Ch02. Constrained Optimization

Ping Yu

Faculty of Business and Economics
The University of Hong Kong

Overview of This Chapter

- We will study the first order **necessary** conditions for an optimization problem with equality and/or inequality constraints.
- The former is often called the Lagrange problem and the latter is called the Kuhn-Tucker problem (or nonlinear programming).
- We will not discuss the unconstrained optimization problem separately but treat it as a special case of the constrained problem because the unconstrained problem is rare in economics.

Maximum/Minimum and Maximizer/Minimizer

- A function $f : X \to \mathbb{R}$ has a global maximizer at x^* if $f(x^*) \geq f(x)$ for all $x \in X$ and $x \neq x^*$. Similarly, the function has a global minimizer at x^* if $f(x^*) \leq f(x)$ for all $x \in X$ and $x \neq x^*$.
- If the domain X is a metric space, usually a subset of $\mathbb{R}^n$, then f is said to have a local maximizer at the point x^* if there exists $r > 0$ such that $f(x^*) \geq f(x)$ for all $x \in B_r(x^*) \cap X \setminus \{x^*\}$, where $B_r(x^*)$ is an open ball with center x^* and radius r. Similarly, the function has a local minimizer at x^* if $f(x^*) \leq f(x)$ for all $x \in B_r(x^*) \cap X \setminus \{x^*\}$.
- In both the global and local cases, the value of the function at a maximizer is called the maximum of the function and the value of the function at a minimizer is called the minimum of the function.
  - The maxima and minima (the respective plurals of maximum and minimum) are called optima (the plural of optimum), and the maximizer and minimizer are called the optimizer. The optimizer and optimum without any qualifier means the global ones. [Figure here]
  - A global optimizer is always a local optimizer but the converse is not correct.
- In both the global and local cases, the concept of a strict optimum and a strict optimizer can be defined by replacing weak inequalities by strict inequalities.

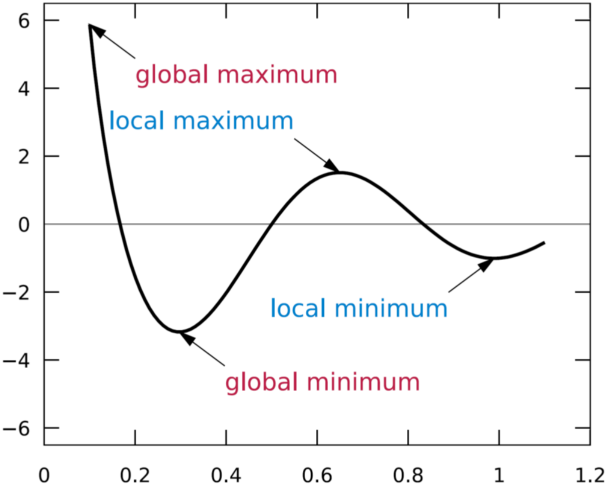


Figure: Local and Global Maxima and Minima for $\cos(3\pi x)/x$, $0.1 \le x \le 1.1$

Notations

- The problem of maximization is usually stated as

$$\max_x f(x)$$
$$\text{s.t. } x \in X,$$

where "s.t." is a short for "subject to",[1] and X is called the constraint set or feasible set.

- The maximizer is denoted as

$$\arg\max \{f(x) | x \in X\} \text{ or } \arg\max_{x \in X} f(x),$$

where "arg" is a short for "arguments".

- The difference between the Lagrange problem and Kuhn-Tucker problem lies in the definition of X.

[1] "s.t." is also a short for "such that" in some books.

Equality-Constrained Optimization

Consumer's Problem

- In microeconomics, a consumer faces the problem of maximizing her utility subject to the income constraint:

$$\max_{x_1,x_2} u(x_1,x_2)$$
$$\text{s.t. } p_1x_1+p_2x_2-y=0$$

- If the indifference curves (i.e., the sets of points (x_1,x_2) for which $u(x_1,x_2)$ is a constant) are convex to the origin, and the indifference curves are nice and smooth, then the point (x_1^*,x_2^*) that solves the maximization problem is the point at which the indifference curve is tangent to the budget line as given in the following figure.

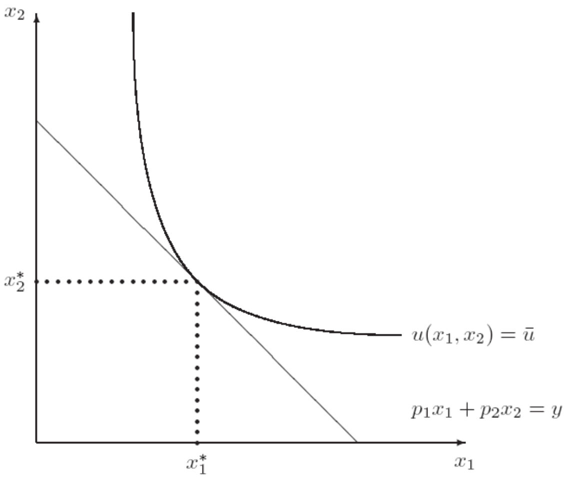


Figure: Utility Maximization Problem in Consumer Theory

Economic Condition for Maximization

- At the point (x_1^*, x_2^*) it must be true that the marginal utility with respect to good 1 divided by the price of good 1 must equal the marginal utility with respect to good 2 divided by the price of good 2.
- For if this were not true then the consumer could, by decreasing the consumption of the good for which this ratio was lower and increasing the consumption of the other good, increase her utility.
- Thus we have

$$\frac{\frac{\partial u}{\partial x_1}(x_1^*, x_2^*)}{p_1} = \frac{\frac{\partial u}{\partial x_2}(x_1^*, x_2^*)}{p_2},$$

or

$$\frac{p_1}{p_2} = \frac{\frac{\partial u}{\partial x_1}(x_1^*, x_2^*)}{\frac{\partial u}{\partial x_2}(x_1^*, x_2^*)}.$$

- What does this mean in the figure? See below.

Mathematical Arguments

- Let x_2^u be the function that defines the indifference curve through the point (x_1^*, x_2^*), i.e.,

$$u(x_1, x_2^u(x_1)) \equiv \bar{u} \equiv u(x_1^*, x_2^*).$$

- Now, totally differentiating this identity gives

$$\frac{\partial u}{\partial x_1}(x_1, x_2^u(x_1)) + \frac{\partial u}{\partial x_2}(x_1, x_2^u(x_1)) \frac{dx_2^u}{dx_1}(x_1) = 0.$$

  That is,

$$\frac{dx_2^u}{dx_1}(x_1) = -\frac{\frac{\partial u}{\partial x_1}(x_1, x_2^u(x_1))}{\frac{\partial u}{\partial x_2}(x_1, x_2^u(x_1))}.$$

- Given that $x_2^u(x_1^*) = x_2^*$, the slope of the indifference curve at the point (x_1^*, x_2^*)

$$\frac{dx_2^u}{dx_1}(x_1^*) = -\frac{\frac{\partial u}{\partial x_1}(x_1^*, x_2^*)}{\frac{\partial u}{\partial x_2}(x_1^*, x_2^*)}.$$

- Also, the slope of the budget line is $-\frac{p_1}{p_2}$. Combining these two results again gives the result in the last slide.

Necessary Conditions for Maximization

- Two functions and two unknowns:

$$\begin{aligned}\frac{\frac{\partial u}{\partial x_1}(x_1^*,x_2^*)}{p_1} &= \frac{\frac{\partial u}{\partial x_2}(x_1^*,x_2^*)}{p_2}, \\ p_1x_1^*+p_2x_2^* &= y.\end{aligned}$$

- We can solve out (x_1^*,x_2^*) if we know $u(\cdot,\cdot)$, p_1, p_2 and y.
- Reformulation to get general conditions: denote the common value of the ratios in the first condition by λ,

$$\frac{\frac{\partial u}{\partial x_1}(x_1^*,x_2^*)}{p_1} = \lambda = \frac{\frac{\partial u}{\partial x_2}(x_1^*,x_2^*)}{p_2},$$

and we can rewrite the two necessary conditions as

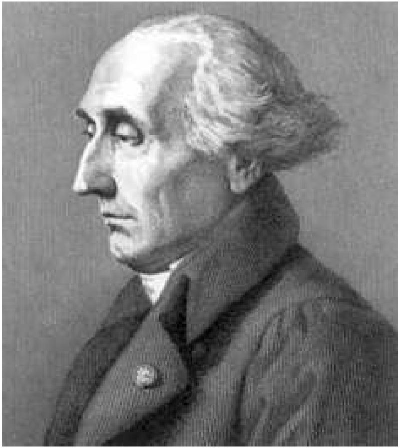
$$\begin{aligned}\frac{\partial u}{\partial x_1}(x_1^*,x_2^*)-\lambda p_1 &= 0, \\ \frac{\partial u}{\partial x_2}(x_1^*,x_2^*)-\lambda p_2 &= 0, \\ y-p_1x_1^*-p_2x_2^* &= 0.\end{aligned}$$

Lagrangian

- Define the Lagrangian as

$$\mathscr{L}(x_1, x_2, \lambda) = u(x_1, x_2) + \lambda(y - p_1 x_1 - p_2 x_2).$$

- Calculate $\frac{\partial \mathscr{L}}{\partial x_1}$, $\frac{\partial \mathscr{L}}{\partial x_2}$, and $\frac{\partial \mathscr{L}}{\partial \lambda}$, and set the results equal to zero we obtain exactly the three equations in the last slide.
- Three equations and three unknowns, so we can solve out $(x_1^*, x_2^*, \lambda^*)$ in principle.
- λ is the new artificial or auxiliary variable, and is commonly called Lagrange multiplier.



Joseph-Louis Lagrange (1736-1813), Italian[2]

[2] but worked at Berlin and Paris during most of his life.

General Necessary Conditions for Maximization

- Suppose that we have the following maximization problem

$$\max_{x_1,\cdots,x_n} f(x_1,\cdots,x_n)$$
$$\text{s.t. } g(x_1,\cdots,x_n) = c$$

- Let

$$\mathscr{L}(x_1,\ldots,x_n,\lambda) = f(x_1,\ldots,x_n) + \lambda(c - g(x_1,\ldots,x_n)).$$

- If $(x_1^*,\ldots,x_n^*)$ solves this maximization problem, there is a value of λ, say λ^* such that

$$\frac{\partial \mathscr{L}}{\partial x_i}(x_1^*,\ldots,x_n^*,\lambda^*) = 0,\ i = 1,\ldots,n, \tag{1}$$
$$\frac{\partial \mathscr{L}}{\partial \lambda}(x_1^*,\ldots,x_n^*,\lambda^*) = 0. \tag{2}$$

More Explanations on the Necessary Conditions

- The conditions (1) are precisely the first order conditions for choosing $x_1, \ldots, x_n$ to maximize $\mathscr{L}$, once λ^* has been chosen.
- From conditions (1), there are two equivalent ways to interpret the constrained maximization problem.
  - The decision maker must satisfy $g(x_1, \ldots, x_n) = c$ and that she should choose among all points that satisfy this constraint the point at which $f(x_1, \ldots, x_n)$ is greatest.
  - The decision maker chooses any point she wishes but that for each unit by which she violates the constraint $g(x_1, \ldots, x_n) = c$ we shall take away λ units from her payoff.
- We must be careful to choose λ to be the correct value.
- If we choose λ too small, the decision maker may choose to violate her constraint. E.g., if we made the penalty for spending more than the consumer's income very small the consumer would choose to consume more goods than she could afford and to pay the penalty in utility terms.
- On the other hand, if we choose λ too large the decision maker may violate her constraint in the other direction. E.g., the consumer would choose not to spend any of her income and just receive λ units of utility for each unit of her income.

Multiple Constraints

- The technique above can be extended to multiple constraints case:

$$\max_{x_1,\cdots,x_n} f(x_1,\cdots,x_n)$$
$$\text{s.t. } g_1(x_1,\cdots,x_n) = c_1,$$
$$\vdots$$
$$g_m(x_1,\cdots,x_n) = c_m,$$

  where $m \leq n$ or $m < n$ (why?).
- The Lagrangian

$$\mathscr{L}(\mathbf{x},\lambda) = f(\mathbf{x}) + \lambda\cdot(\mathbf{c}-\mathbf{g}(\mathbf{x})), \tag{3}$$

  where $\mathbf{x} = (x_1,\ldots,x_n)'$, $\lambda = (\lambda_1,\cdots,\lambda_m)'$, and $\mathbf{c}$ and $\mathbf{g}$ are similarly defined.
- If $\mathbf{x}^* = (x_1^*,\ldots,x_n^*)'$ solves (3), there are values of λ, say $\lambda^* = (\lambda_1^*,\ldots,\lambda_m^*)'$ such that

$$\frac{\partial\mathscr{L}}{\partial x_i}(\mathbf{x}^*,\lambda^*) = 0,\ i = 1,\ldots,n$$
$$\frac{\partial\mathscr{L}}{\partial \lambda_j}(\mathbf{x}^*,\lambda^*) = 0,\ j = 1,\cdots,m,$$

  which are labeled as "first order conditions" or "FOCs" for the corresponding maximization problem.

Existence of Maximizer

- We have not even claimed that there necessarily is a solution to the maximization problem.
- One example of an unconstrained problem with no solution is

$$\max_x 2x,$$

maximizing over the choice of x the function $2x$. Clearly the greater we make x the greater is $2x$, and so, since there is no upper bound on x there is no maximum.

- Thus we might want to restrict maximization problems to those in which we choose x from some bounded set. Again, this is not enough.
- Consider the problem

$$\max_{0 \le x \le 1} 1/x .$$

The smaller we make x the greater is $1/x$ and yet at zero $1/x$ is not even defined.

- We could define the function to take on some value at zero, say 7. But then the function would not be continuous. Or we could leave zero out of the feasible set for x, say $0 < x \le 1$. Then the set of feasible x is not closed.

Weierstrass Theorem

- We shall restrict maximization problems to those in which we choose x to maximize some continuous function from some closed and bounded set (which is compact from the Heine-Borel Theorem).
- Is there anything else that could go wrong? No. The following result says that if the function to be maximized is continuous and the set over which we are choosing is both closed and bounded, i.e., is compact, then there is a solution to the maximization problem.

Theorem (The Weierstrass Theorem)

Let S be a compact set and $f : S \to \mathbb{R}$ be continuous. Then there is some x^* in S at which the function is maximized. More precisely, there is some x^* in S such that $f(x^*) \geq f(x)$ for any x in S.

- We will give an example later.

Karl T.W. Weierstrass (1815-1897), German[3]

[3] cited as the "father of modern analysis", leaving university without a degree.

Extension to Nonequality Constraints

- In defining the compact sets in the Weierstrass theorem, we typically use inequalities, such as $x \geq 0$.
- However, we did not consider such constraints in the above discussion, but rather considered only equality constraints.
- However, even in the example of utility maximization at the beginning of this section, there were implicitly constraints on x_1 and x_2 of the form

$$x_1 \geq 0, \qquad x_2 \geq 0.$$

- We shall return to this question in the next section.

Extension to Minimization Problem and Unconstrained Problem

- We could have transformed the minimization problem into a maximization problem by simply multiplying the objective function by -1.
- That is, if we wish to minimize $f(x)$ we could do so by maximizing $-f(x)$.
- As an exercise write out the necessary conditions for the case that we wanted to minimize $u(x)$ in the consumer's problem.
- Notice that if x_1^*, x_2^*, and λ satisfy the original equations then x_1^*, x_2^*, and $-\lambda$ satisfy the new equations. Thus we cannot tell whether there is a maximum at (x_1^*, x_2^*) or a minimum.
- This corresponds to the fact that in the case of a function of a single variable over an unconstrained domain at a maximum we require the first derivative to be zero, but that to know for sure that we have a maximum we must look at the **second** derivative which will be discussed in the next chapter.

- For the unconstrained problem, set $\lambda^* = 0$, i.e., since no constraints exist, no penalty is imposed on constraints.

Inequality-Constrained Optimization (Nonlinear Programming)

Introduction to Nonlinear Programming

- Formulation of a simple nonlinear programming problem:

$$\max_x f(x)$$
$$\text{s.t. } x \geq 0,$$

where $\dim(x) = 1$ for simplicity.

- Without the constraint $x \geq 0$, the FOC for the maximization problem is $\frac{df}{dx}(x^*) = 0$.
- When the inequality constraint is added in, either the solution could occur when $x^* > 0$ or it could occur when $x^* = 0$.
- When $x^* > 0$, the FOC should still be $\frac{df}{dx}(x^*) = 0$. When $x^* = 0$, the necessary condition should be

$$\frac{df}{dx}(x^*) \leq 0.\text{(why?)}$$

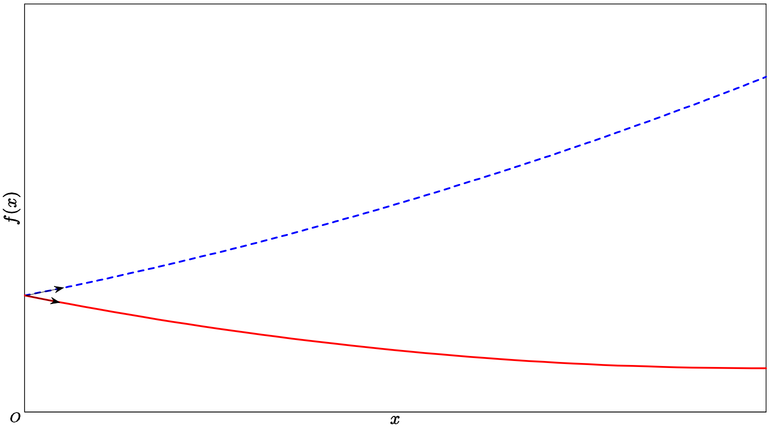


Figure: Illustration of Why $\frac{df}{dx}(x^*) \leq 0$ When $x^* = 0$

Nonlinear Programming with Possible Corner Solution

- So the FOC is

$$\frac{df}{dx}(x^*)\begin{cases} \leq 0, & \text{if } x^* = 0, \\ = 0, & \text{if } x^* > 0, \end{cases}$$

which can be expressed in a compact form as in the following theorem.

Theorem

Suppose that $f : \mathbb{R} \to \mathbb{R}$ is a C^1 function. Then, if x^* maximizes $f(x)$ over all $x \geq 0$, x^* satisfies

$$\begin{aligned} \frac{df}{dx}(x^*) &\leq 0 \\ x^*\frac{df}{dx}(x^*) &= 0 \\ x^* &\geq 0 \end{aligned}$$

- A pair of inequalities, not both of which can be strict (or **slack**) (i.e., at least one of them is effective), is said to show complementary slackness.

Reformulation of the FOCs

- As in the equality-constrained problem, we introduce a Lagrange multiplier.
- If we form the Lagrangian

$$\mathscr{L}(x,\lambda) = f(x) + \lambda x,$$

then we can express these FOCs as (check!)

$$\begin{aligned} \frac{\partial \mathscr{L}}{\partial x}(x^*,\lambda^*) &= \frac{df}{dx}(x^*) + \lambda^* = 0, \\ \frac{\partial \mathscr{L}}{\partial \lambda}(x^*,\lambda^*) &= x^* \geq 0, \\ \lambda^* &\geq 0, \lambda^* \frac{\partial \mathscr{L}}{\partial \lambda}(x^*,\lambda^*) = \lambda^* x^* = 0. \end{aligned}$$

General Inequality-Constrained Problem

- Suppose we want to

$$\max_{x_1,\cdots,x_n} f(x_1,\cdots,x_n)$$
$$\text{s.t.} \quad g_j(x_1,\cdots,x_n) \geq 0,\ j = 1,\cdots,J,$$

  or more compactly,

$$\max_{\mathbf{x}} f(\mathbf{x})$$
$$\text{s.t. } \mathbf{g}(\mathbf{x}) \geq \mathbf{0}.$$

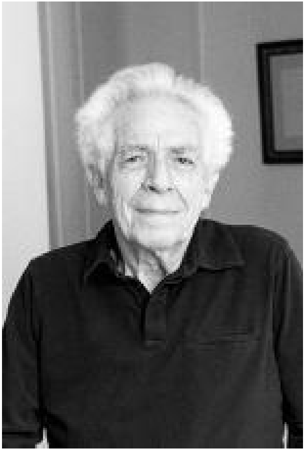
- Form the Lagrangian

$$\mathscr{L}(\mathbf{x},\lambda) = f(\mathbf{x}) + \lambda \cdot \mathbf{g}(\mathbf{x}).$$

  and express the FOCs as

$$\begin{aligned}
\frac{\partial \mathscr{L}}{\partial \mathbf{x}}(\mathbf{x}^*,\lambda^*) &= \frac{\partial f}{\partial \mathbf{x}}(\mathbf{x}^*) + \frac{\partial \mathbf{g}(\mathbf{x}^*)'}{\partial \mathbf{x}}\lambda^* = \mathbf{0}, \\
\frac{\partial \mathscr{L}}{\partial \lambda}(\mathbf{x}^*,\lambda^*) &= \mathbf{g}(\mathbf{x}^*) \geq \mathbf{0}, \\
\lambda^* &\geq \mathbf{0}, \lambda^* \odot \frac{\partial \mathscr{L}}{\partial \lambda}(\mathbf{x}^*,\lambda^*) = \lambda^* \odot \mathbf{g}(\mathbf{x}^*) = \mathbf{0},
\end{aligned}$$

  where $\odot$ is the element-by-element product.
- These FOCs are called the Kuhn-Tucker conditions due to Kuhn and Tucker (1951).



H.W.Kuhn (1925-2014, Princeton)

A.W. Tucker (1905-1995, Princeton)[4]

[4] Albert W. Tucker is the supervisor of John Nash, the Nobel Prize winner in Economics in 1994, and Lloyd Shapley, the Nobel Prize winner in Economics in 2012.

Mixed Constrained Problem

- Combine the equality-constrained and inequality-constrained problem to define the mixed constrained problem:

$$\begin{aligned} &\max_{x_1,\cdots,x_n} f(x_1,\cdots,x_n) \\ \text{s.t. } & g_j(x_1,\cdots,x_n) \geq 0,\ j=1,\cdots,J, \\ & h_k(x_1,\cdots,x_n) = 0,\ k=1,\cdots,K, \end{aligned}$$

  or more compactly,

$$\begin{aligned} &\max_{\mathbf{x}} f(\mathbf{x}) \\ \text{s.t. } & \mathbf{g}(\mathbf{x}) \geq \mathbf{0}, \\ & \mathbf{h}(\mathbf{x}) = \mathbf{0}, \end{aligned}$$

  where $K \leq n$.
- The term "mixed constrained problem" is only for convenience because any equality constraint can be transformed to two inequality constraints, e.g., $h_k(\mathbf{x}) = 0$ is equivalent to $h_k(\mathbf{x}) \geq 0$ and $h_k(\mathbf{x}) \leq 0$.
- Form the Lagrangian

$$\mathscr{L}(\mathbf{x}, \lambda, \mu) = f(\mathbf{x}) + \lambda \cdot \mathbf{g}(\mathbf{x}) + \mu \cdot \mathbf{h}(\mathbf{x}).$$

Theorem of Kuhn-Tucker

Theorem (Theorem of Kuhn-Tucker)

Suppose that $f : \mathbb{R}^n \to \mathbb{R}$, $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^J$ *and* $\mathbf{h} : \mathbb{R}^n \to \mathbb{R}^K$ *are* C^1 *functions. Then, if* $\mathbf{x}^*$ *maximizes* $f(\mathbf{x})$ *over all* $\mathbf{x}$ *satisfying the constraints* $\mathbf{g}(\mathbf{x}) \geq \mathbf{0}$ *and* $\mathbf{h}(\mathbf{x}) = \mathbf{0}$, *and if* $\mathbf{x}^*$ *satisfies the nondegenerate constraint qualification (NDCQ) as will be discussed below, then there exists a vector* (λ^*, μ^*) *such that* $(\mathbf{x}^*, \lambda^*, \mu^*)$ *satisfies the Kuhn-Tucker conditions given as follows:*

$$
\begin{aligned}
&\frac{\partial \mathscr{L}}{\partial \mathbf{x}}(\mathbf{x}^*, \lambda^*, \mu^*) = \mathbf{0}, \ \frac{\partial \mathscr{L}}{\partial \mu}(\mathbf{x}^*, \lambda^*, \mu^*) = \mathbf{h}(\mathbf{x}^*) = \mathbf{0}, \\
&\frac{\partial \mathscr{L}}{\partial \lambda}(\mathbf{x}^*, \lambda^*, \mu^*) = \mathbf{g}(\mathbf{x}^*) \geq \mathbf{0}, \ \lambda^* \geq \mathbf{0}, \\
&\lambda^* \odot \frac{\partial \mathscr{L}}{\partial \lambda}(\mathbf{x}^*, \lambda^*, \mu^*) = \lambda^* \odot \mathbf{g}(\mathbf{x}^*) = \mathbf{0}.
\end{aligned}
$$

- The $\mathbf{x}^*$'s that satisfy the Kuhn-Tucker conditions are called the critical points of $\mathscr{L}$.
- Usually, critical points mean the points that satisfy the FOCs; the Kuhn-Tucker conditions are a special group of FOCs.
- Parallel to Lagrange multipliers in the Lagrange problem, (λ^*, μ^*) are called Kuhn-Tucker multipliers.
- Finally, note that the Kuhn-Tucker conditions are necessary conditions for "local" optima, and of course are also necessary conditions for global optima.

Nondegenerate Constraint Qualification

- A constraint $g_j(\mathbf{x}) \geq 0$ is binding at $\mathbf{x}^*$ if $g_j(\mathbf{x}^*) = 0$.
- Suppose the first J_0 inequality constraints are binding at $\mathbf{x}^*$; then the NDCQ states that the rank at $\mathbf{x}^*$ of the Jacobian matrix of the equality constraints and the binding inequality constraints

$$\mathbf{J} \equiv \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(\mathbf{x}^*) & \cdots & \frac{\partial g_1}{\partial x_n}(\mathbf{x}^*) \\ \vdots & \ddots & \vdots \\ \frac{\partial g_{J_0}}{\partial x_1}(\mathbf{x}^*) & \cdots & \frac{\partial g_{J_0}}{\partial x_n}(\mathbf{x}^*) \\ \frac{\partial h_1}{\partial x_1}(\mathbf{x}^*) & \cdots & \frac{\partial h_1}{\partial x_n}(\mathbf{x}^*) \\ \vdots & \ddots & \vdots \\ \frac{\partial h_K}{\partial x_1}(\mathbf{x}^*) & \cdots & \frac{\partial h_K}{\partial x_n}(\mathbf{x}^*) \end{pmatrix}$$

  is $J_0 + K$ - as large as it can be.
- When for some $\mathbf{x}$'s the NDCQ does not hold, compare the values of $f(\cdot)$ at critical points and also these $\mathbf{x}$'s to determine the ultimate maximizer.

Failure of the Constraint Qualification

Example

Suppose we want to maximize $f(x_1,x_2)=x_1$ s.t. $g(x_1,x_2)=x_1^3+x_2^2\leq 0$. From the following figure, the constraint set is a cusp and it is easy to see that $(x_1^*,x_2^*)=(0,0)$. However, at (x_1^*,x_2^*), there is no λ^* satisfying the Kuhn-Tucker conditions. To see why, set the Lagrangian

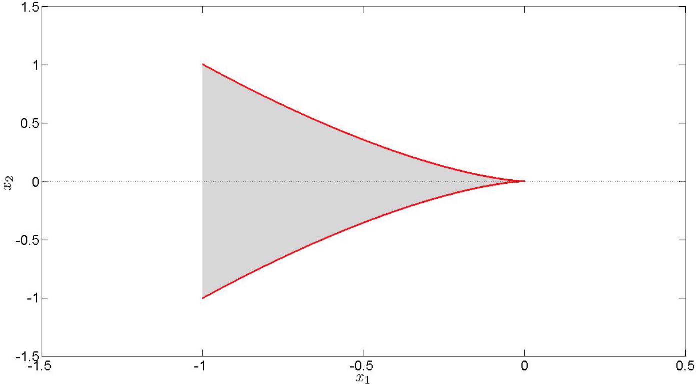
$$\mathscr{L}(\mathbf{x},\lambda)=x_1-\lambda\left(x_1^3+x_2^2\right),$$

and then the Kuhn-Tucker conditions are

$$\begin{aligned} 1-3\lambda x_1^2 &= 0, 2\lambda x_2=0, \\ x_1^3+x_2^2 &\leq 0, \lambda\geq 0, \lambda\left(x_1^3+x_2^2\right)=0. \end{aligned}$$

It is not hard to see that there is no λ^* satisfying these conditions when $(x_1^*,x_2^*)=(0,0)$. What can we learn from this example? Note that $g(x_1,x_2)$ is binding at $(0,0)$, while $(0,0)$ is the critical point of $g(x_1,x_2)$ (i.e., $\frac{\partial g_1}{\partial x_1}(0,0)=\frac{\partial g_1}{\partial x_2}(0,0)=0$), so the constraint qualification fails. If we compare $f(\cdot)$ at the critical values of $\mathscr{L}$ (which is empty) and $(0,0)$, we indeed get the correct maximizer $(0,0)$. $\square$

- The LN provides some intuition on why the NDCQ is required; essentially, the NDCQ guarantees that local to $\mathbf{x}^*$, the binding constraints and their first order approximations are equivalent.



The Constraint Set $\left\{(x_1,x_2)|x_1^3+x_2^2\leq 0\right\}$

An Illustrating Example of Finding the Maximizer

Example

max $x_1^2 + (x_2 - 5)^2$ s.t. $x_1 \geq 0$, $x_2 \geq 0$, and $2x_1 + x_2 \leq 4$.

Solution

First, since the objective function is continuous and the constraint set is compact (why?), by the Weierstrass theorem, the maximizer exists. We then check the NDCQ. $g_1(x) = x_1$, $g_2(x) = x_2$ *and* $g_3(x) = 4 - 2x_1 - x_2$, *so the Jacobian of the constraint functions is*

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -2 & -1 \end{pmatrix},$$

whose any one or two rows are linearly independent. Since at most two of the three constraints can be binding at any one time, the NDCQ holds at any solution candidate.
The Lagrangian is

$$\mathscr{L}(x, \lambda, \mu) = x_1^2 + (x_2 - 5)^2 + \lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 (4 - 2x_1 - x_2),$$

and the Kuhn-Tucker conditions are

$$\begin{aligned} 2x_1 + \lambda_1 - 2\lambda_3 &= 0,\ 2(x_2 - 5) + \lambda_2 - \lambda_3 = 0, \\ x_1 &\geq 0,\ x_2 \geq 0,\ 4 - 2x_1 - x_2 \geq 0, \lambda_1 \geq 0, \lambda_2 \geq 0, \lambda_3 \geq 0, \\ \lambda_1 x_1 &= 0,\ \lambda_2 x_2 = 0,\ \lambda_3 (4 - 2x_1 - x_2) = 0. \end{aligned}$$

Solution (continue)

Totally eight possibilities depending whether $\lambda_j = 0$ *or not,* $j = 1, 2, 3$.

(i) $\lambda_1 > 0, \lambda_2 > 0$ *and* $\lambda_3 > 0 \Longrightarrow x_1 = 0, x_2 = 0$, *and* $2x_1 + x_2 = 4$. *Impossible.*

(ii) $\lambda_1 = 0, \lambda_2 > 0$ *and* $\lambda_3 > 0 \Longrightarrow x_1 \geq 0, x_2 = 0$, *and* $2x_1 + x_2 = 4$. *So* $(x_1, x_2) = (2, 0)$. *From* $4 - 2\lambda_3 = 0$ *and* $-10 + \lambda_2 - \lambda_3 = 0$, *we have* $(\lambda_1, \lambda_2, \lambda_3) = (0, 12, 2)$.

(iii) $\lambda_1 > 0, \lambda_2 = 0$ *and* $\lambda_3 > 0 \Longrightarrow x_1 = 0, x_2 \geq 0$, *and* $2x_1 + x_2 = 4$. *So* $(x_1, x_2) = (0, 4)$. *From* $\lambda_1 - 2\lambda_3 = 0$ *and* $-2 - \lambda_3 = 0$, *we have* $(\lambda_1, \lambda_2, \lambda_3) = (-4, 0, -2)$. *Impossible.*

(iv) $\lambda_1 = \lambda_2 = 0$ *and* $\lambda_3 > 0 \Longrightarrow x_1 \geq 0, x_2 \geq 0$, *and* $2x_1 + x_2 = 4$. *So from* $2x_1 - 2\lambda_3 = 0, 2(x_2 - 5) - \lambda_3 = 0$, *and* $2x_1 + x_2 = 4$, *we have* $(x_1, x_2) = (-2/5, 24/5)$. *Impossible.*

(v) $\lambda_1 > 0, \lambda_2 > 0$ *and* $\lambda_3 = 0 \Longrightarrow x_1 = 0, x_2 = 0$, *and* $2x_1 + x_2 \leq 4$. *So* $(x_1, x_2) = (0, 0)$. *From* $\lambda_1 = 0$ *and* $-10 + \lambda_2 = 0$, *we have* $(\lambda_1, \lambda_2, \lambda_3) = (0, 10, 0)$. *Impossible.*

(vi) $\lambda_1 = \lambda_3 = 0, \lambda_2 > 0 \Longrightarrow x_1 \geq 0, x_2 = 0$, *and* $2x_1 + x_2 \leq 4$. *From* $2x_1 = 0$ *and* $-10 + \lambda_2 = 0$, *we have* $(x_1, x_2) = (0, 0)$ *and* $(\lambda_1, \lambda_2, \lambda_3) = (0, 10, 0)$.

(vii) $\lambda_1 > 0, \lambda_2 = \lambda_3 = 0 \Longrightarrow x_1 = 0, x_2 \geq 0$, *and* $2x_1 + x_2 \leq 4$. *So from* $\lambda_1 = 0$ *and* $2(x_2 - 5) = 0$, *we have* $(x_1, x_2) = (0, 5)$ *and* $(\lambda_1, \lambda_2, \lambda_3) = (0, 0, 0)$. *Impossible.*

(viii) $\lambda_1 = \lambda_2 = \lambda_3 = 0 \Longrightarrow x_1 \geq 0, x_2 \geq 0$, *and* $2x_1 + x_2 \leq 4$. *So from* $2x_1 = 0$ *and* $2(x_2 - 5) = 0$, *we have* $(x_1, x_2) = (0, 5)$. *Impossible.*

Candidate maximizers are $(2, 0)$ *and* $(0, 0)$. *The objective function values at these two candidates are 29 and 25, so* $(2, 0)$ *is the maximizer and the associated Lagrange multipliers are* $(0, 12, 2)$. $\square$

- Caution: Never blindly apply the Kuhn-Tucker conditions.

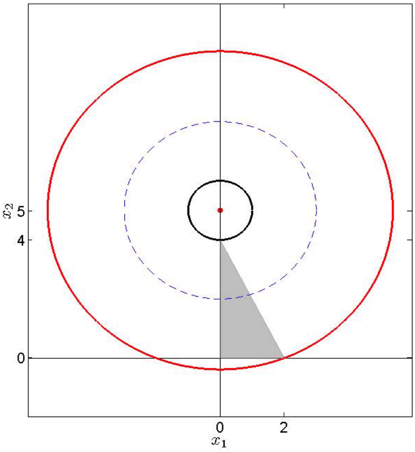


Figure: Intuitive Illustration of Example

A "Cookbook" Procedure of Optimization

- Define the **feasible set** of the general mixed constrained maximization problem as

$$G = \left\{\mathbf{x} \in \mathbb{R}^n | \mathbf{g}(\mathbf{x}) \geq \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0}\right\}.$$

Step 1: Apply the Weierstrass theorem to show that the maximum exists. If the feasible set G is compact, this is usually straightforward; if G is not compact, truncate G to a compact set, say G_o, such that there is a point $x_o \in G_o$ and $f(x_o) > f(x)$ for all $x \in G \backslash G_o$.

Step 2: Check whether the constraint qualification is satisfied. If not, denote the set of possible violation points as Q.

Step 3: Set up the Lagrangian and find the critical points. Denote the set of critical points as R.

Step 4: Check the value of f on $Q \cup R$ to determine the maximizer or maximizers.

Caution

- It is quite often for practitioners to apply Step 3 directly to find the maximizer. Although this may work in most cases, it is possible to fail in some cases.
- First, the Lagrangian may fail to have any critical points due to nonexistence of maximizers or failure of constraint qualification.
- Second, even if the Lagrangian does have one or more critical points, this set of critical points need not contain the solution still due to these two reasons.
- Let us repeat our caveat, "Never blindly apply the Kuhn-Tucker conditions"!

- This cookbook procedure works well in most cases, especially when the set $Q \cup R$ is small, e.g., $Q \cup R$ includes only a few points.
- If this set is large, it is better to employ more necessary conditions (e.g., the second order conditions (SOCs)) to screen the points in $Q \cup R$.
- Another solution is to employ sufficient conditions, i.e., as long as $\mathbf{x}^*$ satisfies these conditions, it must be the maximizer.
  - Sufficient conditions are very powerful especially combined with the uniqueness result because as long as we find one solution, it is **the** solution and we can stop.
- These topics are the main theme of the next chapter.