**Principle of Optimal Allocation and Its Application**

Yushi Mao

China Energy Research Society

China Academy of Railway Sciences

Beijing, August 1981

This article was originally published in August 1981, as a mimeograph. There were a total of 12 chapters with over twenty thousand Chinese words in the original text, and now in order to reduce the length, I compress it into eight chapters with 13, 000 Chinese words. The four chapters that were cut down are as follows: Chapter 5, Optimal Allocation and Macroeconomic; Chapter 6, Applications of the Principles of Optimal Allocation; Chapter 7, The Optimal Allocation of Reliability and Chapter 12, the application of Kukn Tucker optimal conditions to seek revenue function. This published edition strictly follows the original version without any change, except for the serial number of the formulas and charts.

# 1. The Principle of Optimal Allocation

Allocation problem occurs frequently in economic activities. The so-called allocation problem means allocating the limited supply of resources to a number of sectors in order to achieve the optimal economic return. The resources here may refer to energy, capital, foreign exchange, ore, timber, and even manpower, time and land.

In order to determine whether an allocation plan is better than another, there must be an evaluation criterion. It will make no sense to decide a better allocation plan if the evaluation criterion isn’t established. Establishing the evaluation standard is one of the fundamental problems in economics. This article, however, will not dwell on this issue. It will discuss the approaches to formulate the allocation plan on the assumption that the evaluation criterion has already been set up.

If there is only one sector in allocation, the best allocation amount is the amount achieving maximum return. If the revenue is given in advance in the allocation function, we are able to obtain the optimal input X\*, where the tangent is horizontal, as in Figure 1. If the quantity of resource is less than X\*, then obviously all the resources should be allocated as the increase of allocation will lead to revenue increase.

We can see from the example that, there is more than one method to calculate the “Return”. For instance, the allocation amount of ore can be the output of cast iron, the allocation amount of electric energy can be obtained from the output of the sector which uses the electric energy, and the allocation amount of funds can be the amount of profits. However, all these benefits must be net one, namely return after deducting cost. Otherwise, for unprofitable enterprises, the larger amount is allocated, the greater their loss will be, although the gross yield increases with the allocation.

Return

Best allocation

X \*

Allocation

Figure 1 Inputs and outputs—income and allocation

When there are several departments to participate allocation, it is obvious that we cannot get the optimal solution with a simple differential method. Consider the case of only two departments to participate in distribution, we first define the derivative of distribution x for the output g as , which is called the marginal rate of return, or marginal yield. For these two departments we have their own revenue function and . For any allocation scheme, the two departments have their marginal rate of return and , as shown in figure 2. If these two marginal output rate is not equal, the allocation scheme can be further improved. Assume adding a small incremental allocation on the first sector with a larger marginal output ratio and reducing allocation on the second sector with a smaller one. Then with the condition of the total quota does not change, the total output of the two departments can obtain an increment . So an allocation scheme is not optimal and can continue to improve when . Only when the two departments have the same marginal output, an allocation scheme can be optimal.

When there are more than two departments to participate in distribution, the above conclusion is also true. The reason is that if there is an optimal allocation scheme, any two departments should have an equal marginal output. Otherwise, we can improve the amount of total output by just adjusting the distribution of these two sectors. Thus, an equal marginal output for all departments is a necessary condition for the optimal allocation.

Now considering the output function as a concave one, as shown in figure 1 and 2, we improve allocation scheme through gradual adjustment. For the same allocation , when *x* is bigger, the marginal return is smaller, ie. if *a<b,*, the incremental return .

Output

I 

II 

  Input

Figure 2 Marginal productivity

We increase allocation for which has a largerto improve the distribution, when *x* increases, of corresponding departments must decrease. After we decrease the allocation of departments with larger, increases, which is a condition of concave function. We change allocation when . After adjustments the large become smaller as the small become larger. So finally marginal output ratios of all departments converge to the same value, i.e., where is the number of departments.

If we divide the total amount of resources X into several small parts, and [distribute](http://dict.youdao.com/w/distribute/) gradually to the department with best return. This is an optimal allocation. The department first assigned must be the one that has steepest gain function near the origin. And then we [distribute](http://dict.youdao.com/w/distribute/) to the department with the largest. But as  is micro-increment, the results of all allocations must be that each department has the same marginal output ratio. X experiences countless allocation scheme from begin to the end, and for each scheme, we have . They are the optimal allocation schemes. So, an equal marginal productivity is also a sufficient condition of optimal allocation. At last we can make a conclusion: for concave output function, it is a sufficient and necessary condition of optimal allocation that each department has an equal distribution of marginal output. And then we also assume output curve passes the origin, i.e. return is zero when allocation is zero.

Based on this conclusion, we can easily get the optimal allocation method for all departments. That is determining the slope of any tangent line as and making parallel tangents for the output function of all departments. The abscissa of each tangential point  is the amount of resources that the department deserves. If, we can change the value of, when the tangent slope get larger, since the function is concave, all the  become smaller. Otherwise, when  decreases all the get larger. Finally we can get .The realization of this process is very easy on computer. Another method to get the optimal allocation scheme is that we choose a scheme which meets  and optimize this program constantly. In other words, we reduce the amount of allocation of departments with smaller  and give them to departments with larger. When  of all departments are equal, we get the optimal allocation scheme.

If the output function is convex, the result of gradually improving distribution can’t converge to the marginal output ratio. The reason is that for the convex function, after increasing allocation for departments with larger, their  will get larger, and the others get smaller, and in the end, all resources will be distributed to one department. The allocation result is related to the initial scheme because all resources are distributed to one department and the relative order of all departments’  depends on the choice of the initial scheme. However, we can use the method getting the peak of concave function which is discussed before to get the minuteness of convex function. In an economic problem, the cost function is often convex for various factors of production (i.e. invested capital, labor, materials, production schedules, etc.). So we can get the optimal allocation of the factor of production in the condition that marginal rate of change in cost to allocation of factors of production is equal.

# 2. Duality Principle of Optimal Allocation

In the case that 0utput functions of various departments have been fully identified, a certain amount of resources x corresponds to a certain marginal income ratioand certain total output . So when x changes, so does and G. Define  as the set of resources amount, as unified marginal output ratio, as total output. Then there is a one to one mapping between these three sets, which is shown in Figure 3. We define a map, and a map. They are all determined by the concavity of the output function. For example, if there are resources of 100 units and the unified marginal output ratio for all departments is 0.3, then the total output is 70 units. If the resources increase to 101 units, as the function is concave, marginal income ratio must decrease, which for example we choose as 0.29, and the total revenue will increase with the increase of total amount of resources invested, which for example we choose as 71. This example illustrates that there is a one to one map between X, G and .

X= = G=



Figure 3 Relationship of X、、G sets

Correspondence between these three sets implies duality of Principle of Optimal Allocation, i.e. by unified marginal output ratio one total-resources amount can just correspond to one maximum total-output. Meanwhile, one total-output can just correspond to one minimum amount of total-resources. With different allocation schemes, one amount of total-resources can get different total-output. But there is only one maximum total-output, which corresponds to numbers in G obtained by unified marginal output ratio. Similarly, to achieve certain total-output, there is only one minimum amount of total-resources, which corresponds to numbers in G obtained by unified marginal output ratio.

The reason that X corresponds to the minimum resource required for G is shown by the above example. In order to obtain 71 units of total-output, minimal resource required cannot be larger than 101 units, for the reason that we indeed get 71 units of total output by only 101 total-resources from the map between  and . Meanwhile minimal resource required can not be less than 101 units, which is deduced by ascent stage of concave function, i.e. only when x increases will decrease. In ascent stage,  will be always positive no matter how it decreases, which means g changes in the same direction with x. Now x is reduced from 101, so g shall be reduced from 71. We can see minimal resource required is neither above 101 nor below 101, so it is 101. Then we have proved X corresponds to the minimum resource required of G. But this conclusion is true only in ascent stage of output function. If the amount of allocation is more than X\* in Figure 1, output function gets into descent stage, the conclusion is no longer applicable. But in practice output functions of problems are always in the ascent stage.

# 3. The Meaning of Unified Marginal Output

There is only one unified marginal output ratio for each certain output amount. In the energy distribution problems, the unit of marginal output ratio can be industrial output per ton of standard coal; in distribution of iron ore, it can be pig iron capacity produced per ton of ore; in the distribution of funds, it can be the interest rate, where  is marginal capital-output ratio (MCOR). When  is represented in the form of annual rate, its reciprocal is the investment payback period.

As there is a one to one map between X, G and, there is also a one to one map between MCOR and capital. When capital is adequate, the corresponding MCOR is low, so investment payback period can be relaxed. In the case of market regulation, an adequate supply of capital will reduce interest rates of the bank. Therefore, interest rate is actually linked with MCOR. If a company is engaged in business loans from banks, loan interest rate shall not exceed the capital-output ratio of enterprises. Otherwise, not only will it never be able to repay the principal, even the interest can not be paid.

Optimal allocation of capital construction investment should be measured with MCOR or investment payback period. A certain amount of total funds corresponds a certain investment payback period. Projects whose investment payback periods are longer than that should be postponed. Allocating investment by this simple principle, we can get an optimal allocation. Of course this is only measured in the economic view. Other factors like political and military ones should be considered in the decisions on other criteria.

Unified investment payback period mentioned above should be adjusted annually in accordance with the funds available. When capital is tight we can reduce unified investment payback period, that is, only invest the most profitable projects. When capital is adequate we can also invest suboptimal profitable projects.

Because the MCOR corresponds to the total amount of capital, the amount of loans a country obtains from abroad should correspond to the related interest. The higher the interest rate, the less the loan. Quantities of the rate and loan should be objectively determined by specialized research, and monitor the international financial dealings. Obviously, the loan amount not only corresponds to the interest, but also to the number N of the relevant projects participating in the distribution. Under the same conditions of the interest rate, the larger the N , which means more construction projects, the larger loan amount, and so as the total output revenue. Improving production opportunities and economic development advantages are the specific measures to increase N.

Revenue corresponds to the total amount of resources available. In the case of the distribution of capital, the total amount of profits determines the total amount of capital should be prepared. If the capital is not enough, the total amount of profits can not be reached. On the other hand, if the capital is too much, although profits can be reached, profit rate on capital will drop. We can not guarantee investment results. So estimating minimum amount of resources needed based on the intended target is an important task in economic planning.

The unified marginal output ratio  marks degree of scarcity of resources. The larger the , the more serious the shortage of resources, because a larger  means more output can be obtained with input resources, or more revenue losses as the lack of resources. If we use output value per ton of coals to indicate the degree of coal shortage, the number should be the unified marginal output ratio finally reached after an optimal allocation. If there is a lot of waste of coal nearby coal mines, and in the industrial area away from coal mines the coal is seriously shorted, marginal income ratio of the two places differ widely, which is not an optimal allocation, i.e. there is no unified marginal output ratio, so we can not measure the extent of the shortage of coal. Only by Optimal Allocation, distributing wasted coal to the coal-shorten area, can we estimate if the output value can be improved by increasing coal production, and how much the output value per ton of coal may increase.

The unified marginal output ratio can be used to measure the economic value of scarce supplies. For products that are not subject to resource constraints, they can be manufactured as long as there are enough capital and labor. But for some other products like oil, some non-ferrous metals, tropical crops and so on, their production has monopolistic nature due to resource and other constraints. So when the demand of them exceeds which the resource can be provided, their prices are times of the value determined under normal conditions. Under the conditions of capitalist production, this is the reason why generalized differential rent occurs. Under the conditions of socialism, how to correctly measure the value of such scarce supplies and make judgments in technical and economic comparison so that we can get an optimal allocation? We should calculate with the help of unified marginal output ratio in an allocation of such supplies to all departments. This unified marginal output ratio is their shadow price. We will discuss the mathematical definition and meaning of it later.

Due to the different natural conditions, product costs of socialist enterprises have considerable differences under the same management and level of effort. This unfair phenomenon must be corrected using resource tax. Generally, unified marginal output ratio can be used to decide the resource tax rate with which to be considered when comparable costs are to be properly formulated. Thus we can see, it is not the cost of production to be used in the technical and economic comparison, but the cost plus profit and resource taxes, or unified marginal output ratio reached finally in an optimal allocation. .

Marginal output ratio is the marginal ratio of input and output increments. When it is used to explain the scarcity of resources, we must note that with the increase in the supply of resources, marginal output ratio will decrease, which indicates that if the energy is in shortage we can not use indices like ‘increase of revenue per one hundred million tons inpout’ because it is far from a marginal ratio, but an average ratio.

# 4. Concavity of Revenue Function in Economic Activity

We can get a unified marginal input-output ratio in an optimal allocation only when the output function is concave. But in practice is the revenue function of economic activity always concave? Why? What if there are exceptions? This is subject to further discussion.

We believe that in general revenue function is concave, that is, output growing slower than the growth og input. The fundamental reason is that output depends on a number of factors. When any one of the factors fails to change along with others due to certain restrictions, the effect of input resources will be affected and as a result the revenue gradually decreases. Harvest crops can increase with the increasing amount of fertilizer applied. But it can not increase proportionally with the increasing amount of fertilizer when the land area is constant. The yield of plants is related to energy supply, but can not increase proportionally with energy supply due to limitations of plant equipment, labor, space, and so on. Limitations of space, time, manpower and resources are objective. They will eventually manifest in economic activity in some form.

Secondly, progressive allocation phenomenon in economic relationship also makes the revenue function concave. Assume investment allocation of agriculture, light industry, heavy industry to be , then for agriculture, can be distributed to food, forestry, research, and education and so on. Their return constitutes agricultural investment output expressed as. Then we can make an optimal allocation of  to these different departments, firstly allocate to the department with the largest return, then the second. So we can see when  increases, as the department allocated last, must be the one with the worst return, the growth of  gets more slowly, which indicates thatis concave. There is progressive allocation not only in agriculture, but also in the industry and transportation etc, even in a specific product. For example, when investing to improve authenticity of a TV parts, this investment will be firstly used on the procedure which is most sensitive to authenticity. As investment increases procedures not so sensitive to authenticity are considered. This example indicates that progressive allocation phenomenon is common. Optimal allocation of each level makes the revenue function of a previous level concave.

But not all revenue function is concave. For an example, the amount of work done by two people’s cooperation is more than twice of one person. Since the development of the social division of labor, the volume of output may increase more than input quantity under very large scale collaboration. Due to the improved management methods (like using an optimal allocation), or the improvement of technology, we can even increase revenue with no more input quantity. These will all make output function convex. From the point of view of the total process of development of human productivity, improved organization of production and technological progress are always the most active factors. However, in a specific period of time when the mode of production is relatively identified, output function is concave.

In a capitalist society, as businesses and individuals chase profits, and because of the market regulation, process of allocation will improve constantly. In the case of fund allocation, it becomes a spontaneous process of funding margins tend to unified. This phenomenon illustrates the revenue function of distribution of funds is indeed concave, because for a convex revenue function, allocation optimization results will not converge to the unified marginal input-output ratio, which is proved before. However,due to the lack of a unified planning and scheduling in capitalist production, its optimal process is under the conditions of opposing interests. Technological monopoly and confidential information prevent achieving optimal conditions. The actual allocation result is around the optimal allocation scheme with an oscillation. As a result of the huge monopolies, or simply due to random combination of adverse conditions, this oscillation amplitude may sometimes be great, which may lead to serious social consequences. Under production conditions of socialism, theoretically we do not have to rely on the market and the role of competition. Just by optimizing the program we can achieve the best economic results. However, due to the fact that the original data can not be very detailed and is impossible to be transferred without delay, mathematical programs of optimization problem are far from perfect. So we have to rely on the feedback effect of market regulation to supplement plan. The worse we do in the plan optimization work, the more we rely on market regulation. The higher the proportion of optimized adjustment plan exists in the economic regulation, the less random phenomena occur in economic activity. If we use the entropy to measure the degree of randomization in economic activity, improving and strengthening the role of regulation plan constantly to make economic activity more targeted, more cooperative and more perfect, is actually a constant process of entropy reduction. Therefore, the highly developed socialist production is the process of gradual strengthening of the effect of organizational optimization. Conversely, if we rely neither on plan optimization, nor the use of feedback mechanisms of market regulation, economic work must be in serious chaos and the consequences would be catastrophic. Recalling the history of the world economy with gradual commercialization dependence, many major economic crisis occur when these two phenomena exist, although the reason of these two phenomena may be economy itself or crucial political or military unrest.

# 5. The Principle of Optimal Allocation and Lagrange Multiplier Method

Resource allocation problems and reliability allocation problems discussed earlier are all problems to seek an extreme value under constraints. Such problems can generally be resolved by Lagrange multiplier method. Yet we do not rely on the same Lagrange multiplier method to obtain a solution, and economic significance of this method is obvious. In this section we will further prove general results of Lagrange Multiplier Method can be deducted with the principle of optimal allocation.

Lagrange Multiplier Method in a resource allocation problem is described as:

Maximize the objective function:、、……、)with 2、……、N) and subject to.

Solving by Lagrange multiplier method, we assume……，where the necessary condition that function L can get a extreme value is , i.e. .

Earlier we discussed the economic significance of optimal allocation. The main idea is: when resources among various departments make micro adjustments among each other (this adjustment cannot violate the specified constraints, i.e., the objective function can improve constantly. When we have reached optimal allocation no further improvement can be found. This is a necessary condition of optimal allocation. The direction of adjustment of the incremental amounts of resource allocation can be positive or negative, that is, we may give  from A department to B or conversely. If both adjustments do not make the objective function increase, or the objective function does not change, we must arrive the optimal allocation. In other words, if reducing from A changes the objective function with, then increasing to B will bring a change to the objective function, and their sum is zero, i.e. marginal output of A and B has the same margin with opposite signs

When the constraint is not a simple, the above deduction is still right. What is different is that whenreduces dx,  does not increase  but a trace which depends on specific requirements of constraints.

From the above discussion, we get a very important result of optimal allocation: the condition that marginal input-output ratios of all departments are equal is necessary condition of optimization problems. Specifically describes as follows: an optimization problem under constraints can be regarded as an allocation problem of each variable. When reaching an optimal allocation, any two variables make micro adjustment subject to constraints then marginal input-output ratio of the objective function has the same margin with opposite signs. Here we use this to prove the Lagrange multiplier method.

Maximize the objective function:、、……、) 5-1

subject to(、、……、 5-2

Let changes with, changes with , other  does not change. Ignore higher order terms of , variation of R can be defined as

 5-3

5-3 must be equal to zero, otherwise the constraint is violated.

From 5-3, we get that

 5-4

5-4 shows how muchshould change whenchange in order to satisfy the constraints.

The marginal output ofandfrom the objective function are  and . They are equal but with opposite signs, so

 5-5

Substitute 5-4, we get

 5-6

Where andare arbitrarily selected, so

2、……、N） 5-7

Also can be written as

 ( i = 1、2、……、N) 5-8

5-8 is the Lagrange conditions. It has N equations. With constraints 5-2, we can get all and . This solution is a necessary condition of the optimal allocation. It is a sufficient condition when the objective function is concave, or when we seek for the minimum, the objective function is convex.

5-6 also can be written as

 5-9

This formula clearly demonstrates the significance of the multiplier, which means the ratio between increment of the objective function and constraints when the constraints change a little, or the degree how much the objective function value depends on the constraints. As a result, it is known as sensitivity coefficients.  has economic implications in economic problems. For example, indicates resource constraints and  indicates economic returns, then  is shadow price of this kind of resource, i.e. how much returns can get with increasing a unit in . As we have discussed, the price of monopoly resources does not depend on cost plus average profit. Rather, it is related to the shadow price.

In a more generally constrained extremium problem, there are M constraints, i.e.

、、……、 5-10

、2、……M， MN

For each constraint, do the same derivation as above, presenting in a matrix form as:

 5-11

Let  5-12

5-11 multiply 5-12：

 5-13

i.e.

 5-14

or

 5-15

Then we get

 5-16

 5-17

here.

Re-write as algebraic formula：

 、2、……、N 5-18

Formula 5-18 are Lagrange optimization conditions for multiple constraints. So far we have completed all the proof. From the discussion above, we can prove Lagrange multiplier method using the principle of optimal allocation and vice versa. But the principle of optimal allocation demonstrates an explanation of economic sense to the conditions which optimal solution must meet. That is the objective function must have an equal marginal input-output ratio to all variables while change of variables is under the constraints.

An equal marginal input-output ratio is a very important concept. When there are no constraints, that is,, we can suppose there exists a constraint does not work, or the sensitivity coefficient of the objective function to constraints is zero. The necessary condition for the optimal solution become  when . This is the necessary condition of seeking an extreme value using ordinary differential method.

Recalling figure 3, setconsists of the variation of a coefficient in constraints. It corresponds to the objective function value set  via sensitivity coefficient set . In figure3 setconsists of the variation of constant term in constraints. In fact, any factor changes in constraints generate a set, and will correspond to  via . is the marginal input-output ratio. We can see that this marginal ratio corresponds to a definitive objective function value, no matter what the constraint is.

# 6. Application of Principle of Optimal Allocation in Continuous Allocation Problem

Within allocation problems discussed above, the distribution sector is countable. When there are infinite departments and allocation for each department infinitely approaches zero, it becomes a continuous distribution problem. Principle of Optimal Allocation can also be used to continuous distribution problem and get more general results. Solution of the continuous distribution problem has a very wide range of applications in engineering technology. This will be illustrated by the following three examples.

## 6.1 Distribution of speed of a train or car to save energy

We have known the resistance of train running is a convex monotonically increasing function of speed. Assume time of the train driving from station A to B is T. How should the driver control the speed of the train to make the energy to overcome the resistance in driving a minimum?

Now the objective function is, with a constraint

In the formula, S is the distance between two stations. We want to find out the relationship between speed and distance .

R

V

Figure4 Relationship between train resistance and speed

We divide S into equal N segments, each of a length of. So the problem is how to allocate T to these N segments. Assume time allocated to each segment is, then we get and within each segment the speed . When  is identified, the corresponding  is determined. From figure 4, as the distance of each segment is all , the relationship between  and  is the same curve for all segments. Relative to Figure 2 now there are many curves like I、II、……But these curves are coincident. We can use the principle of concave function maximization to solve a convex function minimization problem. In an optimal allocation marginal ratio is all the same, that is, the abscissas of tangent points of parallel tangents corresponding to the optimal allocation are the same. Now all the curves of are coincident and tangent points of parallel tangents are the same, so each segment should have a same speed, that is, all . The above discussion is also true when, where it becomes a continuous problem of distribution.

## 6.2 The production schedule with lowest cost

A problem similar to the previous example is: production cost is known as a convex function of production speed. Now we want to produce a quantity of products N in a certain time T. How to arrange the production schedule to make total cost of products minimal?

This problem corresponds exactly to the above example. It also allocates time T to N products. The relationship between cost and production schedule of each product is just like curves in figure 4, and the total cost is the sum of all products’ costs. From the above example, we know that the lowest cost production schedule is uniform production. Any change will increase the cost. No matter what the function between cost and production speed is and as long as the function is convex, the above conclusion is always true. Of course, here assumes that all other production conditions are unchanged.

Strictly speaking, the problem is different with the above example in which resistance is only a function of speed and has nothing to with acceleration. However production costs relate with production acceleration. When re-adjust the production schedule there will be extra cost. So the above example can be solved totally in a static allocation model while in this case we should make a model including the production adjustment costs. Nonetheless in this case the conclusion is obviously correct, because uniform production does not cause any extra cost with adjustment.

## Optimal allocation of boiler combustion rate

As we all know, heat absorbed in a boiler is a concave function of combustion rate (calculated by the weight of the fuel burning out per unit time per unit area of the grate). As each point throughout the grate has different ventilation, distance to heat transfer surface and pathway of generated gas, a certain amount of coal burning in different location producing different heat to the boiler. In other words, the heat G boiler can absorb is a function of burning rate u and coordinate position 



Now the problem is: in the constraint that the total amount of combustion U is certain, that is,

 6-1

we want to figure out the distribution functionwhich makes the total heat G that boiler absorbs reaches the largest where G is

** 6-2

In this example there is a basic difference compared with the optimal allocation problem discussed before. In the earlier problem, the revenue function g is only a function of allocation; however, in this problem g is not only a function of allocation u, but also of the location  of u. So function g cannot be indicated as the plane curve shown in figure1, but a surface in space.

y

x u

y

x

Figure 5 Optimal distribution of burning rate

These problems can be solved by using the model of variation problem having isoperimetric conditions, and the result is to solve a differential equation degenerating into a general algebraic equation. From the discussion of the principle of optimal allocation, we know that the optimal solution must have an equal marginal productivity. Now the marginal productivity is , so the optimal solution must satisfy

 6-3

Whereis figured out from constraint 6-1. The larger the u, the smaller the. U is equivalent to the constraint of total amount of resources discussed before.

6-3 has its economic implications. It shows in an optimal allocation, putting a little coal on wherever a small area  of grate can make the boiler heat absorption increase by the same amount. So if one area has a larger, we should put the extra coalon it. By continuous adjustment, as functionis concave, finally all marginal return of combustion rate is equal. Now function  has been known, so 6-3 is the solution of this problem. This form of solution requests the distribution of resource u is a function changing with coordinates, which is different with the result discussed before. Further promotion makes the application of the principle of allocation have more possibility.

Application of the principle of optimal allocation in continuous distribution problems is mainly used for solving the problem in engineering technology.

# 7. Application of Principle of Optimal Allocation in Dynamic Problem

The fundamental difference between dynamic and static problems is that optimal solution of the dynamic problem not only considers the return of decisions, but also relates to the condition of prescript starting and ending points. When these condition change, the optimal solution changes accordingly. Dynamic problems can often be divided into several phases in the space and time. Adopting different strategies in each phase affects not only the output of this phase, but also the ending state, and then affects the solution in the next phase. Take national economic planning as an example, conditions of starting point are economic parameter at the start of planning, like fixed assets and population. The purpose of planning may be that people spending within the total five-year plan is maximum and fixed assets and population reaches a predetermined target in the end of the phase. We need to figure out an optimal savings rate and a population growth rate. Obviously, the more fixed assets are required to achieve in the end of the phase, the less the total consumption of the people would be in this phase, but this creates the conditions to improve the living in next phase. Therefore, the savings rate should be chosen not only to make the objective function maximum, but also to guarantee economic parameters at the end of this phase. This kind of economic planning problem often has a series of constraints, like constraints of resource and economic ties between departments. Economic planning is a very complex issue. Now we will discuss the simplest case of dynamic problem, the brachistochrone. This is a typical problem tackled with classical variation method. It requires the solution of an equation of the curve from the start point  to the end point , along which a particle sliding by gravity from the start point to the end takes the least time. Here we ignore the friction.

The straight line connecting is the shortest line between these two points. But the time cost may not be the least, because particle does not get a big acceleration at the starting point. If moving vertically downward in the beginning, the particle’s acceleration is the largest, but the distance traveled increases, too. So the curve we need may be between the two cases above. The brachistochrone theory can be used in the design of hump vertical section in marshalling stations with a restriction of coupler angle between adjacent vehicles.



x





 

y 







Figure 6 The brachistochrone problem

From figure6, assume  is the curve we seek. The curve passes through  and . According to physics theorem, the speed on  is , whereis the ordinate of . According to the principle of integral calculus, the travelling time  from  to  is:

 7-1

Or be written as the form of sum

 7-2

Generally consider the integrand as a normal function, then

 7-3

Treat the brachistochrone problem as an optimal allocation problem. We will naturally regard  as the amount of resources held. Because is the vertical distance, which provides the speed for the particle. Now the problem is how to allocate vertical distanceto all segments so that the return function is maximum. Once the vertical distanceallocated to each segments is identified, we can get the curve. Since the total vertical distance is limited, so the problem is subject to a constraint when gets the optimal solution:

 7-4

At the same time

、1、……、N-1) 7-5

7-5 is also a constraint.

7-3、7-4、7-5 have constituted the brachistochrone of variation problem using optimal allocation theory. But it is still not a common form of variation problem, as integrand or function of 7-2or7-3 does not include, which is to say that coordinate positions of income and distribution sectors have nothing to do with . In a more general variation problem integrand or return function are related to ，，at the same time. We should choose the objective functionas:

 7-6

Boundary conditions are

 7-7

、、、 are given.

Now we can understand the variation problem in the sense of distribution and return output and use the principle of optimal allocation to solve the problem.  is the return function of one of unlimited departments to be allocated.  is the cumulative amount of resources invested. The new total amount of resources invested is limited to. Coordinate position of the department to be allocated is between  and .

Write the objective function in the form of summation:

 7-8

Constraints 7-4 and7-5 are still holding.

So we transfer the unconstrained variation problem to a generalized resource allocation problem with constraints.

From the physical or economic significance of the problem, we assume that this polynomial converges in the range of variation ofand. Now we use Lagrange multiplier method constituting a new function:

 7-9

Here all are different. They change with (or) and is a function of .

To figure out an extreme value of, we can change eachor. Although from 7-5 they are essentially the same, they have different forms in math. We will discuss from these two aspects. Assume partial derivative of toandexists, and 7-9 is uniformly convergent. Firstly we change  to makereach extremum, that is:



or、1、2……、N-1) 7-10

7-10 is true for all intermediate point, so it can be written as

 7-11

Hereis a function of

Then we changeto makereach extremum, we can get



、1、2……、N-1) 7-12

When summing up  in , all items with a subscript greater than or equal tocontains. So 7-12 has N equalities. Take the limit:

 7-13

Assume the primitive of integrandis, then 7-13 can be written as:

 7-14

In 7-14, the first term is a constant; the second and the third term are valid for all points  in . Finding the derivative of 7-14 toand noting  is a constant, we get:

 7-15

Find the derivative of 7-14 to:

 7-16

Substituting the 7-15 to 7-16, we get Euler equation:

 7-17

So far, we manage to get Euler equation by principle of optimal allocation. Principle of optimal allocation can used to solve both static problem and dynamic problem. Resource constraints may be understood broadly; isoperimetric conditions, boundary conditions can be treated as resource constraints. So this principle can be used to many problems in different descriptions. Optimal allocation gives a common explanation for different optimization problems. We consider it a universal basis of optimization method, so it is called Principle of Optimal Allocation.

7-11 shows that marginal output ratio  of return functionis no longer a constant, but a function  of .

The above derivation also indicates a way to find numerical solutions of the Euler equations. 7-13 can be written as:

 7-18

So,

 7-19

But  is constant and can merger with. Assume

 7-20

Write 7-19 as a limit summation, we get:

 、1、2……、N-1 7-21

Combined 7-21, 7-5 and 7-10, when set a , for each  we can solve ,  and . Proceeding gradually till , and then all can be solved. If allmeet, then the set of is right, otherwise we will reset a value ofuntil all meet. In actual calculation, we can plot the curve of with regard to. Choosingin the location where will reduce computational workload of trial-and-error. Each corresponding curve  calculated tois an optimal solution. What’s different is only their end condition.

# 8. Economic Meaning of the Euler Equation

In the research of brachistochrone we regard the vertical distanceas resource. Allocating in the abscissa, we can get the solution. Generally, when regarding the vertical distanceas resource and deriving Euler equations with the principle of optimal allocation, as discussed before, we need to use Lagrange multiplier method. In fact, according to the result of optimal allocation, that is, an equal marginal output is the necessary condition of an optimal solution, we can get Euler equations directly. The description of equal marginal output is shown in Section 5. Now we use it in the dynamic problem, and get Euler equations of variation method directly.

Now we discuss the fundamental variation problem, it is regulated by 7-6 and 7-7. Assume in figure 7,  and  are two given points, and is the optimal solution of 7-6.  is an arbitrary point on the optimal curve. The objective function 10-6 can be written as the form of a summation:

 8-1

Any point on the curve passing through  and  has a coordinate, and so. Substituting these three values into we can get a micro-unit of . But only in the optimal curve can the sum of 11-1 be maximum (or minimum).

Refer to Figure 7. At any point  on the extreme curve, keeping its ordinate (x,y), only change  with a minute quantity. The micro-unit  of  on point  will change accordingly, which is ignoring higher order terms:

 8-2

y

 

 

 

 

x

Figure 7 Deriving Euler equations with the concept of equal marginal output

This is the impact on the objective function due to the change of  along the direction of the curve at the point, which can also be treated as the marginal output. But as the direction of the curve at  has changed, ordinates of the curve after  also change. At first the abscissa increases and the ordinate changes. Now will change, which is the second-order difference of. So the change of micro unit is:



It is so for every micro units after, so the total amount change ofis

 8-3

This is a marginal loss due to the changes of the direction of the curve at, as constraint is that the curve must pass finally.

If is an extreme curve, we can assert that marginal output must be equal to marginal losses. If they are not equal, this change will produce net increase, i.e. can still be improved, and then it is not an extreme curve. So

 8-4

Extreme conditions above are valid for all, so specific point can be exchanged for any point on the curve. Further as, so

 8-5

Differentiate and note thatis a constant, so

 8-6

This is the classical Euler equation for the calculus of variation.

From the derivation above we can see that the Euler equations have its economic implications. It shows that the extremum curve has a characteristic, that is, the marginal return (loss) gotten by changing direction slightly at any point of the curve will be offset just by the marginal loss (return) due to changes of the coordinates of the curve afterwards. At any point of the curve, any direction of change cannot improve the objective function.

**Reference**

Bellman, R, and Dreyfus, S. Applied Dynamic Programming. Princeton, New Jersey: Princeton University Press, 1962.