Constrained optimisation

Problems in economics typically involve maximising some quantity, such as utility or profit, subject to a constraint – for example income. We shall therefore need techniques for solving such constrained optimisation problem.

Typically, we will have an objective function $F(X_1,X_2,\ldots,X_n)$, where $X_1\ldots X_n$ are the choice variables, and one or more constraint functions $G_1(X_1,X_2,\ldots,X_n),\ldots G_k(X_1,X_2,\ldots,X_n)$. The problem is typically formulated as:

Maximise/Minimise $F(X_1,X_2,\ldots,X_n)$ subject to $G_1(X_1,X_2,\ldots,X_n)\leq 0$, $G_2(X_1,X_2,\ldots,X_n)\leq 0,\ldots, G_k(X_1,X_2,\ldots,X_n)\leq 0$.

In this section, we will consider techniques for solving problems of this type.

Constrained optimisation in one variable

We will start by considering constrained optimisation problems in one variable.

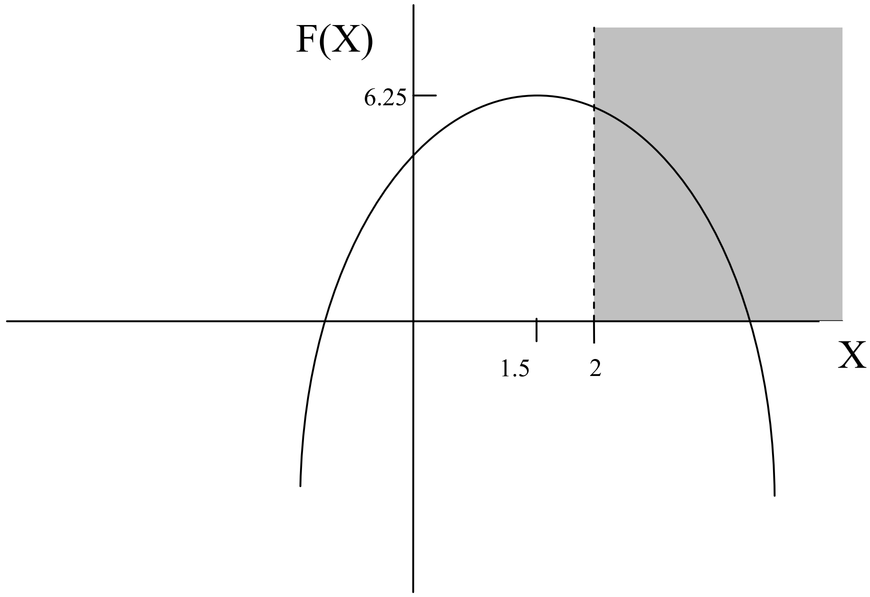
For example, consider the problem:

Maximise $F(x)=4+3x-x^2$

Subject to the condition $x\leq 2$

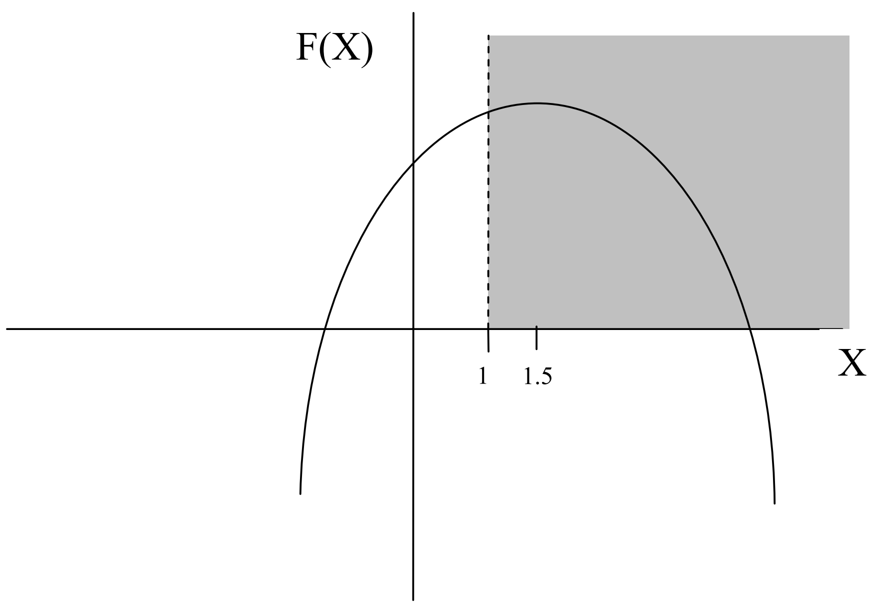
We can rewrite the constraint as $G(x)=x-2\leq 0$, to get it into the form described above.

We can easily solve this problem using differentiation, and see the solution graphically:



We have that dF/dx=3-2x. Setting this to 0 gives x=1.5, F(x)=6.25, and consideration of the second differential shows this is a local maximum. The second differential is equal to -2, so the function is concave for all real values, so this is a global maximum. Finally, the resulting value of x is within the constraint, so that this is the solution to the constrained optimisation problem as well as to the unconstrained problem.

In this case, the constraint, x⩽2, is <u>non-binding</u> or <u>slack</u>. Suppose that instead we had imposed the constraint G(x)=x-1≤0, i.e. x⩽1



We can now see from the graph that the optimum solution is x*=1, giving F(x)=6. This time the constraint is binding. Although it is easy to see what is happening in this case, in general we need to be able to distinguish between binding and non-binding constraints.

Constrained optimisation in more than one variable: the method of Lagrange Multipliers

The most important method for solving constrained optimisation problems in more than one variable is the method of Lagrange Multipliers.

Consider the problem of a consumer seeking to maximise their utility subject to a budget constraint. They must divide their income M between food (F) and clothes (C), with prices P_F and P_C, so as to maximise the following 'Stone-Geary' utility function:

$U(F,C) = \alpha \mathrm{Ln}(F-F_0) + (1-\alpha)\mathrm{Ln}(C-C_0)$

So their budget constraint can be written

$G(F,C) = P_F F + P_C C - M = 0$

Our problem is to maximise U(F,C), subject to the constraint G(F,C)=0.

To solve this we introduce an auxiliary variable λ, the Lagrange Multiplier, and form the Lagrangian function

$L(F,C,\lambda) = U(F,C) - \lambda G(F,C)$

To maximise U(F,C) subject to our constraint, we instead solve the *unconstrained* maximisation problem for $L(F,C,\lambda)$.

To do this, we must set all three partial derivatives to zero. Thus,

1) $\frac{\partial U}{\partial F} = \frac{\alpha}{F - F_0} - \lambda P_F = 0$

2) $\frac{\partial U}{\partial C} = \frac{(1-\alpha)}{C - C_0} - \lambda P_C = 0$

3) $\frac{\partial U}{\partial \lambda} = P_F F + P_C C - M = 0$

The third condition is of course simply the original constraint.

It is worth taking a moment to look at the economic significance of this approach. We can rewrite equations 1) and 2) to say that $\delta U/\delta F = \lambda P_F$ and $\delta U/\delta C = \lambda P_C$, whereupon, eliminating λ, we get:

$$\frac{\partial U}{\partial F} \Big/ \frac{\partial U}{\partial C} = \frac{P_F}{P_C}$$

In other words, that the ratio of marginal utilities to price must be the same for both goods. This is a familiar result from elementary consumer choice theory, and illustrative of a general economic principle: an economic quantity (utility, output, profits, etc.) is optimised where the ratio of marginal benefits of different uses of resources is equal to the ratio of marginal costs.

Solving, we obtain:

$$F = F_0 + \frac{\alpha(M - P_C C_0 - P_F F_0)}{P_F}$$

$$C = C_0 + \frac{(1-\alpha)(M - P_C C_0 - P_F F_0)}{P_C}$$

Which says that, after the minimum quantities C_0 and F_0 have been bought, remaining spending is allocated in the proportions α:(1-α) between food and clothing – this is of course a particular property of this utility function, rather than any general law.

We obtain $\lambda = \frac{1}{M - P_F F_0 - P_C C_0}$

What does λ signify? Well, if we feed back our solutions for F and C into the Utility function, we find that

$$U^* = \alpha Ln\left(\frac{\alpha(M - P_C C_0 - P_F F_0)}{P_F}\right) + (1-\alpha)Ln\left(\frac{(1-\alpha)(M - P_C C_0 - P_F F_0)}{P_C}\right)$$

Which can be rearranged to give

$U^*=\alpha(Ln(\alpha)-Ln(P_F))+(1-\alpha)(Ln(1-\alpha)-Ln(P_C)) + Ln(M-P_FF_0-P_CC_0)$

Whereupon $\delta U^*/\delta M = 1/(M-P_FF_0-P_CC_0) = \lambda$.

Thus, λ gives the marginal utility from extra income. More generally, the Lagrange Multiplier λ gives the *marginal increase in the objective function from a unit relaxation of the constraint.*

Lagrange multipliers; a formal treatment

We now extend the treatment of Lagrange Multipliers to functions of several variables, and to allow for both non-negativity constraints and non-binding constraints.

Thus, we consider the following problem:

Maximise $F(X_1,\ldots,X_n)$ subject to

$G_1(X_1,\ldots,X_n)\leqslant 0$
….
$G_k(X_1,\ldots,X_n)\leqslant 0$

$X_i\geqslant 0$, for each $i=1,\ldots,n$.

Thus we have n variables, and k constraints, each of which may be binding or non-binding. We also have n non-negativity constraints.

We form the Lagrangian:

$L(X_1,\ldots,X_n,\lambda_1,\ldots,\lambda_n)=F(X_1,\ldots,X_n)-\lambda_1G_1(X_1,\ldots,X_n)-\ldots-\lambda_kG_k(X_1,\ldots,X_n)$

Note there are now k Lagrange Multipliers, one for each constraint. The Kuhn-Tucker theorem states that, at the optimum solution, $(X_1^*,\ldots,Xn^*)$ where F takes its maximum value, there exist values for $\lambda 1^*,\ldots,\lambda n^*$ for $\lambda 1,\ldots,\lambda n$ such that:

1) For each X_i, $\frac{\partial F}{\partial X_i} \leq 0$, with equality if $X_i>0$
2) For each $j=1,\ldots,k$, $G_j(\lambda_j^*)\leqslant 0$, $\lambda_j^*>0$, and *either* $\lambda_j^*=0$ or $G_j(\lambda_j^*)=0$.

The second condition is worth looking at more closely. It says that, first of all, the Lagrange multiplier must always take a positive value (this is natural if we consider the role of the LM as the marginal benefit from relaxing the constraint – this must be positive); secondly, that the constraint must be satisfied; and thirdly that either the constraint must be *just* satisfied (a binding constraint), or the value of the LM must be zero, in which case we have a slack constraint. Again this is natural, since if the constraint is slack, then there is no marginal benefit from relaxing it.

Note that these are *necessary* conditions for the existence of a local maximum. It is possible to state *sufficient* conditions that specify cases when we can guarantee that a point that satisfies conditions 1) and 2) will be a global maximum, but these conditions are quite complex, and beyond the scope of this course.

In general, it may be necessary to look at all the different possible combinations of binding and slack constraints, and of boundary and interior solutions.

Exact constraints

If one of the constraints is *exact*, that is requiring $G(X_1,\ldots,X_n)=0$, then condition 2) for this constraint does not apply, instead it is required, of course, that the constraint is satisfied.

Non-negativity conditions

We have framed the problem on the assumption that all the variables must be non-negative. If a particular variable X_i does not have to be non-negative, then condition 1) for that variable simply becomes $\delta L/\delta X_i=0$.

Constrained minimisation

We have formulated the Kuhn-Tucker theorem in terms of maximising a function. Of course, it is easy to minimise a function $F(X_1,\ldots,X_n)$ by maximising $-F(X_1,\ldots,X_n)$. However, more usually, we solve a minimisation problem by forming the Lagrangian as $L(X_1,\ldots,X_n,\lambda_1,\ldots,\lambda_k)=F(X_1,\ldots,X_n)+\lambda_1 G_1(X_1,\ldots,X_n)+\ldots+\lambda_k G_k(X_1,\ldots,X_n)$, and proceeding as above.

Example

A manufacturing firm produces two models of Widget, A and B. Let X and Y denote the quantity of models A and B produced in a week respectively. Model A requires 2 hours of machine time per item, while model B requires 1.5 hours of machine time. Each hour of machine time costs £2, whether for type A or type B. The total labour and material costs for producing X units of type A is $4X-0.1X^2+0.02X^3$, while for Y of type B, the cost is $4.5Y-0.1Y^2+0.02Y^3$. The two are strong substitutes, so that the demand curve for types A and B are given by $X = 80 – 0.5P_A + 0.3P_B$ and $Y = 70 + 0.25P_A – 0.4P_B$, where P_A and P_B are the price in pounds of A and B respectively.

The two constraints on (short-term) production are firstly, that there is only a maximum of 80 hours available machine time per week, (The rest being required for maintenance), and that the firm is under contractual obligations to produce a total of at least 40 widgets per week.

What is the optimal quantity of types A and B for the firm to produce to maximise profits?

First of all, we solve the demand functions to work out price in terms of X and Y, giving P_B=220-X-2Y, and P_A = 292-2.6X-1.2Y. Thus, total revenue is equal to $292X-2.6X^2+220Y-2Y^2-3.2XY$.

Total costs (machining, labour and materials) come to $8X-0.1X^2+0.02X^3 + 7.5Y – 0.1Y^2 + 0.02Y^3$. Hence, we can write the profit function as:

$$\Pi(X,Y)=284X-2.5X^2-0.02X^3+212.5Y-1.9Y^2 -0.02Y^3 – 3.2XY$$

The constraints on machine time and production give: (putting them in the required form)

$$G_1(X,Y)=2X+1.5Y-80\leqslant 0$$
$$G_2(X,Y)=40-X-Y\leqslant 0$$

We also have the non-negativity constraints $X\geqslant 0$ and $Y\geqslant 0$, as we can't have negative production.

We form the Lagrangian

$$L(X,Y,\lambda,\mu)= \Pi(X,Y)=284X-2.5X^2-0.02X^2+212.5Y-1.9Y^2 -0.02Y^3–3.2XY – \lambda(2X+1.5Y-80) – \mu(40-X-Y).$$

We thus have the conditions:

1) $\delta L/\delta X = 284 - 5X - 0.06X^2 - 3.2Y - 2\lambda + \mu \leqslant 0$, with equality if $X>0$.

2) $\delta L/\delta Y = 212.5 - 3.8Y - 0.06Y^2 - 3.2X - 1.5\lambda + \mu \leqslant 0$, with equality if $Y>0$

3) $\lambda \geqslant 0$, $G_1(X,Y) \leqslant 0$, and either $\lambda=0$ or $G_1(X,Y)=0$
4) $\mu \geqslant 0$, $G_2(X,Y) \leqslant 0$, and either $\lambda=0$ or $G_2(X,Y)=0$

Let us start by looking for interior solutions, so that $X,Y>0$.

Let us also start by looking for solutions where both constraints are slack, that is where $\lambda=\mu=0$. Solving some ugly equations for conditions 1) and 2) gives $X = 22.37$, and $Y= 26.22$. (Ignoring the fact that you can't have non-integer quantities of widgets for now). However, this does not satisfy the constraint on machine time, so this is impossible.

Let us now consider solutions where the first constraint is slack, so $\lambda=0$, but the second is binding, so $X+Y=40$, and $\mu \geqslant 0$. Conditions 1) and 2) now become

$284 - 5X - 0.06X^2 - 3.2(40-X) + \mu = 0$

So that

3) $156 - 1.8X - 0.06X^2 + \mu = 0$

And $212.5 - 3.8(40-X) - 0.06(40-X)^2 - 3.2X + \mu = 0$

So that

3) $-35.5 + 5.4X - 0.06X^2 + \mu = 0$

Which gives

$191.5 - 7.2X = 0$, so $X = 26.6$, whereupon $Y = 13.4$. This satisfies the constraint on machine time, and also the non-negativity conditions. We must check that it gives a positive value for μ. With these values, $\mu = -35.5+5.4*26.6-0.06*26.6*26.6 = -51.3<0$. Hence this violates the condition that the LM be non-negative, so it is not a possible solution.

We can now consider the possibility that the machine-time constraint is binding, so that $2X + 1.5Y = 80$, but that the production constraint is slack, so that $\mu=0$ and $X+Y \geqslant 40$. We now have

1) $284 - 5X - 0.06X^2 - 3.2Y - 2\lambda = 0$ and
2) $212.5 - 3.8Y - 0.06Y^2 - 3.2X - 1.5\lambda = 0$.

Substituting using $2X + 1.5Y = 80$, so $X = 40 - 0.75Y$ gives

3) $2\lambda = -.03375Y^2 + 4.15Y - 12$ and
4) $1.5\lambda = -.06Y^2 - 1.4Y + 84.5$

Which gives $.04625Y^2 + 6.01667Y - 124.6667 = 0$

One solution is negative, the other gives $Y = 18.18$, whence $X = 26.365$. We need to confirm that this gives a non-negative value of λ. These values give $\lambda = 52.29$, which is OK. Hence, $(X,Y) = (26.365,18.18)$ is a possible solution to our optimisation problem.

We may now suppose both constraints are binding, so that $X+Y=40$ and $2X+1.5Y=80$. This is only possible with $X=40$ and $Y=0$. Then conditions 1) and 2) become

$284 - 200 - 96 - 2\lambda + \mu = 0$, so $-12 - 2\lambda + \mu = 0$ and

$212.5 - 128 - 1.5\lambda + \mu = 0$, so $84.5 - 1.5\lambda + \mu = 0$.

Hence, $0.5\lambda + 96.5 = 0$, giving a negative value for λ, which is impossible.

We have thus exhausted all possibilities for internal solutions, the only one being $(X,Y) = (26.365,18.18)$. We may now try for boundary solutions. We may first try $X=Y=0$, which gives the conditions:

1) $284 - 2\lambda + \mu \leqslant 0$
2) $212.5 - 1.5\lambda + \mu \leqslant 0$

Again, we may consider binding or non-binding constraints. If both are non-binding, so that $\lambda=\mu=0$, then clearly 1) and 2) are not satisfied. However, if either constraint is binding, then X or Y must be strictly positive, which contradicts our assumption.

What if $X=0$, but $Y \geqslant 0$? In that case, the conditions become

1) $284 - 3.2Y - 2\lambda + \mu \leqslant 0$
2) $212.5 - 3.8Y - 0.06Y^2 - 1.5\lambda + \mu = 0$

Let us try both constraints non-binding, so that $\lambda=\mu=0$. This gives Y = 35.75 as the non-negative solution to 2), but that fails to satisfy 1). If the machine constraint is binding but the production constraint is slack, then 2) gives a negative value for λ, which is impossible. If the production constraint is binding but the machine constraint slack (so Y=40 and $\lambda=0$), then 2) gives $\mu=35.5$, but this fails to satisfy 1). Finally we cannot have both constraints binding, as then X is positive. Hence, there is no solution with X=0 but $Y \geqslant 0$.

Finally, we may look for solutions where $X \geqslant 0$ and Y=0. Our first two conditions now become

1) $284 - 5X - 0.06X^2 - 2\lambda + \mu = 0$
2) $212.5 - 3.2X - 1.5\lambda + \mu \leqslant 0$

The constraints must now be either both binding or both slack, as they are both precisely satisfied when X=40. In this case, the conditions become

$-12 - 2\lambda + \mu = 0$
$84.5 - 1.5\lambda + \mu = 0$

But this makes λ negative. If both constraints are slack, so that $\lambda=\mu=0$, we have from 1) that X = 38.77. However, this fails to satisfy condition 2).

Thus, we have ruled out all possible boundary solutions, leaving only the interior solution (X,Y)=(26.365,18.18), where the machine-time constraint is binding, and the production constraint is slack. As this is the only possibility, and as there logically must be some profit-maximising combination of outputs subject to the constraint (as infinite profits are clearly impossible), then this must in fact be the global maximum solution.

This has been a rather cumbersome process of checking all possibilities. In fact, consideration of the properties of the function would enable us to rule out a lot of the possible solutions very easily, but this would take rather more theoretical machinery to demonstrate. You are not likely to meet such awkward cases in your MA programme, but this example illustrates how the process can be carried out if necessary.