

Laundry Strategy Model Based on Nonlinear Programming

Summary

Washing clothes is something that people have to do every day, in practice, because of the variety of detergents and stains, how to use a lower cost method to wash clothes clean is a problem that everyone will consider. The purpose of this paper is to solve a cost-effective mathematical model of washing clothes, while meeting the needs of sustainable development.

For problem 1, the amount of dirt attached to clothes and the unit water consumption and total water consumption under the condition solubility of problem were first set. Constraint conditions were established through nonlinear programming, and simulated annealing was adopted to solve the problem. Given the specific value of each parameter, bring into the model calculation and finally get the best cleaning scheme was to clean twice, and the best washing scheme was when water was 43.048 liters. The increase of water consumption and coefficient dirt solubility can increase the washing effect.

For problem 2, the washing scheme that saves the most time is required, that is, the washing frequency is the lowest. At this time, the annealing model is modified and the water consumption no longer affects the objective function, that is, the decision variable is completely dependent on the washing frequency. Finally, the optimal washing scheme with 66.873 liters of water used for two washing times is obtained. Then the control variable method was used to change the specific value of the parameter, and it was found that the higher the solubility of dirt and water consumption, the better the influence on the washing effect, while the initial dirt amount and the proportional coefficient of solubility decay had little influence on the washing effect.

For problem 3, in order to find a cost-effective and good cleaning plan, the clothes were cluster analyzed according to the types of clothes stains, and the best cleaning plan was obtained when the clothes were divided into two categories. The first type of clothes was best washed with the second detergent in Table 3, and the second type of clothing is best washed with a mixture of the second and fifth detergents.

For problem 4, we classify the clothes that can be cleaned together, screen out the four categories currently divided, and then perform the same operation according to problem 3, obtain n possible optimal solutions according to the clustering results again, and finally determine the final scheme.

Keywords: Nonlinear programming ; Simulated annealing algorithm; Multi-objective programming

Contents

1 Introduction	4
1.1 Problem Background	4
1.2 Restatement of the Problem	4
1.3 Our Work.....	4
2 Assumptions and Justifications.....	5
3 Notations	5
4 Solution to Problem 1	6
4.1 The Establishment of Model 1	6
4.1.1	6
The initial amount of dirt in the k-th washing is determined by the remaining amount after the k-1-th washing.....	6
4.1.2	6
4.1.3 Determination of cleaning program.....	7
4.2 Test Model 1.....	7
4.2.1 Simulated annealing algorithm ^[2]	7
4.2.2 The solution results of the model	8
5 Solution to Problem 2	11
5.1 The Establishment of Model 2	11
5.1.1	11
5.1.2 Determination of cleaning program.....	11
5.2 Test Model 2.....	11
6 Solution to Problem 3	13
6.1 The Establishment of Model 3	13
6.1.1 Clothing grouping.....	13
6.1.2	13
6.1.3 Determination of cleaning program.....	14
6.2 Test Model 3.....	14
6.2.1 K-means clustering algorithm ^[3]	14
6.2.2 The solution results of the model	15
7 Solution to Problem 4	16
7.1 The Establishment of Model 4	16
7.1.1 Clothing grouping.....	16
7.1.2 Determination of cleaning program.....	16

7.2 Test Model 4.....	17
8 Model Evaluation and Possible Improvements.....	18
8.1 Strengths	18
8.2 Disadvantages	18
8.3 Possible Improvements	18
9 Conclusion.....	19
References	20

1 Introduction

1.1 Problem Background

Washing clothes is something we do every day, and it involves the use of chemicals called surfactants. Surfactants help water penetrate and remove dirt by creating an electric push effect between dirt and water. Surfactants have two distinct parts: one is lipophilic, which attracts dirt and repels water, and the other is hydrophilic, which attracts water and repels oil. When surfactants are added to water, they stick to any surface that is not wet, such as cloth or items to be washed. The hydrophilic part of the surfactant reduces the surface tension of the water, allowing it to enter the gap between the dirt and the surface. The mechanical action of washing removes the dirt surrounded by the oil-wet part of the surfactant and rinses it off with water. This is the math hidden in laundry.

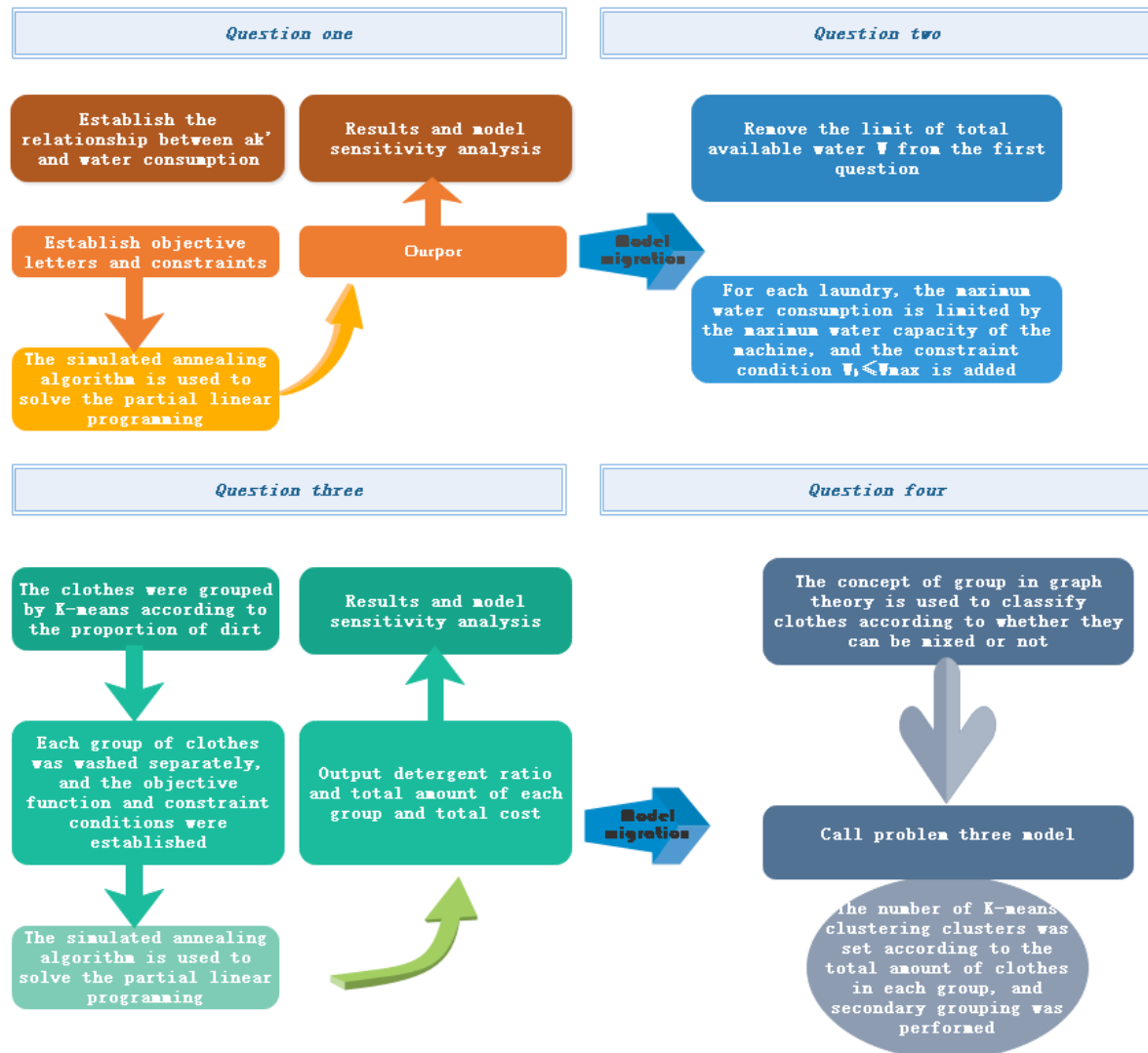
1.2 Restatement of the Problem

In practice, whether for household or professional use, there is a question to consider: How to achieve the best cleaning results with the lowest cost? This seemingly simple question actually contains deep mathematical principles. Based on this, we need to address the following issues:

- Problem 1: Given the amount of dirt, the amount of water and the solubility of dirt in water, find the best cleaning method, including the number of cleaning and water consumption each time, and analyze the influence of different factors on the results.
- Problem 2: Assuming that each cleaning time is the same, the amount of water is unlimited, and the final dirt residue cannot exceed one thousandth of the initial amount, find the most time-saving cleaning program, and analyze the influence of different factors on the results.
- Problem 3: There are many clothes, the type and amount of dirt on each piece of clothing are different, ten different detergents are given, each detergent has different solubility and price for different dirt, and the solution that saves cost and can be cleaned is found.
- Problem 4: There are many pieces of clothing, the material of each piece of clothing is different, the type and amount of dirt on each piece of clothing is different, taking into account the material, color and other factors, some clothes can not be mixed cleaning, find both cost saving and time saving program.

1.3 Our Work

The topic asks us to give the best solution for washing clothes according to the nature of detergent, solubility, etc. Our work mainly includes the following aspects:



2 Assumptions and Justifications

To simplify the problem, we make the following basic assumptions, each of which is properly justified.

- Ignore the influence of water quality, temperature and other factors.
- When there is a variety of dirt on a garment, the solubility of the detergent to a certain dirt is not affected by other dirt.
- The dirt solubility set in Table 2 is the weight of dirt that can be dissolved per gram of detergent. Each gram of detergent requires 0.75 liters of water.

3 Notations

The key mathematical notations used in this paper are listed in Table 1.

Table 1: Notations used in this paper

Symbol	Description
D_k	the amount of dirt after the k-th washing
W_k	the amount of water used for the k-th washing
C_i	the i-th piece of clothing
$Q_n(i)$	the amount of the nth kind of dirt on the i-th piece of clothing
$Wipe_m^n$	the removal rate of the m-th kind of detergent for the n-th kind of dirt
P_m	the unit price of the m-th kind of detergent

4 Solution to Problem 1

4.1 The Establishment of Model 1

4.1.1

Water is considered as a base detergent when the garment is simply washed with water without the use of additional detergent. In order to make the model fit the reality, a'_k is further explained here :

represents the proportion of dirt dissolved by an equal amount of detergent during the k-th washing relative to the initial dirt amount. When the detergent is doubled, the amount of dissolved dirt will not simply be doubled, but meet a certain relationship, which can be set according to the relevant articles^[1]

$$a'_k = 1 - b_k \cdot e^{\beta(-\frac{1}{k}\alpha + 1)} \quad (1)$$

Where $b_k = 1 - a_k$ is defined as the residual rate, β is the undetermined coefficient.

$$\alpha = \frac{\text{Actual detergent consumption}}{\text{Standard dosage}} \geq 1 \quad (2)$$

The initial amount of dirt in the k-th washing is determined by the remaining amount after the k-1-th washing.

4.1.2

Let the amount of dirt attached to clothes be D_0 , so that the unit water consumption of dirt meeting the solubility of a is W_0 , and the total available water is W_{\max} .

From this, the following relationship can be derived:

$$\begin{cases} a_k = \frac{a_{k-1}}{2} \\ \alpha = \frac{W_k}{W_0} \\ D_k = (1 - a'_k) D_{k-1} \end{cases} \quad (3)$$

4.1.3 Determination of cleaning program

To get the optimal solution, solve for $\min f1 = \left[\frac{k}{k_{\max}} \cdot \varepsilon + \frac{\sum_{n=0}^k W_k}{W_{\max}} (1 - \varepsilon) \right]$.

Can be further established:

$$s.t. \left\{ \begin{array}{l} \frac{D_k}{D_0} < 0.002 \\ \sum_{n=1}^k W_k \leq W_m \\ 1 \ll k \ll k_{\max} \\ a_k = \frac{a_{k-1}}{2} \\ \alpha = \frac{W_k}{W_0} \\ D_k = (1 - a'_k) D_{k-1} \end{array} \right. \quad (4)$$

(ε represents the weight of evaluation parameter)

4.2 Test Model 1

4.2.1 Simulated annealing algorithm^[2]

Based on the above model, we can use simulated annealing algorithm(SA) to solve the problem.

The simulated annealing algorithm is a random search optimization algorithm, which is inspired by the solid annealing process in physics, and seeks a global optimal solution or approximate optimal solution by simulating the heat equilibrium state of the solid at high temperature and the cooling state at low temperature.

As a general optimization method and an extension of local search method, different from local search algorithm, SA selects the inferior solution with large target value in the neighborhood with a certain probability.

For the first question, we manually set the unit water consumption, available water, initial dirt amount, a_1 and attenuation coefficient, and the ratio of the final remaining dirt amount to the initial dirt amount (that is, the cleaning effect). We set the maximum cleaning times to ten times. Through the simulated annealing algorithm, the water consumption, the amount of residual dirt, the cleaning times and the value of the optimal solution.

The algorithm flow is as follows:

1. Generate the initial solution, the solution sequence is the sequence of water consumption each time, and the initial solution only needs to ensure that the total water consumption does not exceed the available water;
2. Distinguish between the solution sequence and the solution result, which is the value of the decision variable. In the annealing process, we calculate the initial solution first, and then perturbate the current solution to make it close to the optimal solution.

3. The disturbance principle is as follows: for the current solution sequence, under the premise of not exceeding the available water quantity, the water consumption fluctuates up and down each time to obtain a new solution sequence; Then the new solution sequence is obtained and compared with the original solution. If the new solution is better than the original solution, it is replaced directly. If it is not better than the original solution, the current solution is accepted according to a certain probability, the probability is as follows:

$$P = \begin{cases} 1, & E_{t+1} < E_t \\ e^{\frac{-(E_{t+1}-E_t)}{kT}}, & E_{t+1} \geq E_t \end{cases} \quad (5)$$

4. Because the disturbance is affected by the initial solution, if the initial solution is not feasible, in some cases the final feasible solution cannot be obtained even if the initial solution is disturbed several times. Therefore, before annealing, we repeatedly generate the initial solution sequence and solve the initial solution results until a feasible solution is obtained, and then carry out the subsequent annealing process on it;
5. Similarly, because the annealing result is greatly affected by the initial solution, we need to carry out multiple annealing to weaken the influence of the initial solution.

4.2.2 The solution results of the model

Given $W_0=1L$, $W_{\max}=60L$, $D_0=200g$, the model is solved as follows:

Figure 1: The model runs ten times the distribution of objective function results

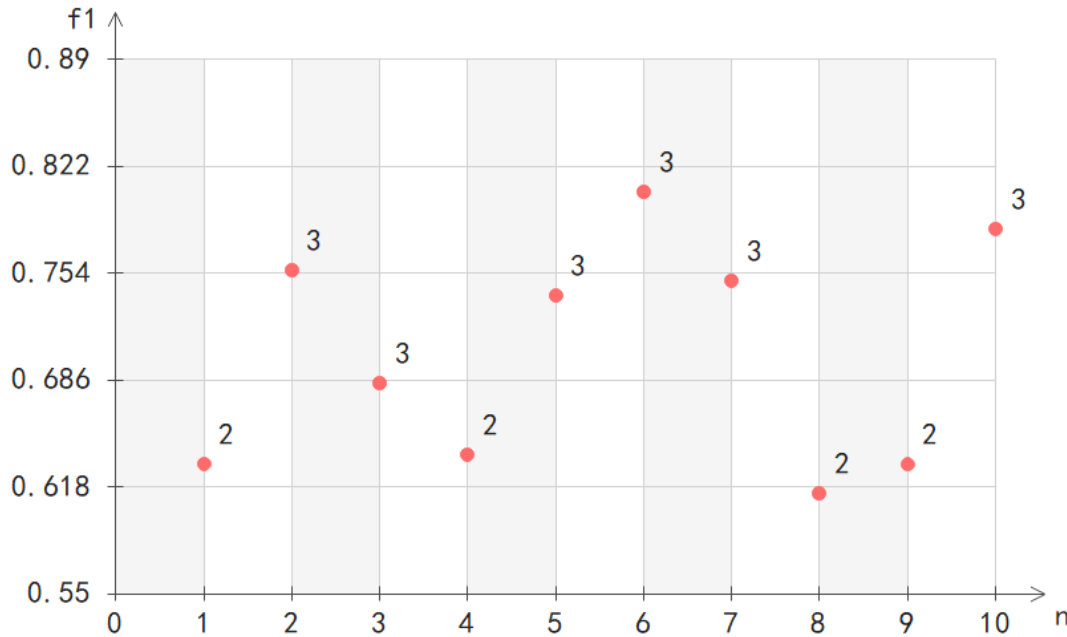
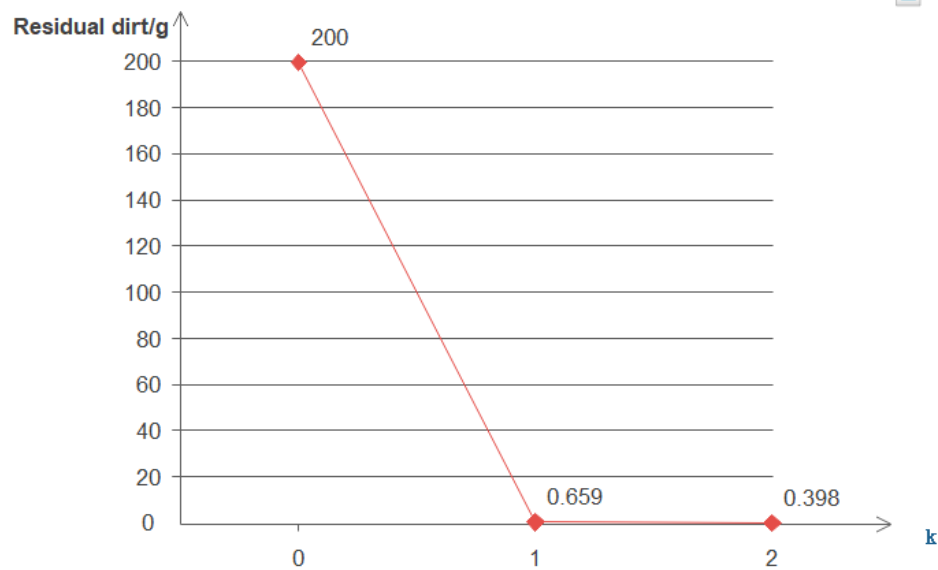


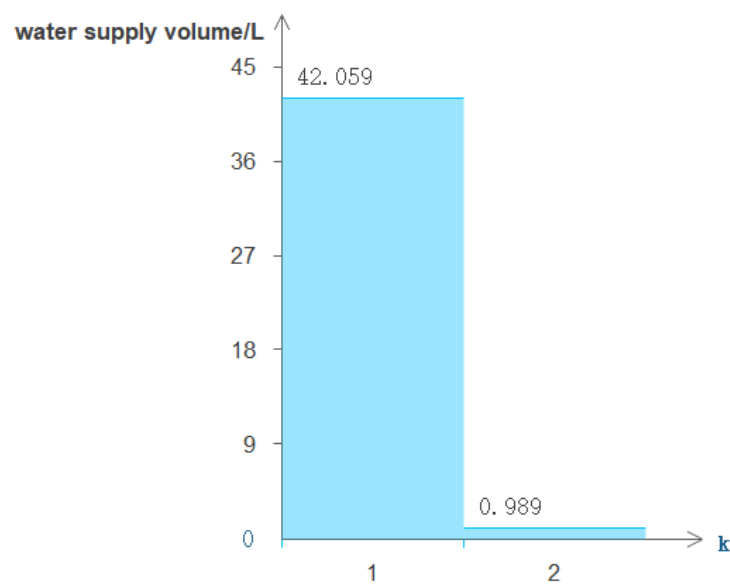
Figure 2: The relationship between residual amount of dirt and washing times in the optimal scheme



According to the results of the model, when the washing frequency is 2, the cleaning effect under the optimal solution can be achieved.

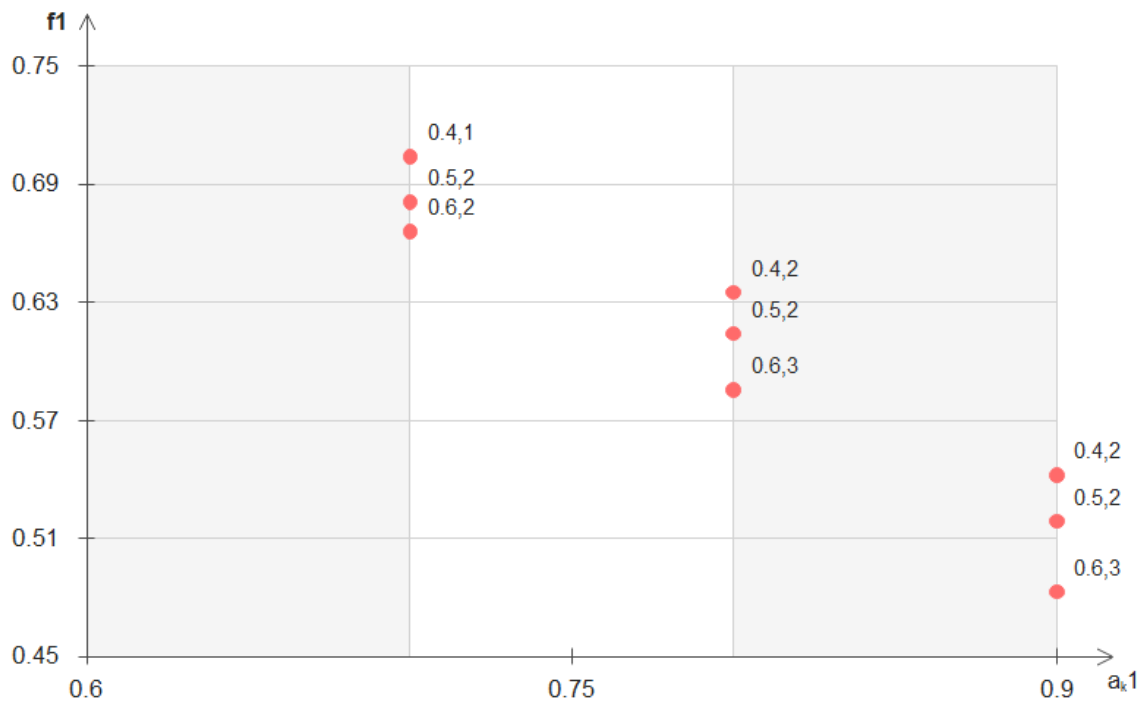
As shown below, the total water consumption at this time is $42.059 + 0.089 = 43.048$ litres.

Figure 3: Optimal solution water consumption per wash



Then discuss the effect of a_1 , the initial amount of dirt, and the amount of water available on the objective, as shown in Figure 4.

Figure 4: a_1 and the influence diagram of the change of attenuation coefficient m on the scheme



In the above table, $f1$ is our evaluation of the washing standard, when the $f1$ is lower, the washing effect is better.

Table 1: Influence of the maximum available water on the cleaning scheme when $D_0 = 200g$

W_m/W_0	70	60	50
$f1$	0.560	0.614	0.748
Number of washes	1	2	2

Table 2: When the maximum available water W is unchanged, the influence of D_0 change on the cleaning scheme

D_0	100	200	300
$f1$	0.617	0.614	0.618
Number of washes	2	2	2

- ✧ According to the chart analysis, the higher the maximum water consumption, the fewer the washing times required to meet the target cleaning requirements;
- ✧ According to the definition of all solubility of the topic, the dirt content has little influence on the cleaning scheme;
- ✧ The larger the parameter, the better the cleaning effect, and the higher the ratio parameter, the better the washing effect.

5 Solution to Problem 2

5.1 The Establishment of Model 2

5.1.1

Problem 2 Removes the limit of total available water w in problem 1. The water consumption of each washing can not be infinite, so the maximum water consumption limit of each washing amount is increased $W_k \leq W_{\max}$.

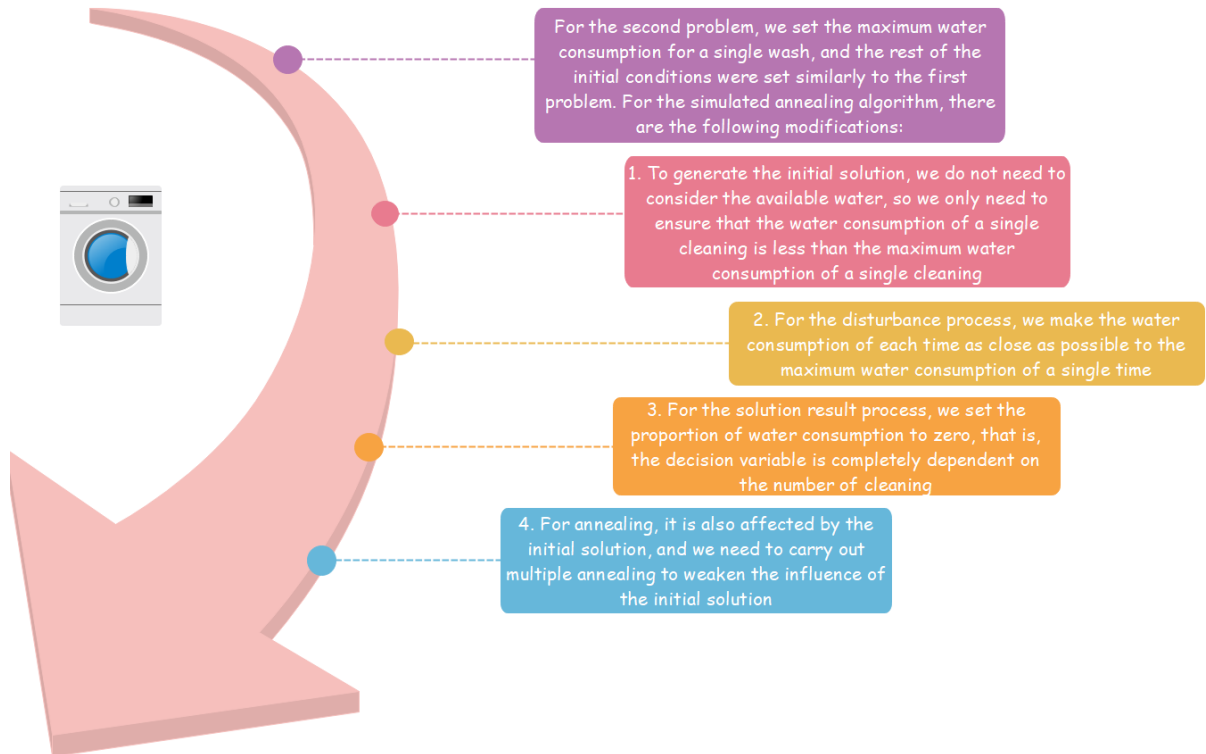
5.1.2 Determination of cleaning program

According to the analysis of the meaning of the problem, the establishment of the objective function can be obtained, that is, solving k_{\min} .

$$s.t. \begin{cases} \frac{D_k}{D_0} \leq 0.001 \\ 1 \leq k \leq k_{\max} \\ W_k \leq W_{\max} \end{cases} \quad (6)$$

5.2 Test Model 2

Based on model analysis, we can get the solution idea of the second model as follows:



Given $W_0=1L$, $D_0=200g$, $W_{\max}=40L$ the model is solved as follows:

Figure 5: The relationship between residual amount of dirt and washing times in the optimal scheme

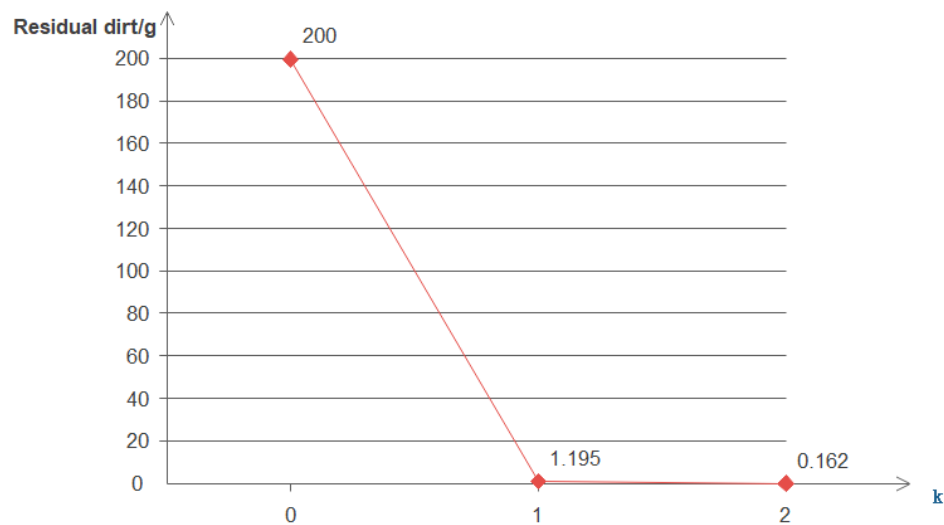
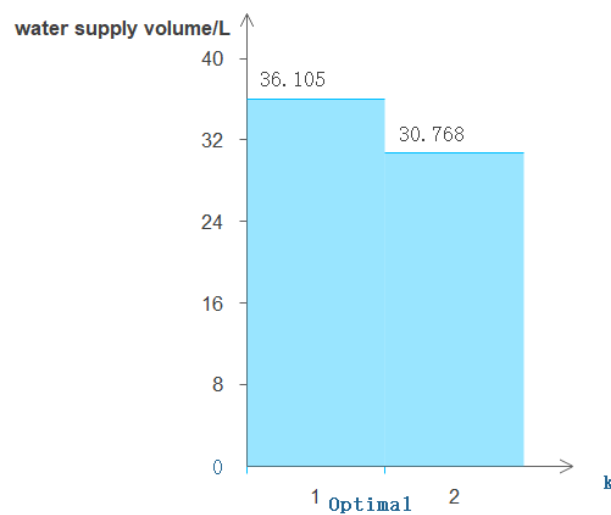


Figure 6: Optimal solution Water consumption per wash



According to the analysis of FIG. 5 and FIG. 6, the washing scheme that saves the most time is washing twice, when the water consumption is $36.105 + 30.768 = 66.873$ liters.

Then analyze the impact of a_k and the initial dirt amount on the optimal solution:

$W_0=1L, W_{\max}=40L$, influence of a_1 change on washing times when attenuation coefficient $m=0.5$ and $D_0=200g$.

A_1	0.97	0.8	0.5
Number of washes	1	2	3

When $W_0=1L, W_{\max}=40L, a_1=0.8, D_0=200g$, the effect of m change on washing times

m	0.7	0.5	0.3
Number of washes	2	2	2

When $W_0=1L$, $a_1=0.8$, attenuation coefficient $m=0.5$, $D_0=200g$ the effect of W_{max} change on washing times

W_{max}	60L	40L	30L
Number of washes	1	2	3

When $W_0=1L$, $W_{max}=40L$, attenuation coefficient $m=0.5$, $a_1=0.8$, the effect of D_0 change on washing times

W_0	300	200	100
Number of washes	2	2	2

It can be obtained from the above table analysis:

- ✧ The larger the a_1 , the better the washing effect;
- ✧ The greater the maximum amount of water used for word washing, the better;
- ✧ The initial amount of dirt and proportion coefficient have little effect on the washing effect.

6 Solution to Problem 3

6.1 The Establishment of Model 3

According to the data of the Blue Moon website, the recommended amount of laundry detergent for 15 liters of water is 20 grams. To simplify the model, set the water requirement for one gram of detergent at 0.75 liters.

The maximum water consumption of the washing machine W_m limits the amount of washing per load. According to the data on the official website of the United States, the maximum water capacity of the machine is set at 144L in this paper.

6.1.1 Clothing grouping

Due to the limited capacity of washing machine, in order to facilitate the choice of washing and detergent, the similarity of dirt content between clothes is used to cluster and group.

6.1.2

Total amount of dirt in Group j clothing $D_n^j = \sum_{i=1}^{36} x_i^j Q_n(i)$.

Among them, $X_i^j = \begin{cases} 1, \text{Clothes } C_i \text{ are divided into groups } j \\ 0, \text{Clothes } C_i \text{ are not divided into groups } j \end{cases}$

The amount of the m detergent in the k washing of group j is $De_m^j(k)$. And the remaining amount of the n th type of dirt after the k -th wash of the j -th group is $D_n^j(k)$.

$$D_n^j(k) = D_n^j - \sum_k Wipe_m^n De_m^j(k) \quad (7)$$

The total cost of detergent is $Ptd = \sum_j \sum_k \sum_m De_m^j(k)$, and the total water cost is

$$Ptw = \frac{Wd \cdot \sum_j \sum_k \sum_m De_m^j(k)}{1000} \cdot Pw . \text{So the determination of cleaning program is}$$

$$Pt = Ptd + Ptw \quad (8)$$

6.1.3 Determination of cleaning program

According to the meaning of the question, according to the determination of the cleaning scheme, the objective function is to find $\min Pt$.

The constraints are as follows:

$$s.t. = \begin{cases} W_d \cdot \sum_m De_m^j(k) \leq Wm \\ \frac{D_n^j(k)}{D_n^j} \leq 0.001 \\ 1 \leq k \leq k_{\max} \end{cases} \quad (9)$$

6.2 Test Model 3

6.2.1 K-means clustering algorithm^[3]

K-means clustering algorithm is an unsupervised learning method that can partition a data set into K clusters, such that the data points within the same cluster have high similarity, and the data points between different clusters have low similarity.

The advantages of the K-means clustering algorithm are that it is simple and easy to understand, has high computational efficiency, and is suitable for processing large-scale data sets. Based on this, for the third question, we first processed the data.

Through cluster analysis, the clothes with similar dirt are divided into three categories (the effect of two categories can also be tested); We wash each kind of clothes separately, that is, several times in several kinds. Secondly, we judge the cleaning ability of different detergents for different dirt, and the judging rules are as follows: We set that 1g of detergent needs 0.75L of water, so as to calculate the water cost of 1g of detergent. By adding the water cost and the unit price of detergent, we get the cost of 1g of detergent (that is, including the water cost). At this time, divide the solubility of detergent for different dirt by the cost of 1g detergent. We can get the solubility of different dirt under the unit price, compare the solubility of different detergents vertically, we can get the best detergent for each type of dirt, that is, get a sequence:

$$\text{best} = [2, 5, 2, 5, 2, 3, 2, 6];$$

For the first type of dirt, the second detergent has the best effect; For the second type of dirt, the fifth detergent works best, and so on.

For the annealing process, there are the following modifications;

- Generate initial solution: For each type of clothing, we calculate the total amount of each type of dirt, and then find the most current type of dirt, so as to select the type of detergent that should be selected (that is, in addition to this type of dirt effect of the best detergent), the number is random within 1-10, and then calculate the amount

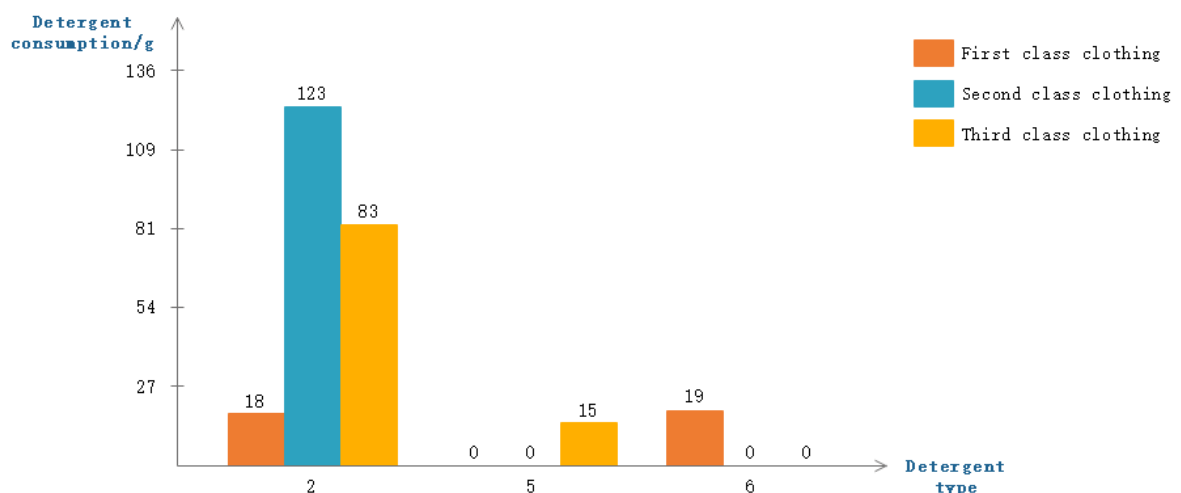
of remaining dirt, and then find the amount of remaining dirt, which type of dirt is the most, and continue to cycle. Until the remaining amount of dirt is 0 (the goal is complete cleaning, but you can also set the cleaning effect to one thousandth).

- We get the initial solution sequence, that is, the amount used by each type of detergent, so the solution result only needs to traverse the solution sequence and calculate the required amount, taking the amount as the decision variable to make it as small as possible.
- For the disturbance, since the process of generating the initial solution sequence is an optimization process, there is no need for disturbance. In the annealing process, at different temperatures, we only need to generate a solution again to be regarded as a new solution.

6.2.2 The solution results of the model

After cluster analysis, clothes can be divided into three categories and two categories for treatment according to the categories of clothing dirt, respectively analyzed as follows:

Figure 7: The amount of detergent used for each type of clothing when it is divided into three types

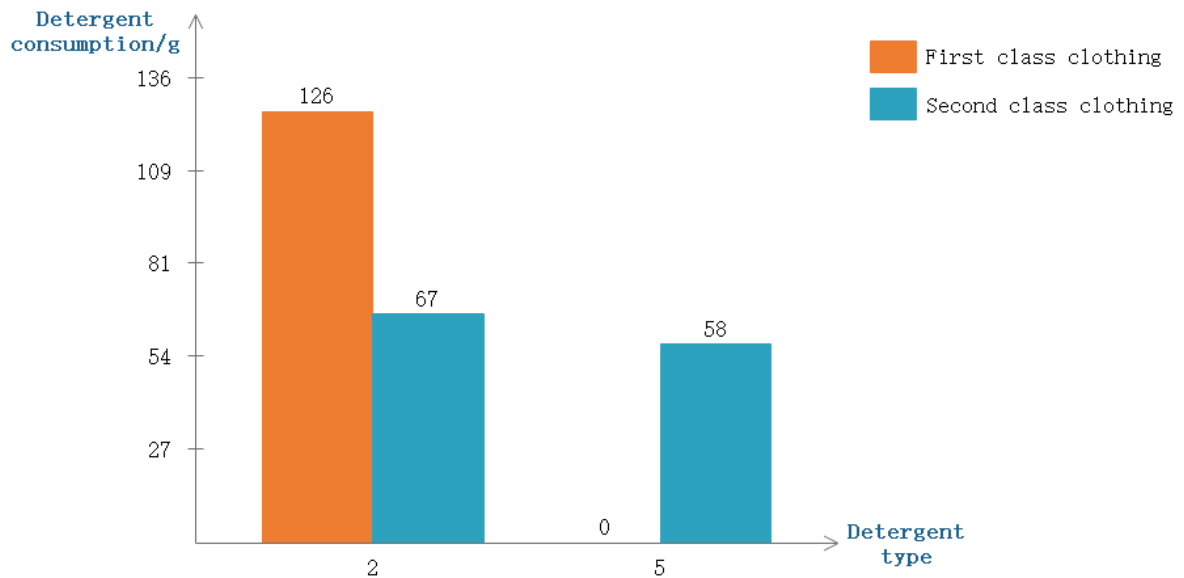


The clothes are grouped into 3 categories, as shown in Figure 7, and the optimal result is as follows:

- ✧ The first type of clothing is best washed with the second and sixth types of detergent;
- ✧ The second type of clothing is best washed with the second detergent;
- ✧ Third types of clothing are best washed with second and fifth detergents.

At this time the total cost of consumption is 24 yuan, the total number of washes is 3.

Figure 8: The amount of detergent used for each type of clothing when it is divided into two types



The clothes are grouped into two categories, as shown in Figure 8, and the optimal result is as follows:

- ✧ The first type of clothing is best washed with the second type of detergent;
- ✧ The second type of clothing is best washed with the second and fifth detergents.

The total price is 22.73 yuan, the total number of washes is 2.

Whether it is according to the number of washing or according to the total cost, it is better to be divided into two categories, and the lowest cleaning cost is 22.73 yuan.

7 Solution to Problem 4

7.1 The Establishment of Model 4

7.1.1 Clothing grouping

Table 3 has a total of eight materials of clothing, Table 4 shows which clothes can be mixed, which can not, these clothes should be washed in groups, so that the clothes in each group can be mixed, we can regard this problem as a graph theory problem.

The matrix can be thought of as an adjacency matrix, where the rows and columns of the matrix represent the vertices in the graph (here the types of clothing), and the elements of the matrix indicate whether these vertices (clothing) are connected (i.e. can be put together).

To find the fewest combinations, we can try to find a maximum clique partition of the graph. A clique is a subgraph in which the vertices are connected to each other. Maximum clipping means dividing the vertices of a graph into as few cliques as possible, such that each vertex belongs to a clique and the vertices within each clique are connected to each other.

7.1.2 Determination of cleaning program

The clothes are divided into x group according to whether they can be mixed.

Each of the x groups can be solved using the model of problem 3. Suppose n alternatives

are obtained, then the final index selected in the scheme is $\min f2 = \varepsilon \cdot \frac{P_i}{P_{\max}} + (1 - \varepsilon) \frac{k_i}{k_{\max}}$.

Where P_i is the cost of the i -th scheme and k_i is the total washing times of the i -th scheme.

$$P_{\max} = \max P_i \quad k_{\max} = \max k_i .$$

7.2 Test Model 4

For problem 4, we need to classify the clothes that can be washed together. Here, we need to use the concept of graph theory. Assuming that the eight types of clothes given by the ques-

tion are eight vertices, We get an 8*8 matrix, $matrix[i][j] = \begin{cases} 0 \\ 1 \end{cases}$, when $matrix[i][j]=1$

means that class i clothes and class j clothes can be washed together, so as to obtain an adjacency matrix:

$$\begin{aligned} matrix = [\\ & [1,0,1,1,1,1,0,0], \\ & [0,1,0,0,0,0,0,1], \\ & [1,0,1,0,0,1,1,1], \\ & [1,0,0,1,0,1,1,1], \\ & [1,0,0,0,1,0,1,1], \\ & [1,0,1,1,0,1,1,1], \\ & [0,0,1,1,1,1,1,1], \\ & [0,1,1,1,1,1,1,1] \\ &] \end{aligned}$$

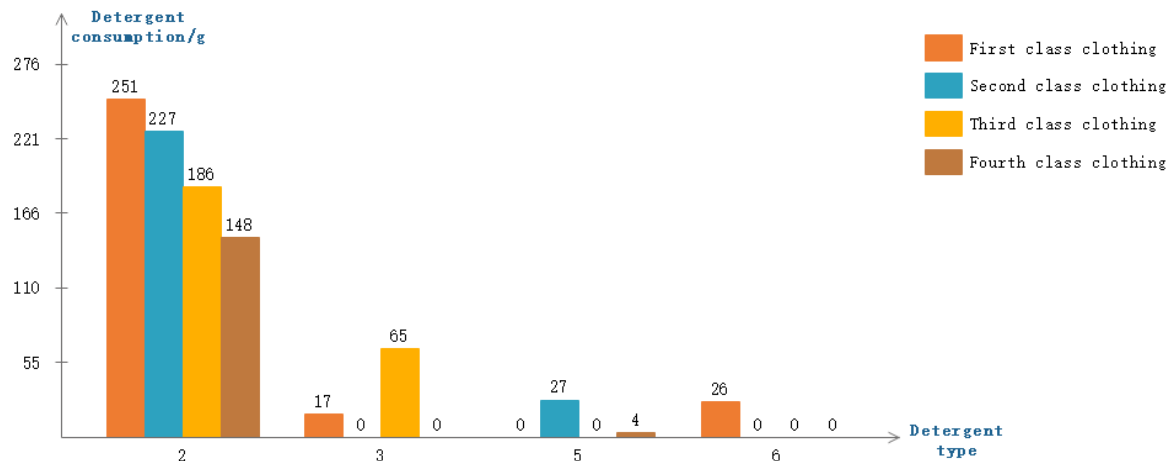
In order to minimize the number of washing times, we need to divide the eight categories of clothes into x categories, and x needs to be as small as possible. Therefore, we need to divide eight vertices into x clusters, and within each cluster, the vertices are connected to each other. Due to the small amount of data, we can solve the optimal classification way through the traversal method: $[[0,2,5], [1,7], [3,6],[4]]$.

Index starts at 0, so the categories given so far need to be all plus one.

After the initial data is processed, four basic categories are selected. For each basic category, the optimal number of classification groups can be selected by referring to the processing method of the third question through cluster analysis.

When ε is 0.8, the final solution is as follows:

Figure 9: The amount of detergent used for each type of clothing when it is divided into four types



As can be seen from Picture 9, cost-effective and efficient cleaning solutions are as follows:

- ✚ the first type of clothing is best with the second, third and sixth kinds of detergents;
- ✚ The second type of clothing is best with the second and fifth detergents;
- ✚ The third type of clothing uses the second and third detergent best;
- ✚ The fourth type of clothing is best with the second and fifth detergents.

The total cost is 90 yuan, and the total number of washes is 7.

8 Model Evaluation and Possible Improvements

8.1 Strengths

- ✧ This model takes into account the change of solubility with the amount of water, and the solution analysis of each problem is combined with sensitivity analysis, which is more comprehensive.
- ✧ This model algorithm has strong global search ability, can find more global optimal solutions, and also has strong robustness, can solve the problem under different objective functions and constraints, and get the optimal washing strategy.
- ✧ Aiming at problem 3, this model uses cluster analysis to classify clothes with similar dirt categories, and obtains a more effective cost saving effect in order to get a cleaning plan.

8.2 Disadvantages

The effect of the concentration of detergent in water on the washing effect is not considered, but in real life, the concentration of detergent may have an unnecessary effect on the washing effect.

8.3 Possible Improvements

- ◆ If we can refer to more papers in the part of solubility model, the analysis of solubility change can be more comprehensive.
- ◆ If more clothing stains and dissolver processing data are encountered, the solution of

the model may fall into the local optimal solution, and some special strategies need to be taken to avoid this situation, such as the design of a suitable state-generating function, which shows the total spatial dispersion or local region of the state according to the needs of the search process.

9 Conclusion

In order to solve the optimal cleaning scheme in different cases, this paper establishes a nonlinear programming model, solves by simulated annealing algorithm and cluster analysis, and uses control variables for sensitivity analysis, and finally obtains the optimal solution in accordance with various cases. According to the above analysis, when washing clothes, if the correct classification according to the type of clothing and dirt type, reasonable arrangement of the corresponding detergent, washing times and water consumption, can save the washing cost and improve the washing efficiency, which has a strong practical reference significance.

References

- [1] Wei Yuhui, Su Zhaowei, Huang Tianci, She Minchuchu, and Zheng Chen. Analysis of the Influencing Factors of Washing Cleanliness Rate of Firefighting Clothing Surface. Journal of Wuhan Textile University, 2022, 35(05): 58-63.
- [2] Ge Xian-Long, Wang Wei-Xin, Li Shun-Yong. Intelligent Algorithm and Application [M]. Southwest Jiaotong University Press, 2017
- [3] Jin, Xin, Jiawei Han, “K-Means clustering.” Encyclopedia of machine learning, edited by Claude Sammut and Geoffrey I. Webb, Springer, 2010, pp. 563-564