Chapter 2. Lagrange's Method

In this chapter, we will formalize the maximization problem with equality constraints and introduce a general method, called *Lagrange's Method* to solve such problems.

2.A. Statement of the problem

Recall, in Chapter 1, the maximization problem with the equality constriant is stated as follows:

$$\max_{x_1 \geq 0,\, x_2 \geq 0} U(x_1, x_2)$$
$$\text{s.t. } p_1 x_1 + p_2 x_2 = I.$$

In this chapter, we will temporarily ignore the non-negativity constraints on x_1 and x_2[1] and introduce a general statement of the problem, as follows:

$$\max_{x} F(x)$$
$$\text{s.t. } G(x) = c.$$

x is a vector of choice variables, arranged in a column: $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$.

As in Chapter 1, we use $x^* = \begin{pmatrix} x_1^* \\ x_2^* \end{pmatrix}$ to denote the optimal value of x.

$F(x)$, taking the place of $U(x_1, x_2)$, is the *objective function*, the function to be maximized.

$G(x) = c$, taking the place of $p_1 x_1 + p_2 x_2 = I$, is the constraint. However, please keep in mind that in general, $G(x)$ could be non-linear.

2.B. The arbitrage argument

The essence of the arbitrage argument is to find a point where "no-arbitrage" condition is satisfied. That is, to find the point from which any infinitestimal change along the constraint does not yield a higher value of the objective function.

[1] We will learn how to deal with non-negativity in Chapter 3.

We reiterate the algorithm of finding the optimal point using the general setup:

(i) Start at any *trial point*, on the constraint.

(ii) Consider a small change of the point along the constraint. If the new point constitutes a higher value of the objective function, use the new point as the new trial point, and repeat Step (i) and (ii).

(iii) Stop once a better new point could not be found. The last point is the optimal point.

Now, we will discuss the arbitrage argument behind the algorithm and derive the "non-arbitrage" condition.

Consider an initial point $x^0 = \begin{pmatrix} x_1^0 \\ x_2^0 \end{pmatrix}$, and the infinitesimal change $\mathrm{d}x = \begin{pmatrix} \mathrm{d}x_1 \\ \mathrm{d}x_2 \end{pmatrix}$.

Since the change in x is infinitesimal, the changes in values could be approximated by the first-order linear terms in Taylor series. Using the subscripts to denote the partial derivatives, we have

$$\mathrm{d}F(x^0) = F(x^0 + \mathrm{d}x) - F(x^0) = F_1(x^0)\mathrm{d}x_1 + F_2(x^0)\mathrm{d}x_2, \tag{2.1}$$

and

$$\mathrm{d}G(x^0) = G(x^0 + \mathrm{d}x) - G(x^0) = G_1(x^0)\mathrm{d}x_1 + G_2(x^0)\mathrm{d}x_2. \tag{2.2}$$

Recall the concrete example in Chapter 1,

$$F_1(x) = MU_1 \text{ and } F_2(x) = MU_2;$$

$$G_1(x) = p_1 \text{ and } G_2(x) = p_2.$$

We continue applying the argitrage argument with the general model. The initial point x^0 is on the constraint, and after the change $\mathrm{d}x$, $x^0 + \mathrm{d}x$ is still on the contraint. Therefore, $\mathrm{d}G(x^0) = 0$. From (2.2), we have the following equation

$$G_1(x^0)\mathrm{d}x_1 = -G_2(x^0)\mathrm{d}x_2.$$

Let $G_1(x^0)\mathrm{d}x_1 = -G_2(x^0)\mathrm{d}x_2 = \mathrm{d}c$. Then

$$\mathrm{d}x_1 = \mathrm{d}c/G_1(x^0) \text{ and } \mathrm{d}x_2 = -\mathrm{d}c/G_2(x^0). \tag{2.3}$$

Substituting into (2.1), we have

$$\begin{aligned} \mathrm{d}F(x^0) &= F_1(x^0)\mathrm{d}c/G_1(x^0) + F_2(x^0)\left(-\mathrm{d}c/G_2(x^0)\right) \\ &= \left[F_1(x^0)/G_1(x^0) - F_2(x^0)/G_2(x^0)\right]\mathrm{d}c. \end{aligned} \tag{2.4}$$

Since we do not impose any boundary for x, so x^0 must be an *interior point*, and $\mathrm{d}c$ could be of either sign. Therefore,

1. If the expression in the bracket $F_1(x^0)/G_1(x^0) - F_2(x^0)/G_2(x^0)$ is positive, then $F(x^0)$ could increase by choosing $\mathrm{d}c > 0$.
2. Similarly, if the expression in the bracket is negative, then $F(x^0)$ could increase by choosing $\mathrm{d}c < 0$.

The same argument holds for all other interior points along the constraint. Therefore, for the interior optimum x^*, we must have

$$F_1(x^*)/G_1(x^*) - F_2(x^*)/G_2(x^*) = 0 \implies F_1(x^*)/G_1(x^*) = F_2(x^*)/G_2(x^*) \tag{2.5}$$

Equation (2.5) is the “non-arbitrage” condition we are looking for.

It is important to distinguish between the interior optimal point x^* and the points that satisfy (2.5). The correct statement is as follows:

Remark. If an interior point x^* maximizes $F(x)$ subject to $G(x) = c$, then (2.5) holds.

Please note that the reverse statement may **not** be true. That is to say, (2.5) is only the **necessary** condition for an interior optimum. We will discuss it in detail in Section 2.E.

Now, we come back to the condition (2.5). Recall that in Chapter 1, the condition

$$F_1(x^*)/G_1(x^*) = F_2(x^*)/G_2(x^*) \iff MU_1/p_1 = MU_2/p_2.$$

We used λ to denote the marginal utility of income, which equals to $MU_1/p_1 = MU_2/p_2$. As in Chapter 1, in the general case, we also define λ as

$$\begin{aligned} \lambda &= F_1(x^*)/G_1(x^*) = F_2(x^*)/G_2(x^*) \\ \Longrightarrow F_j(x^*) &= \lambda G_j(x^*),\ j = 1, 2. \end{aligned} \tag{2.6}$$

Here, similar to the "marginal utility of income" interpretation of λ, λ in (2.6) corresponds to the change of $F(x^*)$ with respect to a change in c. We will learn this interpretation and its implications in Chapter 4.

Before we continue the discussion of *Lagrange's Method* following Equation (2.6), several digressions will be discussed in Sections 2.C, 2.D and 2.E.

2.C. Constraint Qualification

You may have already noticed that when we rewrite $\mathrm{d}x_1$ and $\mathrm{d}x_2$ in (2.3), we require $G_1(x^0) \neq 0$ and $G_2(x^0) \neq 0$. The question now is "what happens if $G_1(x^0) = 0$ or $G_2(x^0) = 0$?"[2] If, say, $G_1(x^0) = 0$, infinitesimal change of x_1^0 could be made without affecting the constraint.[3] Thus, if $F_1(x^0) \neq 0$, it would be desirable to change x_1^0 in the direction that increases $F(x^0)$.[4] Say, if $F_1(x_0) > 0$, then $F(x^0)$ could increase by raising x_1^0. This process could be applied until either $F_1(x) = 0$, or $G_1(x) \neq 0$. Intuitively, for the consumer choice model we discussed in Chapter 1, $G_1(x^0) = p_1 = 0$ means that good 1 is free. Then, it is desirable to consume the free good as long as consuming the good increases the consumer's utility, or until the point where good 1 is no longer free.

Remark. Note x^0 could be any interior point. In particular, if the point of consideration is the optimum point x^*, then, if $G_1(x^*) = 0$, it must be the case that $F_1(x^*) = 0$.

[2] The case $G_1(x^0) = G_2(x^0) = 0$ will be considered later.
[3] See Equation (2.2).
[4] See Equation (2.1).

A more tricky question is that "what if $G_1(x^0) = G_2(x^0) = 0$?" There would be no problem if $G_1(x^0) = G_2(x^0) = 0$ only means that x_1^0 and x_2^0 are free and should be consumed to the point of satiation. However, this case is tricky since it could be arising from the quirks of algebra or calculus. As a concrete example, let's reconsider the consumer choice model in Chapter 1. That problem has an equivalent formulation as follows:

$$\max_{x_1, x_2} U(x_1, x_2)$$
$$\text{s.t. } (p_1 x_1 + p_2 x_2 - I)^3 = 0.$$

Now,

$$G_1(x) = 3p_1(p_1 x_1 + p_2 x_2 - I)^2 = 0,$$
$$G_2(x) = 3p_2(p_1 x_1 + p_2 x_2 - I)^2 = 0.$$

However, the goods are not free at the margin. The contradiction of $G_1(x) = G_2(x) = 0$ and $p_1, p_2 > 0$ makes our method not working.

To avoid running into such problems, the theory assumes the condition of **Constraint Qualification**. For our particular problem, *Constraint Qualification* requires $G_1(x^*) \neq 0$, or $G_2(x^*) \neq 0$, or both.

Remark. Failure of *Constraint Qualification* is a rare problem in practice. If you run into such a problem, you could rewrite the algebraic form of the constraint, just as in the budget constraint example above.

2.D. The tangency argument

The optimization condition (2.5) could also be recovered using the tangency argument. Recall in our Chapter 1 example, the optimality requires the tangency of the budget line and the indifference curve. In the general case, similar observation is still valid. And we could obtain the optimality condition with the help of the graph, Figure 2.1.

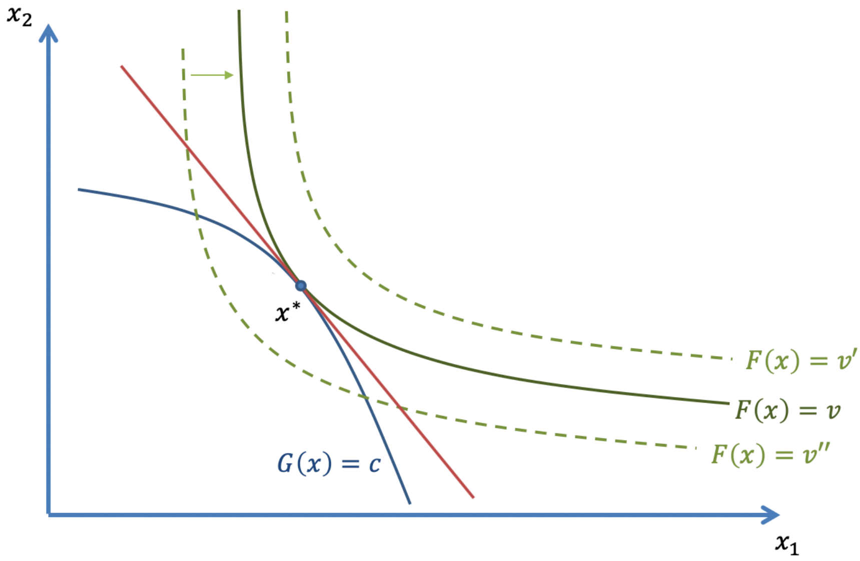

Figure 2.1: Tangency argument

The curve $G(x) = c$ is the constraint. The curves $F(x) = v$, $F(x) = v'$, $F(x) = v''$ are samples of indifference curves. The indifference curves to the right attains higher value for the objective function $F(x)$ compares to those on the left. That is, in the graph, $v' > v' > v''$. It could be seen from the graph, the optimal x^* is attained when the constraint $G(x) = c$ is tangent to an indifference curve $F(x) = v$.

We next look for the tangency condition. For $G(x) = c$, tangency means $\mathrm{d}G(x) = 0$. From (2.2), we have

$$\mathrm{d}x_2/\mathrm{d}x_1 = -G_1(x)/G_2(x). \tag{2.7}$$

Similarly, for the indifference curve $F(x) = v$, tangency means $\mathrm{d}F(x) = 0$. From (2.1), we have

$$\mathrm{d}x_2/\mathrm{d}x_1 = -F_1(x)/F_2(x). \tag{2.8}$$

Since $G(x) = c$ and $F(x) = v$ are mutually tangential at $x = x^*$,

$$F_1(x^*)/F_2(x^*) = G_1(x^*)/G_2(x^*).$$

The above condition is equivalent to (2.5).

Note that if $G_1(x) = G_2(x) = 0$, the slope in (2.7) is not well defined.[5] And we avoid this problem by imposing the *Constraint Qualification* condition as discussed in Section 2.C.

2.E. Necessary vs. Sufficient Conditions

Recall, in Section 2.B, we have established the following result:

> If an interior point x^* maximizes $F(x)$ subject to $G(x) = c$, then (2.5) holds.

In other words, (2.5) is only a necessary condition for optimality. Since the first-order derivatives are involved, it is called the *first-order necessary condition*.

First-order necessary condition helps us narrow down the search for the maximum. However, does not guarantee the maximum. Consider the following graph of unconstrained maximization problem with a single variable.

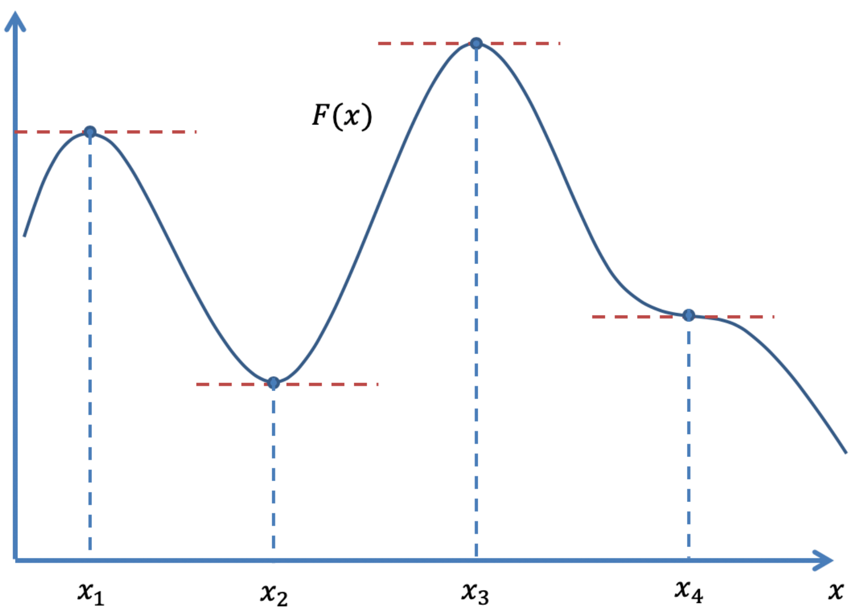

Figure 2.2: Stationary points

[5]Only $G_2(x) = 0$ is not a serious problem. It only means that the slope is vertical.

We want to maximize $F(x)$ in Figure 2.2. The first-order necessary condition for this problem is

$$F'(x) = 0. \tag{2.9}$$

All x_1, x_2, x_3 and x_4 satisfy condition (2.9). However, only x_3 is the global maximum that we are looking for.

Here,

(i) x_1 is a local maximum but not a global one. The problem occurs since when we apply first-order approximation, we only check whether $F(x)$ could be improved by making infinitesimal change in x. Therefore, we obtain a condition for local peaks.

(ii) x_2 is a minimum. This problem occurs since first-order necessary condiition for minimum is the same as that for maximum. More specifically, minimizing $F(x)$ is the same as maximizing $-F(x)$. First-order necessary condition for minimizing $F(x)$ (or maximizing $-F(x)$) and maximizing $F(x)$ are both $F'(x) = 0$.

(iii) x_4 is called a *saddle point*. You could think of $F(x) = x^3$ as a concrete example. We have $F'(0) = 0$, but $x = 0$ is neither a maximum nor a minimum.

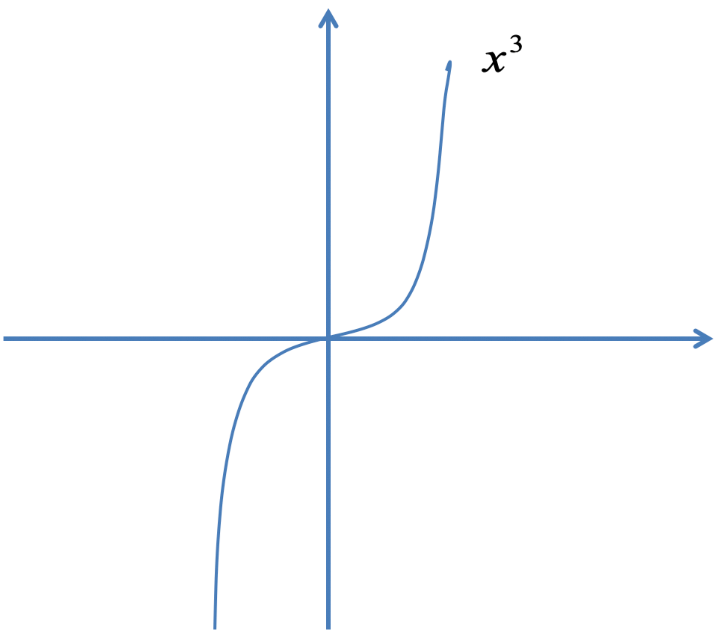

Figure 2.3: $F(x) = x^3$

We used unconstrained maximization problem for easy illustration. The problems remain for constrained maximization problem.

As the caption of Figure 2.2 shows, any point satisfying the *first-order necessary conditions* is called *a stationary point*. The global maximum is one of these points. We will learn how to check whether a point is indeed a maximum in Chapters 6 to Chapter 8.

2.F. Lagrange's Method

In this section, we will explore a general method, called *Lagrange's Method*, to solve the constrained maximization problem restated as follows:

$$\max_x F(x)$$
$$\text{s.t. } G(x) = c.$$

We introduce an unknown variable λ and define a new function, called the Lagrangian:[6]

$$\mathcal{L}(x,\lambda) = F(x) + \lambda\,[c - G(x)] \tag{2.10}$$

Partial derivatives of $\mathcal{L}$ give

$$\mathcal{L}_j(x,\lambda) = \partial\mathcal{L}/\partial x_j = F_j(x) - \lambda G_j(x)$$
$$\mathcal{L}_\lambda(x,\lambda) = \partial\mathcal{L}/\partial\lambda = c - G(x)$$

The first-order necessary condition (2.5) is equivalent to (2.6), and now becomes just

$$\mathcal{L}_j(x,\lambda) = 0.$$

And the constraint is simply

$$\mathcal{L}_\lambda(x,\lambda) = 0.$$

The result is restated formally as the following thoreom:

[6] You would see in a minute that this λ is the same as that in Section 2.B.

Theorem 2.1 (Lagrange's Theorem)**.** *Suppose x is a two-dimensional vector, c is a scalar, and F and G functions taking scalar values. Suppose x^* solves the following maximization problem:*

$$\max_x F(x)$$
$$s.t.\ \ G(x) = c,$$

and the constraint qualification holds, that is, if $G_j(x^*) \neq 0$ for at least one j. Define the function $\mathcal{L}$ as in (2.10):

$$\mathcal{L}(x, \lambda) = F(x) + \lambda \left[c - G(x)\right]. \tag{2.10}$$

Then there is a value of λ such that

$$\mathcal{L}_j(x^*, \lambda) = 0 \text{ for } j = 1, 2 \qquad \mathcal{L}_\lambda(x^*, \lambda) = 0. \tag{2.11}$$

Please always keep in mind that the theorem only provide **necessary conditions** for optimality. Besides, Condition (2.11) do not guarantee existence or uniqueness of the solution. Note that if conditions in (2.11) have no solution, it may be that the maximization problem itself has no solution, or the *Constraint Qualification* may fail so that the first-order conditions are not applicable. If (2.11) have multiple solutions, we need to check the second-order conditions.[7]

In most of our applications, the problems will be well-posed and the first-order necessary condition will lead to a unique solution.

In the next section, we will apply the *Lagrange's Theorem* in examples.

2.G. Examples

Example 2.1: Preferences that Imply Constant Budget Shares. Consider a consumer choosing between two goods x and y, with prices p and q respectively. His income is I,

[7] We will learn Second-Order Conditions in Chapter 8.

so the budget constraint is

$$px + qy = I.$$

Suppose the utility function is

$$U(x, y) = \alpha \ln(x) + \beta \ln(y).$$

What is the consumer's optimal bundle (x, y)?

Solution. First, state the problem:

$$\begin{aligned} &\max_{x,y} U(x, y) \equiv \max_{x,y} \alpha \ln(x) + \beta \ln(y) \\ &\text{s.t. } px + qy = I. \end{aligned}$$

Then, we apply *Lagrange's Method*.

i. Write the Lagrangian:

$$\mathcal{L}(x, y, \lambda) = \alpha \ln(x) + \beta \ln y + \lambda \left[I - px - qy\right].$$

ii. First-order necessary conditions are

$$\partial\mathcal{L}/\partial x = \alpha/x - \lambda p = 0, \tag{2.12}$$

$$\partial\mathcal{L}/\partial y = \beta/y - \lambda q = 0, \tag{2.13}$$

$$\partial\mathcal{L}/\partial \lambda = I - px - py = 0. \tag{2.14}$$

There are various ways to solve the equation system. Here, we introduce one of them. By (2.12) and (2.13):

$$\left.\begin{aligned} \alpha/x - \lambda p = 0 &\implies \alpha/x = \lambda p \\ \beta/y - \lambda q = 0 &\implies \beta/y = \lambda q \end{aligned}\right\} \implies \frac{\alpha}{\beta}\frac{y}{x} = \frac{p}{q} \implies y = \frac{\beta p}{\alpha q} x \tag{2.15}$$

Plugging (2.15) into (2.14):

$$I - px - q\frac{\beta p}{\alpha q} x = 0 \implies x = \frac{\alpha I}{(\alpha + \beta)p}. \tag{2.16}$$

Plugging (2.16) back into (2.12) and (2.15), we could obtain $\lambda = \frac{(\alpha+\beta)}{I}$ and $y = \frac{\beta I}{(\alpha+\beta)q}$.

To conclude,

$$x = \frac{\alpha I}{(\alpha+\beta)p}, \qquad y = \frac{\beta I}{(\alpha+\beta)q}, \qquad \lambda = \frac{(\alpha+\beta)}{I}.$$

We call this demand implying constant budget shares since the share of income spent on the two goods are constant:

$$\frac{px}{I} = \frac{\alpha}{\alpha+\beta}, \qquad \frac{qy}{I} = \frac{\beta}{\alpha+\beta}.$$

Example 2.2: Guns vs. Butter. Consider an economy with 100 units of labor. It can produce guns x or butter y. To produce x guns, it takes x^2 units of labor; likewise y^2 units of labor are needed to produce y butter. Therefore, the economy's resource constraint is $x^2 + y^2 = 100$.

Let a and b be social values attached to guns and butter. And the objective function to be maximized is $F(x, y) = ax + by$.

What is the optimal amount of guns and butter?

Solution. First, state the problem:

$$\max_{x,y} F(x,y) \equiv \max_{x,y} ax + by$$
$$\text{s.t. } x^2 + y^2 = 100.$$

Then, we apply *Lagrange's Method*.

i. Write the Lagrangian:

$$\mathcal{L}(x, y, \lambda) = ax + by + \lambda \left[100 - x^2 - y^2\right].$$

ii. First-order necessary conditions are

$$\partial\mathcal{L}/\partial x = a - 2\lambda x = 0,$$
$$\partial\mathcal{L}/\partial y = b - 2\lambda y = 0,$$
$$\partial\mathcal{L}/\partial \lambda = 100 - x^2 - y^2 = 0.$$

Solving the equation system, we get

$$x = \frac{10a}{\sqrt{a^2+b^2}}, \qquad y = \frac{10b}{\sqrt{a^2+b^2}}, \qquad \lambda = \frac{\sqrt{a^2+b^2}}{20}.$$

Here, the optimal values x and y are called *homogeneous of degree* 0 *with respect to* a *and* b since if we increase a and b in equal proportions, the values of x and y would not change. In other words, x would increase only when a increases relatively more than the increment of b.

Remark. It is always useful to use graphs to help you think. The graphical illustration of the current problem is shown in Figure 2.4 below.

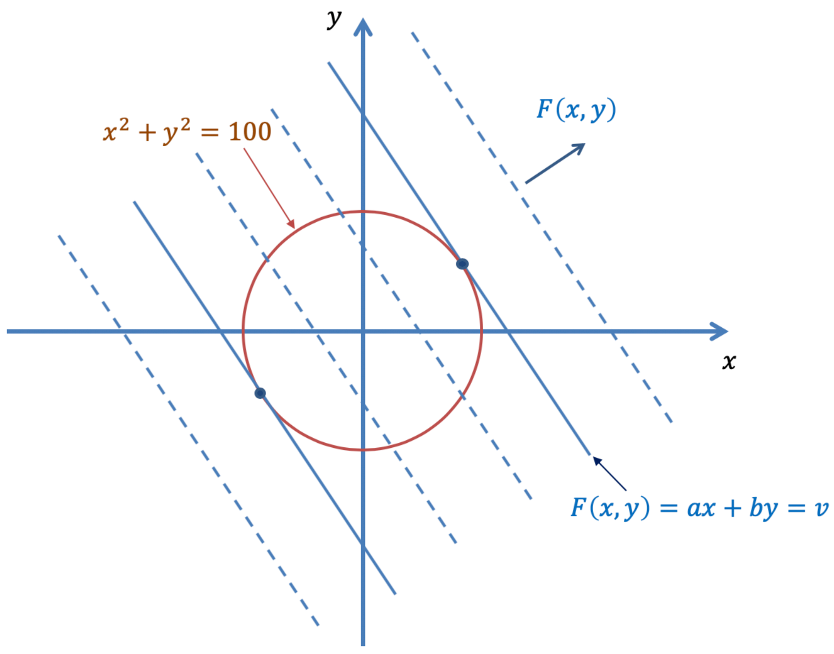

Figure 2.4: The maximization problem

It is not hard to see that the maximum is attained at the intersection to the right.