596
Shanxi	1.574	-0.016	1.558
Shaanxi	1.530	-0.015	1.514
Shanghai	1.998	-1.998	0.000
Sichuan	12.368	-0.124	12.244

Tianjin	0.930	-0.146	0.784
Tibet	0.000	0.000	0.000
Xinjiang	0.377	-0.004	0.373
Yunnan	3.851	-0.039	3.813
Zhejiang	3.779	-3.438	0.341

According to the above table, we can divide the provinces of our country into the following three categories according to whether they can be self-sufficient:

1. Provinces that cannot solve the demand for pork by raising pigs in the province need to rely on the province to transport a certain amount of pork. These provinces mainly include Beijing, Tianjin, Shanghai, Zhejiang, and Guangdong.

2. Other provinces that are self-sufficient will deliver a certain amount of pork to the provinces, including Inner Mongolia, Heilongjiang, Xinjiang, Hebei, Shandong, and Guangxi.

3. Qinghai Province and Tibet Province were not affected by this pork price volatility because of the special geographical location and religious beliefs of the two provinces, resulting in lower demand for pork.

This results in a transport replenishment schedule between provinces:

TABLE 4 INTER PROVINCE TRANSPORTATION SUPPLY PLAN

Exporting provinces	*Transported provinces*	*Transport volume(Unit: 10000 tons)*
Jiangsu	Zhejiang	2.975
Guangxi	Guangdong	2.032
Jiangsu	Shanghai	1.978
Hainan	Guangdong	1.679
Hebei	Beijing	0.737
Anhui	Zhejiang	0.425
Hebei	Tianjin	0.137

According to the two schedules, the provinces can plan the actual aquaculture volume and the external transport (import) volume of the province.

5.2.2 Part B

In the past 10 years, China has been a net importer of pork products, and its import and export has shown a significant deficit. [12] Taking 2016 as an example, the import and export deficit of Chinese pork products was 2.949 million tons, of which the import volume reached 3.096 million tons and the export volume was 147,000 tons.

According to the data, the breeding cycle of pigs usually has a certain periodicity, which should be attributed to its biological mechanism, that is, the complete production cycle needs to go through the breeding sow, farrowing and fattening stages, about 6 month. That is to say, from the perceived high price of pork to the next six months, in order to ensure the stability of domestic pork prices, it is necessary to consider importing pork from abroad.

According to the available data, this paper calculates the pork demand gap in each province in each month, indicating the province's demand gap, the pork gap in each province is getting the pork gap in the country every month, and the monthly pork gap is multiplied by the breeding cycle. Get the total demand gap, the total import volume

$$A_{import} = T_b \times D = T_b \cdot \sum_{i=1}^{n} D_i$$

According to the monthly data of the provinces in part A, the national monthly demand gap D=116.7 tons can be calculated, and the total import volume is 700.2 tons.

When importing pork, in addition to the domestic demand gap, we should consider the real factors. On the one hand, we must give priority to the inspection of imported pork products and increase the customs clearance rate at the port. In addition, it is also necessary to establish a strict registration system for overseas production enterprises. It is not possible to blindly import pork in order to fill the demand gap and allow unqualified pork to flow into the country.

In Canada, for example, data shows that Canada's pork exports to China increased by 80% in March, but the good times are not long. Because China has repeatedly tested ractopamine (commonly known as "clenbuterol") from pork products imported from Canadian companies. Since June, Canadian pork has been stopped by China. [13]

As far as historical data is concerned, China's pork is mainly imported from Germany, Spain, the United States, Canada, Denmark and the Netherlands. Among them, the market share of pork imported from the United States reached 53.98% in 2011, which is the highest market share of pork imports from the United States in the past decade. It has continued to decline since then, only 13.62% in 2017. In 2017, Spain surpassed the United States to become the country's largest source of imported pork. The Danish market share has also been decreasing in recent years, falling to 7.29% in 2017; China has been importing pork from the Netherlands since 2011, the Dutch market share has increased in recent years, reaching 7.1% in 2017. [14]

TABLE 5 CHINA'S PORK IMPORT SOURCE COUNTRY AND MARKET SHARE CHANGE

Country of origin	*2011*	*2014*	*2017*
U.S.A	53.98	20.76	13.62
Spain	9.69	16.2	19.52
Germany	5.25	18.92	17.4
Denmark	12.86	11.98	7.29
Canada	10.59	9.24	13.7

Note: data from UN COMTRADE database

An analysis of the degree of dependence on the international market of pork-derived countries is a more effective means of discovering the country of origin and potential source of pork imports in the future. From the perspective of the export output elasticity of the country of origin of pork imports, Germany and Spain have a high degree of dependence. China should import pork from these countries, which is conducive to reducing the risk of pork imports. [15]

5.3 Question three

5.3.1 Part A

Step1: Basic definition

First of all, according to the results of Part I, the four quarters of the year correspond to the decline, recovery, short-term fluctuations and rising stages of pork demand. So this part needs to further consider the farming plan for different time zones in different province .

Step2: Establish a linear programming model

Considering the inclusion of the time dimension in this question, it is necessary to improve the linear programming model in combination with the pig cycle calculated in Part I.

Constraints: For each province, the total amount of pig breeding in the adjacent two quarters should not exceed the local breeding limit. In each quarter, the total number of pigs transported from the province to the province should be close to the current season’s demand.

$$\begin{cases} \sum_k x_{t,i,k} + \sum_k x_{t+1,i,k} \le m_i \\ \sum_k x_{t,k,j} \le (1+\delta) * Q_{t+2,j} \\ \sum_k x_{t,k,j} \ge (1-\delta) * Q_{t+2,j} \\ x_{t,i,j} \ge 0 \end{cases}$$

Objective function: Under the premise of satisfying the above constraints, the total freight rate is the lowest.

$$\min z = \sum_t \sum_i \sum_j \omega_{i,j} * x_{t,i,j}$$

symbol:

Set the number of live pigs transported from province i to province j at the time t, indicating the length of the shortest path from province i to province j, indicating the upper limit of the number of pigs in province i, indicating that province i needs to increase in the t quarter The supply of pigs, indicating the tolerance coefficient, is taken as 0.02.

Step3: Operation result：

TABLE 6 QUARTERLY PORK PRODUCTION OF PROVINCES (UNIT: 10000 TONS)

Province	*The first quarter*	*The second quarter*	*The third quarter*	*The fourth quarter*
Anhui	25.36	25.96	33.74	27.99
Beijing	3.55	6.05	3.55	6.05
Chongqing	22.4	23.75	30.81	25.51
Fujian	23.74	22.65	27.91	24.96
Gansu	6.93	6.81	8.86	7.18
Guangdong	72.72	66.28	72.72	66.28
Guangxi	30.15	30.47	37.16	33.08
Guizhou	22.66	22.24	27.63	24.51
Hainan	5.48	5.41	6.29	5.42
Hebei	20.75	21.55	27.42	22.24

Henan	23.33	23.7	30.88	25.13
Heilongjiang	13.47	13.98	16.56	13.87
Hubei	26.89	29.08	34.83	29.73
Hunan	40.82	42.23	52.63	45.5
Jilin	9.8	10.22	12.67	10.25
Jiangsu	34.7	35.35	44.36	38.44
Jiangxi	21.28	21.65	26.61	23.16
Liaoning	18.37	19.2	23.1	19.45
Inner Mongolia	10	9.67	12.34	10.79
Ningxia	0.96	0.97	1.21	1.03
Qinghai	1.62	1.43	1.84	1.85
Shandong	30.41	29.21	36.83	32.75
Shanxi	8.03	7.88	9.91	8.75
Shaanxi	8.97	8.91	11.18	9.41
Shanghai	7.3	0	7.3	0
Sichuan	64.32	68.87	86.21	73.7
Tianjin	5.59	5.53	5.59	5.71
Tibet	0	0	0	0
Xinjiang	2.11	2.15	2.63	2.28
Yunnan	28.24	29.07	36.45	32.01
Zhejiang	17.8	23.85	17.8	23.85

Main inter provincial transport relations:

TABLE 7 THE FIRST QUARTER

Exporting provinces	*Transported provinces*	*Transport volume(Unit: 10000 tons)*
Jiangsu	Zhejiang	9.62
Jiangsu	Shanghai	3.23
Hebei	Beijing	2.65

TABLE 8 THE SECOND QUARTER

Exporting provinces	*Transported provinces*	*Transport volume(Unit: 10000 tons)*
Jiangsu	Shanghai	10.52
Hainan	Guangdong	6.57
Jiangsu	Zhejiang	4.02
Tian	Beijing	0.18

TABLE 9 THE THIRD QUARTER

Exporting provinces	*Transported provinces*	*Transport volume(Unit: 10000 tons)*
Anhui	Zhejiang	16.59
Guangxi	Guangdong	11.93
Jiangsu	Shanghai	6.18
Hebei	Beijing	4.34
Hainan	Guangdong	3.94
Hebei	Tianjin	1.49

TALBE 10 THE FOURTH QUARTER

Exporting provinces	*Transported provinces*	*Transport volume(Unit: 10000 tons)*
Jiangsu	Shanghai	11.72
Guangxi	Guangdong	6.67
Hainan	Guangdong	6.56
Jiangsu	Zhejiang	6.45
Hebei	Beijing	0.68
Hebei	Tianjin	0.28

5.3.2 Part B

Develop an effective pork reserve strategy to address peak demand for pork in a region
1. The main purpose of the pork reserve is to cope with the high price of pork caused by the serious imbalance between supply and demand. Therefore, the government needs to capture the changes in market prices at all times. When pork prices rise, a small amount of reserve pork is placed, and on the contrary, a small amount of pork is repurchased to ensure a certain balance between supply and demand of pork. In order to ensure the timeliness of this regulation strategy, the government needs to predict the growth trend of pork prices and make decisions in advance.
2. China's current pork stocks are still insufficient enough to cope with the soaring pig prices brought about by this year's outbreak. The demand for pork is relatively stable, and the theoretically optimal pork stock can be calculated by the model. However, the total pork production in China is limited, the consumption demand is difficult to meet, and the pork that can be used for storage is even less. Therefore, in this reality, it is necessary to rely on pork imported from abroad, or rationally increase the reserves of substitutes (such as cattle and lamb) to fill the gap between supply and demand of pork.
3, reserve pork has a certain shelf life, usually reserve 4 months or so will replace the new reserve meat, and timely and low-cost processing of frozen meat for a long time, and the quarantine, preservation and replacement process of the reserve pork also need to consume a large amount Human and financial resources. In order to ensure efficiency, the government's control process can be appropriately simplified, and the situation of multi-sectoral cross-responsibility can be reduced. In addition, the company's proper commercial storage of pork and reasonable arrangement of pork storage sites can quickly replenish pork, which is conducive to price stability.
4. Although controlling the price increase of pork will damage the interests of some merchants in the short term, it will also prevent other farmers from causing greater losses due to follow-up pig raising. The government should provide appropriate subsidies for this part of the losses. It can also be achieved by issuing funds and establishing an insurance system.

6. Sensitivity Analysis

1.Impact of Size of Training Set in Linear Regression Model

In linear regression model, the more data used in training set, the better the fitting effect of the model, while the accuracy of the prediction will be lower. In this case, the model is likely to be over-fitted. Conversely, if the training data is too small, the model cannot learn the characteristics of the data. In general, when the size of the training set is controlled at 70%-80%, the model can achieve satisfactory prediction accuracy.

2.Impact of Tolerance Parameter δ in Linear Programming Model

In linear programming model, tolerance parameter is highly correlated with the balance between supply and demand. The lower tolerance parameter is, the more balanced the supply and demand are, the higher the number of iterations of the model is. However, given the limited pork production of each province, demand may not often be met, and there is no solution to the linear programming problem. Therefore, the tolerance parameter should be is appropriately adjusted higher to get a reasonable solution. We set this parameter to 0.01 in Part II and 0.02 in Part III.

7.Strengths and Weakness

7.1 Strengths

1. We use time-series decomposition method to extract the cycle of pork price and verify the existence of the economic phenomenon of "Pig Cycle".

2. From the three perspectives of supply, demand, and external shock, we find out several variables that highly related to the fluctuation pork price and analyze whether it is reasonable to use these variables in our model.

3. We convert the map of China into an undirected graph structure, and use the distance between the provinces in the objective function of the model, so that remote provinces can be used for pig farming while long-distance transportation of pigs and pork can be avoided

4. We fully take into account the location, the scale of urban area, the proportion of rural population and religious beliefs of each province, for these factors may have restriction on the potential maximum pig farming scale.

7.2 Weakness

1. We simply used linear model to predict the trend of pork price, while deep correlations between indicators cannot be found. When the prediction period is too long, the accuracy of the model decreases.

2. In the objective function of the linear programming model, the differences in cost of pig farming between different regions are not considered.

8.Conclusion

In question I, we use time series decomposition method to extract the "Pig Cycle" in pork price fluctuations, and obtained five indicators: corn price, soybean meal price, live pig stock, total yield of beef and mutton and recent pork price, which affect fluctuations through correlation analysis. Then multivariate linear regression is used to fit the trend of historical price, so the correlation of the indicators above is verified.

However, the recent abnormal fluctuations in pork prices cannot be explained by our model. After analysis, we lead to the conclusion that the fluctuation of pork prices is caused by the unbalance of supply and demand. The main reasons behind this are: the outbreak of African swine fever, Sino-US trade friction and poor transportation of pork production and sales areas.

In question II, we mainly study the fallback mechanism of pork prices. Firstly, the relationship of pork price with supply and demand quantity is obtained through the cobweb model. Then we use the linear programming model, taking into account the demand of the provinces, the potential maximum pork production and the cost of inter-provincial transportation, a reasonable pork farming plan is given.

Specifically, the supply of pork in Beijing, Tianjin and Shanghai is seriously insufficient and needs to be imported from other provinces. In other provinces, such as Inner Mongolia, Shandong and Hainan, pork production is higher enough to provide pork. Besides, Qinghai and Tibet have not been affected by this price fluctuation due to their special geographical location and religious beliefs.

However, pork farming usually has a certain periodicity. To achieve a fall in pork prices in the short term, it is necessary to rely on imported pork. We study the import and export of pork in China in recent years, combined with the elasticity of export output, and select importing countries with higher reliability, such as Germany and Spain, to propose corresponding procurement plans.

In question III, we focus on how to keep long-term stability of pork prices through a reasonable farming plan on the premise of relatively stable demand. Here we improve on the basis of the linear programming model in question II, considering the pig breeding cycle, and propose the optimal pig farming plans according to different provinces and different seasons of a year.

However, in order to cope with some sudden surges in demand, such as animal epidemics and natural disasters, in addition to the farming plans above, some emergency measures should be proposed. Finally, we review some relevant literature and propose an effective pork storage strategy from the perspective of government macro-control to ensure long-term stability of pork prices.

References

[1] Zhang Wei, Research on the dependence of Chinese meat import market in the international market, Chinese Journal of Animal Science, 55(08): 174-180, 2019

[2] http://www.stats.gov.cn

[3]https://zhujia.zhuwang.cc

[4]www.chinafarming.com

[5]http://www.chinabric.com/service.html

[6]http://www.agri.cn

[7]http://cyfd.cnki.com.cn/N2019010190.htm

[8]http://www.askci.com

[9] Li Fangmin, Wang Junxi, Research on the dependence of China's pork international market——Based on the comparative analysis of major meat products, Chinese Journal of Animal Science, 55(07):151-156,2019

[10] Li Su, Bao Zhe, Du Chunmei, Chinese pork product supply and demand forecast from the perspective of partial equilibrium theory, Heilongjiang Animal Husbandry and Veterinary, 08:1-7, 2019

[11] Countries seize the Chinese pork market Pork imports rose, Northern Animal Husbandry, 20:19, 2019

[12]Chris Hurt. Pork prices spike due to African swine fever[J]. Southwest Farm Press,2019.

[13]. Seasonal meat price rallies end; pork producers face uncertain market[J]. Southwest Farm Press,2019.

[14]Chris Hurt. 2019 pork and corn price volatility: Good or bad?[J]. Farm Industry News,2019.

[15]Xiaoxia Dong,Colin Brown,Scott Waldron,Jing Zhang. Asymmetric price transmission in the Chinese pork and pig market[J]. British Food Journal,2018,120(1).

Appendix

[Python Code] quest_1.py

```
import re
import requests
import time
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from pandas import DataFrame,Series
from statsmodels.tsa.arima_model import ARIMA
from statsmodels.tsa.seasonal import seasonal_decompose
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import *
from sklearn.svm import SVR
from scipy.stats import kstest
plt.rcParams['font.sans-serif'] = ['SimHei']
plt.rcParams['axes.unicode_minus'] = False
print "Please run this code with Python 2"
data = pd.read_csv("data-explore.csv",header=0,encoding="gbk",parse_dates=['date'])
print data.shape
print np.array(data.corr())
cols = ["maize","bean","store","nbfsp","lag"]
names = data.columns.values.tolist()
T = 6
Y = data["pork"]
lag = [Y[i-T] if i-T>=0 else 0 for i in range(len(Y))]
data["lag"] = lag
decomposition = seasonal_decompose(Y,freq=12,model="multiplicative")
trend = decomposition.trend
season = decomposition.seasonal
resid = decomposition.resid
plt.figure(41)
axes = plt.subplot(411)
plt.plot(data['pork'],'-',c="black",label="pork")
plt.legend(loc="upper left")
axes.set_xticks([])
axes = plt.subplot(412)
plt.plot(trend,'-',c="darkblue",label="trend")
plt.legend(loc="upper left")
axes.set_xticks([])
```

```
axes = plt.subplot(413)
plt.plot(season,'-',c="darkgreen",label="season")
plt.legend(loc="upper left")
axes.set_xticks([])
plt.subplot(414)
plt.plot(resid,'-',c="darkred",label="resid")
plt.legend(loc="upper left")
plt.xlabel("Time")
plt.show()
data = pd.read_csv("data-train.csv",header=0,encoding="gbk",parse_dates=['date'])
Y = data["pork"]
lag = [Y[i-T] if i-T>=0 else 0 for i in range(len(Y))]
data["lag"] = lag
sdm = seasonal_decompose(Y,freq=12,model="multiplicative")
trend = sdm.trend
season = sdm.seasonal
resid = sdm.resid
Yt = trend[6:-6]
Xt = np.array(data[cols])[6:-6]
X_train, X_test, Y_train, Y_test = train_test_split(Xt, Yt, test_size=0.30, random_state=233)
model = LinearRegression()
model.fit(X_train,Y_train)
a = model.intercept_
b = model.coef_
print a,b
score = model.score(X_test,Y_test)
print score
Y_pred = model.predict(Xt)
MSE = mean_squared_error(Yt, Y_pred)
RMSE = np.sqrt(mean_squared_error(Yt, Y_pred))
print 'R2:',r2_score(Yt,Y_pred)
print 'MSE:',MSE
print 'RMSE:',RMSE
plt.figure(figsize=(10,5))
plt.plot(range(len(Yt)), Yt, 'b', label='test')
plt.plot(range(len(Yt)), Y_pred, 'r', label='predict')
plt.legend()
plt.show()
print season[:12]
res_model = ARIMA(np.array(resid[6:-6]), order=(6,1,0))
print kstest(resid[6:-6], 'norm')
res_fit = res_model.fit(disp=0)
print res_fit.summary()
ans = []
```

```
pre = []
for i in range(6,len(Y)-6):
    ans.append(float(Y[i]))
    pre.append(float(season[i%12]*Y_pred[i-6]))
plt.figure(figsize=(10,5))
plt.plot(range(len(ans)), ans, 'b', label='test')
plt.plot(range(len(ans)), pre, 'r', label='predict')
plt.legend()
plt.show()
data2 = pd.read_csv("data-test.csv",header=0,encoding="gbk",parse_dates = ['date'])
print data2.shape
Y_ = data2['pork']
lag = [Y_[i-T] if i-T>=0 else 0 for i in range(len(Y_))]
data2["lag"] = lag
X_ = data2.loc[:,cols]
trend_pred = model.predict(X_)
Y_pred2 = []
eps_pred = []
for i in range(len(Y_)):
    Y_pred2.append(float(trend_pred[i])*float(season[i%12]))
    eps_pred.append(Y_pred2[-1]-Y_[i])
plt.figure(figsize=(10,5))
plt.plot(range(len(Y_[T:])), Y_[T:], 'b', label='test')
plt.plot(range(len(Y_[T:])), Y_pred2[T:], 'r', label='predict')
plt.ylim(5,45)
plt.legend()
plt.show()
print len(Y_pred)
eps = []
for i in range(6,len(Y)-6):
    eps.append(float(Y_pred[i-6])*float(season[i%12])-Y[i])
eps = eps[T:]
print kstest(eps,'norm')
print np.mean(eps),np.std(eps)
eps_pred = eps_pred[T:]
plt.figure()
plt.plot(range(len(eps_pred)),eps_pred)
print len(eps_pred)
for i in range(len(eps_pred)):
    if abs(eps_pred[i])>3*np.std(eps):
        print "out",i,eps_pred[i]/np.std(eps)
plt.show()
```

[Python Code] quest_2.py

```
import re
import requests
import time
import numpy as np
import pandas as pd
from scipy import optimize
data = pd.read_csv("adjacent.csv",header=0,encoding="gbk")
prov = data['Province']
n = len(prov)
data = np.array(data)
dist = np.array([data[i][1:] for i in range(n)])
fid = {}
idx = 0
for e in prov:
    fid[e] = idx
    idx += 1
for i in range(n):
    for j in range(n):
        if i!=j and dist[i,j]==0:
            dist[i,j] = float("inf")
def floyd(dist):
    for k in range(n):
        for i in range(n):
            for j in range(n):
                if dist[i,j]>dist[i,k]+dist[k,j]:
                    dist[i,j] = dist[i,k]+dist[k,j]
    return dist
dis = floyd(dist)
def funD(p):
    return (-6.843*1e-5*p*p*1e6 + 2.414*p*1e3 + 26616.021)/12.0
def funS(p_):
    return (50628.798 - 6.288*1e7/p_*1e-3)/12.0
data_sp = pd.read_csv("supply.csv",header=0,encoding="gbk")
s = [float("nan") for i in range(n)]
m = [float("nan") for i in range(n)]
for i in range(data_sp.shape[0]):
    e = data_sp["Province"][i]
    if e in fid:
        idx = fid[e]
    else:
        continue
    p0,p0_ = data_sp["P11"][i], data_sp["P10"][i]
    pt_ = data_sp["Pold"][i]
    d0 = data_sp["D2019"][i]*10000.0/12.0
```

```
        s0 = d0 * pt_ / p0_
        smax = data_sp["S2017"][i]*10000.0/12.0
        m[idx] = max(smax - s0,0)/10000.0
        s[idx] = max(d0 - s0,0)/10000.0
m,s = np.array(m), np.array(s)
m[fid["Tibet"]]=0
s[fid["Tibet"]]=0
for e in prov:
        print(e,m[fid[e]],s[fid[e]])
k = 0.01
z,a,b = [],[],[]
bound_list = []
for i in range(n):
        for j in range(n):
                z.append(dis[i,j])
                bound_list.append((0,None))
for i in range(n):
        tmp = np.zeros([n*n])
        for j in range(n):
                tmp[i*n+j] = 1
        a.append(tmp)
        b.append(m[i])
for j in range(n):
        tmp = np.zeros([n*n])
        for i in range(n):
                tmp[i*n+j] = 1
        a.append(tmp)
        b.append(s[j]*(1+k))
        a.append(-tmp)
        b.append(-s[j]*(1-k))
z,a,b = np.array(z), np.array(a), np.array(b)
bound_list = np.array(bound_list)
res = optimize.linprog(z,A_ub=a,b_ub=b,bounds=bound_list)
for i in range(n):
        for j in range(n):
                if res['x'][i*n+j]>0 and i!=j:
                        print(prov[i],prov[j],res['x'][i*n+j])
for i in range(n):
        for j in range(n):
                if res['x'][i*n+j]>0 and i==j:
                        print(prov[i],prov[j],res['x'][i*n+j])
for i in range(n):
        print(prov[i]+"\t"+str(sum(res['x'][i*n:(i+1)*n]))+"\t"+str(s[i]))
```

[Python Code] quest_3.py

```
import re
import requests
import time
import numpy as np
import pandas as pd
from scipy import optimize
data = pd.read_csv("adjacent.csv",header=0,encoding="gbk")
prov = data['Province']
n = len(prov)
data = np.array(data)
dist = np.array([data[i][1:] for i in range(n)])
fid = {}
idx = 0
for e in prov:
    fid[e] = idx
    idx += 1
for i in range(n):
    for j in range(n):
        if i!=j and dist[i,j]==0:
            dist[i,j] = float("inf")
def floyd(dist):
    for k in range(n):
        for i in range(n):
            for j in range(n):
                if dist[i,j]>dist[i,k]+dist[k,j]:
                    dist[i,j] = dist[i,k]+dist[k,j]
    return dist
dis = floyd(dist)
T = 4
L = 90
data_sp = pd.read_csv("supply.csv",header=0,encoding="gbk")
data_price = pd.read_csv("price.csv",header=0,encoding="gbk")
p0 = [float("nan") for i in range(n)]
price = [[float("nan") for i in range(n)] for t in range(T)]
for i in range(data_sp.shape[0]):
    e = data_sp["Province"][i]
    if e in fid:
        idx = fid[e]
    else:
        continue
    p0[idx] = np.nanmean(data_price[e])
    for t in range(T):
        price[t][idx] = np.nanmean(data_price[e][t*L:(t+1)*L])
price = np.array(price)
```

```
s = [[float("nan") for i in range(n)] for t in range(T)]
m = [float("nan") for i in range(n)]
for i in range(data_sp.shape[0]):
    e = data_sp["Province"][i]
    if e in fid:
        idx = fid[e]
    else:
        print(e)
        continue
    d0 = data_sp["D2017"][i]/T
    smax = data_sp["S2017"][i]/T
    for t in range(T):
        s0 = d0*price[t][idx]/p0[idx]
        s[t][idx] = max(s0,0)
    m[idx] = max(smax,0)
m = np.array(m)
s = np.array(s)
m[fid["Tibet"]]=0
for t in range(T):
    s[t,fid["Tibet"]]=0
print(np.nansum(m))
print(np.nansum(s))
m *= (T*0.5)
k = 0.02
z,a,b = [],[],[]
bound_list = []
for t in range(T):
    for i in range(n):
        for j in range(n):
            z.append(dis[i,j])
            bound_list.append((0,None))
for t in range(T):
    for i in range(n):
        tmp = np.zeros([T*n*n])
        for j in range(n):
            tmp[t*n*n+i*n+j] = 1
            tmp[((t+1)%T)*n*n+i*n+j] = 1
        a.append(tmp)
        b.append(m[i])
for t in range(T):
    for j in range(n):
        tmp = np.zeros([T*n*n])
        for i in range(n):
            tmp[t*n*n+i*n+j] = 1
```

```
        a.append(tmp)
        b.append(s[(t+2)%T,j]*(1+k))
        a.append(-tmp)
        b.append(-s[(t+2)%T,j]*(1-k))
z = np.array(z)
a = np.array(a)
b = np.array(b)
bound_list = np.array(bound_list)
optimize.linprog(z,A_ub=a,b_ub=b,bounds=bound_list,method="simplex",options={"maxiter
":1000000})
print(res)
for t in range(T):
    for i in range(n):
        for j in range(n):
            if res['x'][t*n*n+i*n+j]>1e-6 and i!=j:
                print(prov[i],prov[j],res['x'][t*n*n+i*n+j])
for t in range(T):
    for i in range(n):
        sums = 0
        for j in range(n):
            if res['x'][t*n*n+i*n+j]>0 and i==j:
                sums += res['x'][t*n*n+i*n+j]
        print(prov[i],sums)
```