Chapter 2. Lagrange's Method

Xiaoxiao Hu

February 24, 2020

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:

$$\max_{x_1 \geq 0,\, x_2 \geq 0} U(x_1, x_2)$$

$$\text{s.t. } p_1 x_1 + p_2 x_2 = I.$$

Statement of the problem

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:

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

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

Statement of the problem

- 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.

The arbitrage argument

We reiterate the algorithm of finding the optimal point:

(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.

The arbitrage argument

- Now, we will discuss the arbitrage argument behind the algorithm and derive the "non-arbitrage" condition.
- Consider initial point x^0 and infinitesimal change $\mathrm{d}x$.
- Since the change in x^0 is *infinitesimal*, the changes in values could be approximated by the first-order linear terms in Taylor series.

The arbitrage argument

Using subscripts to denote partial derivatives, we have

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

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

Recall the concrete example in Chapter 1,

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

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

The arbitrage argument

- 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, $\mathrm{d}G(x^0) = 0$.

The arbitrage argument

- $\mathrm{d}G(x^0) = 0$ together with (2.2),

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

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

- Then,

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

The arbitrage argument

From (2.3) and (2.1)

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

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

we get

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

The arbitrage argument

$$\mathrm{d}F(x^0) = \Big[F_1(x^0)/G_1(x^0) - F_2(x^0)/G_2(x^0)\Big]\,\mathrm{d}c. \qquad (2.4)$$

- Recall, $\mathrm{d}c = G_1(x^0)\mathrm{d}x_1 = -G_2(x^0)\mathrm{d}x_2$.
- 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.
- If the expression in the bracket is **positive**, then $F(x^0)$ could increase by choosing $\mathrm{d}c > 0$.
- Similarly, if it is **negative**, then choose $\mathrm{d}c < 0$.

The arbitrage argument

$$\mathrm{d}F(x^0) = \Big[F_1(x^0)/G_1(x^0) - F_2(x^0)/G_2(x^0)\Big]\,\mathrm{d}c. \qquad (2.4)$$

- The same argument holds for all other interior points along the constraint.

- Therefore, for the interior optimum x^*, we must the following "non-arbitrage" condition:

$$F_1(x^*)/G_1(x^*) - F_2(x^*)/G_2(x^*) = 0$$

$$\Longrightarrow F_1(x^*)/G_1(x^*) = F_2(x^*)/G_2(x^*) \qquad (2.5)$$

The arbitrage argument

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

- 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.

The arbitrage argument

Remark. If an interior point x^* is a maximum, then (2.5) $F_1(x^*)/G_1(x^*) = F_2(x^*)/G_2(x^*)$ holds.

- 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 Subsection 2.E.

The arbitrage argument

- Now, we come back to Condition (2.5):

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

- Recall in Chapter 1, Condition (2.5) is equivalent to

$$MU_1/p_1 = MU_2/p_2.$$

- We used λ to denote the **marginal utility of income**, which equals to $MU_1/p_1 = MU_2/p_2$.

The arbitrage argument

- Similarly, in the general case, we also define λ as

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

$$\Longrightarrow F_j(x^*) = \lambda G_j(x^*),\ j = 1, 2. \tag{2.6}$$

- Here, λ 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.

A few Digressions

$$F_j(x^*) = \lambda G_j(x^*),\ j = 1, 2. \tag{2.6}$$

Before we continue the discussion of *Lagrange's Method* following Equation (2.6), several digressions will be discussed in Subsections 2.C Constraint Qualification, 2.D Tangency Argument and 2.E Necessary vs. Sufficient Consitions.

2.C. Constraint Qualification

- You may have already noticed that (2.3)

$$dx_1 = dc/G_1(x^0) \text{ and } dx_2 = -dc/G_2(x^0) \qquad (2.3)$$

requires $G_1(x^0) \neq 0$ 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]

[2] The case $G_1(x^0) = G_2(x^0) = 0$ will be considered later.

Constraint Qualification

- If, say, $G_1(x^0) = 0$, then infinitesimal change of x_1^0 could be made without affecting the constraint.

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

- Thus, if $F_1(x^0) \neq 0$, it would be desirable to change x_1^0 in the direction that increases $F(x^0)$.

$$\mathrm{d}F(x^0) = F_1(x^0)\mathrm{d}x_1 + F_2(x^0)\mathrm{d}x_2. \qquad (2.1)$$

- This process could be applied until either $F_1(x) = 0$, or $G_1(x) \neq 0$.

Constraint Qualification

- 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.

Constraint Qualification

- 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 $F_1(x^*) = 0$.

Constraint Qualification

- A more tricky question is

  "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.

Constraint Qualification

- As a concrete example, let's reconsider the consumer choice model in Chapter 1:

$$\max_{x_1, x_2} U(x_1, x_2)$$

$$\text{s.t. } p_1 x_1 + p_2 x_2 - I = 0.$$

- That problem has an equivalent formulation as follows:

$$\max_{x_1, x_2} U(x_1, x_2)$$

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

Constraint Qualification

- Under the new formulation:

$$G_1(x) = 3p_1(p_1x_1 + p_2x_2 - I)^2 = 0,$$

$$G_2(x) = 3p_2(p_1x_1 + p_2x_2 - I)^2 = 0.$$

- 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.

Constraint Qualification

- 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.

Constraint Qualification

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

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

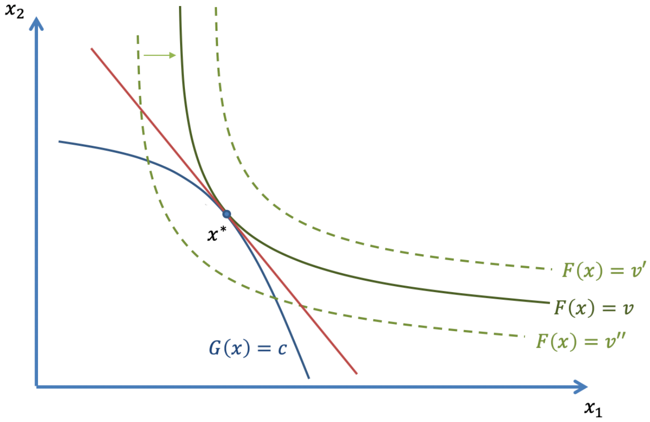
could also be recovered using the tangency argument.

- 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.

The tangency argument

We could obtain the optimality condition with the help of the graph:



The 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 compares to those on the left.
- The optimal x^* is attained when the constraint $G(x) = c$ is tangent to an indifference curve $F(x) = v$.

The tangency argument

- We next look for the tangency condition.

- For $G(x) = c$, tangency means $\mathrm{d}G(x) = 0$. From (2.2), we have

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

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

The tangency argument

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

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

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

The tangency argument

Recall,

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

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

- Since $G(x) = c$ and $F(x) = v$ are mutually tangential at $x = x^*$, we get $\boldsymbol{F_1(x^*)/F_2(x^*) = G_1(x^*)/G_2(x^*)}$.

- The above condition is equivalent to (2.5):

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

The tangency argument

- Note that if $G_1(x) = G_2(x) = 0$, the slope in (2.7) is not well defined.[3]

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

- We avoid this problem by imposing the *Constraint Qualification* condition as discussed in Subsection 2.C.

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

2.E. Necessary vs. Sufficient Conditions

- Recall, in Subsection 2.B, we established the result:

> *Remark.* If an interior point x^* is a maximum, then (2.5) $F_1(x^*)/G_1(x^*) = F_2(x^*)/G_2(x^*)$ holds.

- In other words, (2.5) is only a necessary condition for optimality.

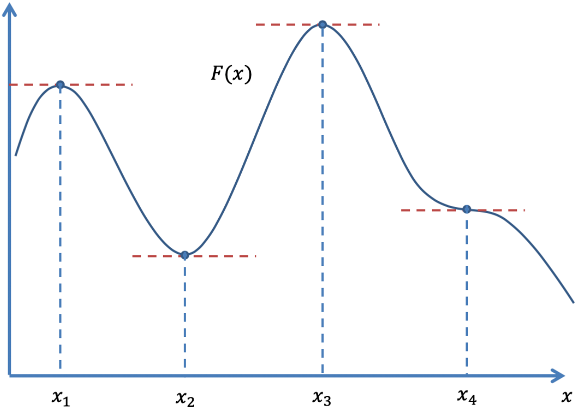
- Since the first-order derivatives are involved, it is called the *first-order necessary condition.*

First-order necessary condition

- *First-order necessary condition* helps us narrow down the search for the maximum.
- However, it does not guarantee the maximum.

First-order necessary condition

Consider the following unconstrained maximization problem:



First-order necessary condition

- We want to maximize $F(x)$.

- 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.

First-order necessary condition: local maximum

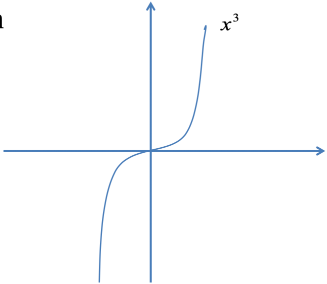
- x_1 is a local maximum but not a global one.
- The problem occurs since when we apply first-order approximation, we only check whether $F(x)$ could be improved by making infinitesimal change in x.
- Therefore, we obtain a condition for local peaks.

First-order necessary condition: minimum

- x_2 is a minimum.
- This problem occurs since first-order necessary condiition for minimum is the same as that for maximum.
- More specifically, this is because minimizing $F(x)$ is the same as maximizing $-F(x)$.
- First-order necessary condiition: $F'(x) = 0$

First-order necessary condition: saddle point

- 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 minim



First-order necessary condition

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

Stationary point

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

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

Lagrange's Method

- We introduce an unknown variable λ[4] and define a new function, called the Lagrangian:

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

- Partial derivatives of $\mathcal{L}$ give

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

$$\mathcal{L}_\lambda(x, \lambda) = \partial \mathcal{L} / \partial \lambda = c - G(x) \tag{$\mathcal{L}_\lambda$}$$

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

Lagrange's Method

- Restate ($\mathcal{L}_j$)

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

- Recall first-order necessary condition (2.5)

$$F_1(x^*)/G_1(x^*) = F_2(x^*)/G_2(x^*) = \lambda \qquad (2.5)$$

- First-order necessary condition is just

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

Lagrange's Method

- Restate ($\mathcal{L}_\lambda$)

$$\mathcal{L}_\lambda(x, \lambda) = \partial \mathcal{L} / \partial \lambda = c - G(x) \qquad (\mathcal{L}_\lambda)$$

- Recall constraint: $G(x) = c$.

- The constraint is simply

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

Lagrange's Method

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.

Theorem 2.1 (continued).

Define function $\mathcal{L}$ as in (2.10):

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

Then there is a value of λ such that

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

- 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.

Lagrange's Method

- 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.[5]

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

Lagrange's Method

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

2.G. Examples

In this subsection, we will apply the *Lagrange's Theorem* in examples.

Example 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, 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)?

Example 1: Solution.

First, state the problem:

$$\max_{x,\,y} U(x,y) \equiv \max_{x,\,y} \alpha \ln(x) + \beta \ln(y)$$

$$\text{s.t. } px + qy = I.$$

Then, we apply *Lagrange's Method.*

i. Write the Lagrangian:

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

Example 1: Solution (continued)

ii. First-order necessary conditions are

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

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

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

Solving the equation system, we get

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

Example 1: Solution (continued)

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

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: 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

$$x^2 + y^2 = 100.$$

Example 2: Guns vs. Butter.

- 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?

Example 2: 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:

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

Example 2: Solution (continued)

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

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

Example 2: Solution (continued)

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

- Here, the optimal values x and y are called *homogeneous of degree* 0 *with respect to* a *and* b.
  - 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.

Example 2: Solution (continued)

Remark. It is always useful to use graphs to help you think.