Equality Constraints and the Theorem of Lagrange

Constrained Optimization Problems.

- It is rare that optimization problems have unconstrained solutions. Usually some or all the constraints matter.
- Before we begin our study of th solution of constrained optimization problems, we first put some additional structure on our constraint set $\mathcal{D}$ and make a few definitions.

Definition. Let U be an open subset of $\mathbb{R}^n$.

- An *equality constrained optimization problem* is an optimization problem in which the constraint set $\mathcal{D}$ can be represented as

  $$\mathcal{D} = U \cap \{x \in \mathbb{R}^n \mid g(x) = 0\},$$

  where $g : \mathbb{R}^n \to \mathbb{R}^k$. We refer to the functions $g = (g_1, \dots, g_k)$ as *equality constraints*.

- An *inequality constrained optimization problem* is an optimization problem in which the constraint set $\mathcal{D}$ can be represented as

  $$\mathcal{D} = U \cap \{x \in \mathbb{R}^n \mid h(x) \geq 0\},$$

  where $h : \mathbb{R}^n \to \mathbb{R}^l$. We refer to the functions $h = (h_1, \dots, h_l)$ as *inequality constraints*.

- An *optimization problem with mixed constraints* is an optimization problem in which the constraint set $\mathcal{D}$ can be represented as

  $$\mathcal{D} = U \cap \{x \in \mathbb{R}^n \mid g(x) = 0 \text{ and } h(x) \geq 0\},$$

  where there are both equality and inequality constraints. ▲

- The given specifications of the constraint set $\mathcal{D}$ are very general.
- For instance, nonnegativity constraints can be easily handled. If a problem requires that $x \in \mathbb{R}^n_+$, we can simply define functions $h_j : \mathbb{R}^n \to \mathbb{R}$ by

  $$h_j(x) = x_j, \qquad j = 1, \dots, n,$$

  and use the n inequality constraints

  $$h_j(x) \geq 0, \qquad j = 1, \dots, n.$$

- Similarly, we can rewrite the constraints on the left as on the right.

$$\begin{aligned} \alpha(x) \geq a \quad &\Leftrightarrow \quad \alpha(x) - a \geq 0, \\ \beta(x) \leq b \quad &\Leftrightarrow \quad b - \beta(x) \geq 0, \\ \phi(x) = c \quad &\Leftrightarrow \quad c - \phi(x) = 0. \end{aligned}$$

Example 1. Consider the budget set

$$\mathcal{B}(p, I) = \{x \in \mathbb{R}^n_+ \mid p \cdot x \leq I\}$$

of the utility maximization problem.

- This can be represented using $n+1$ inequality constraints. Define $h : \mathbb{R}^n \to \mathbb{R}^{n+1}$ by

$$\begin{aligned} h_j(x) &= x_j, & j &= 1, \ldots, n, \\ h_k(x) &= I - p \cdot x & k &= n+1. \end{aligned}$$

- Then
$$\mathcal{B}(p, I) = \{x \in \mathbb{R}^n \mid h_j(x) \geq 0, j = 1 \ldots, n+1\}$$
is the budget set written in the specified form. ♦

Equality Constraints and the Theorem of Lagrange.

- We will study equality constrained problems and begin with a graphical "proof" of the theorem of Lagrange with two choice variables and one constraint.
- Consider the problem

$$\text{Maximize } f(x_1, x_2) \text{ subject to } g(x_1, x_2) = 0,$$

  Here the constraint set is $\mathcal{D} = \{(x_1, x_2) \in \mathbb{R}^2 \mid g(x_1, x_2) = 0\}$.

- We will look at this problem graphically.
  - First we draw the constraint set $\mathcal{D}$ in the x_1x_2-plane. It is represented by the red line.
  - Then we draw, in blue, level curves of the objective function f.
  - Our goal is to find the highest level curve of f which meets the constraint set.
  - The highest level curve of f cannot cross the constraint curve $\mathcal{D}$. If it did, as at point b, nearby higher level sets would cross too.
  - Thus the highest level curve of f to touch the constraint set $\mathcal{D}$ must be tangent at the constrained maximizer, x^*.

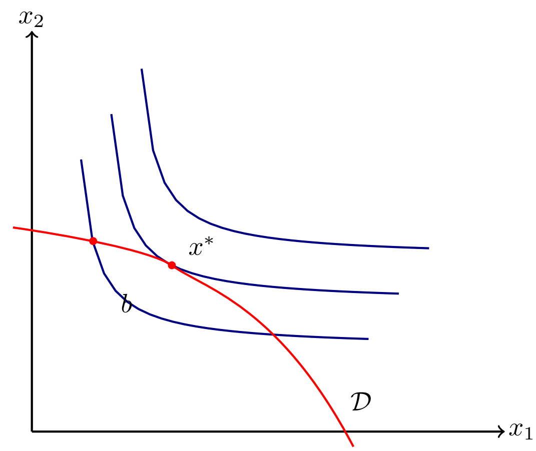

Figure 1: At the constrained maximizer x^*, the highest level curve of f is tangent to the constraint set $\mathcal{D}$.

- We can use our knowledge of the implicit function theorem to represent this condition mathematically.
- Since the level curve of f is tangent to the constraint set $\mathcal{D}$ at the constrained maximizer x^*, the slopes of the level set of f and of the constraint curve must be equal at x^*
- Since the level set at x^* is given by the equation $f(x_1, x_2) = f(x_1^*, x_2^*)$, we can use the implicit function theorem to calculate its slope as

$$-\frac{(\partial f/\partial x_1)(x^*)}{(\partial f/\partial x_2)(x^*)}.$$

- Similarly the constraint set is given by the implicit function $g(x_1, x_2) = 0$ and so its slope at x^* is

$$-\frac{(\partial g/\partial x_1)(x^*)}{(\partial g/\partial x_2)(x^*)}.$$

- Since these two slopes are equal at x^* we have

$$-\frac{(\partial f/\partial x_1)(x^*)}{(\partial g/\partial x_1)(x^*)} = -\frac{(\partial f/\partial x_2)(x^*)}{(\partial g/\partial x_2)(x^*)} = \lambda^*. \tag{1}$$

- We can rewrite this as two equations

$$\frac{\partial f}{\partial x_1}(x^*) + \lambda^* \frac{\partial g}{\partial x_1}(x^*) = 0,$$
$$\frac{\partial f}{\partial x_2}(x^*) + \lambda^* \frac{\partial g}{\partial x_2}(x^*) = 0.$$

- Since we have to solve for three unknowns, (x_1, x_2, λ), we need three equations. The third is the constraint equation $g(x_1, x_2) = 0$.

- So we have a system of three equations in three unknowns:

$$\frac{\partial f}{\partial x_1}(x) + \lambda \frac{\partial g}{\partial x_1}(x) = 0,$$
$$\frac{\partial f}{\partial x_2}(x) + \lambda \frac{\partial g}{\partial x_2}(x) = 0,$$
$$g(x_1, x_2) = 0.$$

- A convenient way of writing this is to form the *Lagrangean (function)*

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g(x_1, x_2).$$

  The critical points of the Lagrangian, are found by computing $\partial L/\partial x_1$, $\partial L/\partial x_2$ and $\partial L/\partial \lambda$ and setting them equal to zero. But this gives the system of three equations above.

- We have transformed a two variable constrained problem into an unconstrained problem of three variables.

- Note, in equation (1), we need $\partial g/\partial x_1$ and/or $\partial g/\partial x_2$ to be nonzero at the constrained maximizer x^*. This restriction is called the *constraint qualification*.

Theorem 1. *Let $f, g : \mathbb{R}^2 \to \mathbb{R}$ be C^1 functions. Suppose that $x^* = (x_1^*, x_2^*)$ is a local maximizer or minimizer of f subject to $g(x_1, x_2) = 0$. Suppose also that $Dg(x_1^*, x_2^*) \neq 0$. Then, there exists a scalar $\lambda^* \in \mathbb{R}$ such that $(x_1^*, x_2^*, \lambda^*)$ is a critical point of the Lagrangean*

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g(x_1, x_2).$$

In other words, at $(x_1^*, x_2^*, \lambda^*)$

$$\frac{\partial L}{\partial x_1} = 0, \ \frac{\partial L}{\partial x_2} = 0 \textit{ and } \frac{\partial L}{\partial \lambda} = 0.$$

- We now present the theorem of Lagrange in the general case of optimization of a function in n variables subject ot k equality constraints.

Theorem 2 (The Theorem of Lagrange)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ and $g_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \ldots, k$ be C^1 functions. Suppose x^* is a local maximizer or minimizer of f on the constraint set*

$$\mathcal{D} = U \cap \{x \in \mathbb{R}^n \mid g(x) = 0\},$$

where $U \subseteq \mathbb{R}^n$ is open. Suppose also that $\operatorname{rank} Dg(x^*) = k$. Then, there exists a vector $\lambda^* = (\lambda_1^*, \ldots, \lambda_k^*) \in \mathbb{R}^k$ such that (x^*, λ^*) is a critical point of the Lagrangean

$$L(x, \lambda) = f(x) + \sum_{i=1}^{k} g_i(x).$$

That is

$$Df(x^*) + \sum_{i=1}^{k} \lambda_i^* Dg_i(x^*) = 0.$$

or writing out the system explicitly

$$\begin{aligned} \frac{\partial L}{\partial x_j}(x^*, \lambda^*) &= 0, && j = 1, \ldots, n \\ \frac{\partial L}{\partial \lambda_i}(x^*, \lambda^*) &= 0, && i = 1, \ldots, k. \end{aligned}$$

- The theorem of Lagrange only provides *necessary* conditions for local optima x^*. Furthermore, these conditions only apply to those local optima x^* which also meet the constraint qualification $\operatorname{rank} Dg(x^*) = k$.
- These conditions are not *sufficient*. That is, the theorem does *not* say that if there exists (x, λ) such that $g(x) = 0$ and $Df(x) + \sum_{i=1}^{k} \lambda_i Dg_i(x) = 0$, then x must be a local maximum or a local minimum even if it also meets the constraint qualification.
- The following example shows that the conditions of the theorem cannot be sufficient.

Example 2. Let f and g be functions defined by $f(x, y) = x^3 + y^3$ and $g(x, y) = x - y$ and consider the equality constrained problem of maximizing or minimizing $f(x, y)$ over the constraint set $\mathcal{D} = \{(x, y) \in \mathbb{R}^2 \mid g(x, y) = 0\}$.

- Let $(x^*, y^*) = (0, 0)$ and let $\lambda^* = 0$. Then $g(x^*, y^*) = 0$, so that (x^*, y^*) is a feasible point.
- It also meets the constraint qualification, since $Dg(x, y) = (1 \;\; -1)$ for any (x, y).
- Finally, since $Df(x, y) = (3x^2 \; 3y^2)$, we have

  $$Df(x^*, y^*) + \lambda^* Dg(x^*, y^*) = (0\;0) + 0(1 \;\; -1) = (0\;0).$$

- Hence, if the conditions of the theorem of Lagrange were also sufficient, then (x^*, y^*) would be either a local minimizer or maximizer of f on $\mathcal{D}$. However, we have neither.
- We have $f(x^*, y^*) = 0$. But for every $\varepsilon > 0$, it is the case that $(-\varepsilon, -\varepsilon) \in \mathcal{D}$ and $(\varepsilon, \varepsilon) \in \mathcal{D}$. Futhermore

  $$f(-\varepsilon, -\varepsilon) = -2\varepsilon^3 < f(x^*, y^*)$$

  and

  $$f(\varepsilon, \varepsilon) = 2\varepsilon^3 > f(x^*, y^*),$$

  so that (x^*, y^*) is not a local maximizer or minimizer. ♦

The Lagrangean Multipliers.

- The vector $\lambda^* = (\lambda_1^*, \ldots, \lambda_k^*)$ in the theorem of Lagrange is called the vector of *Lagrangean multipliers* corresponding to the local optimum x^*.
- The ith multiplier λ_i^* measures the sensitivity of the value of the objective function at x^* to a small relaxation of the ith constraint g_i.
- For simplicity, we will look at the optimization of a function f in two variables subject to one constraint, which we will write as $g(x, y, c) = c - g(x, y) = 0$. Then a relaxation of the constraint corresponds to an increase in c.
- The Lagrangean for this problem is

$$L(x, y, \lambda, c) = f(x, y) + \lambda[c - g(x, y)]$$

  with c entering as a parameter.
- By Lagrange's theorem, a local optimizer $(x^*(c), y^*(c), \lambda^*(c))$ satisfies

$$\begin{aligned}
\frac{\partial f}{\partial x}(x^*(c), y^*(c)) - \lambda^* \frac{\partial g}{\partial x}(x^*(c), y^*(c)) = 0, \\
\frac{\partial f}{\partial y}(x^*(c), y^*(c)) - \lambda^* \frac{\partial g}{\partial y}(x^*(c), y^*(c)) = 0.
\end{aligned}$$

- Also, since $g(x^*(c), y^*(c)) = c$ for all c, we have

$$\frac{\partial g}{\partial x}(x^*(c), y^*(c))\frac{dx^*}{dc}(c) + \frac{\partial g}{\partial y}(x^*(c), y^*(c))\frac{dy^*}{dc}(c) = 1$$

- Now by the chain rule, and the equations above, we have

$$\begin{aligned}
\frac{df}{dc}(x^*(c), y^*(c)) &= \frac{\partial f}{\partial x}(x^*(c), y^*(c))\frac{dx^*}{dc} + \frac{\partial f}{\partial y}(x^*(c), y^*(c))\frac{dy^*}{dc}. \\
&= \lambda^* \\
&= \frac{\partial L^*}{\partial c}(x^*(c), y^*(c), \lambda^*(c)) \text{ (Envelope Theorem)}
\end{aligned}$$

Example 3. Consider the consumer's utility maximization problem in which the budget constraint holds with equality:

$$\text{Maximize } u(x_1, x_2) \text{ subject to } I - p_1x_1 - p_2x_2 = 0.$$

- Here we can look at relaxing the budget constraint by increasing income by one unit.
- By the preceding calculations, at the optimum, an increase in income by one unit will raise utility by λ^* units, where λ^* is the Lagrange multiplier at the optimum.
- Thus λ^* represents the consumer's marginal utility of income.

Second Order Conditions.

- Now we present the second order conditions for two variable optimization subject to a single equality constraint.

Theorem 3. *Let $f, g : \mathbb{R}^2 \to \mathbb{R}$ be C^2 functions. Let $\mathcal{D} = \{(x, y) \in \mathbb{R}^2 \mid g(x, y) = 0$. Suppose that at (x^*, y^*, λ^*)*

$$\frac{\partial L}{\partial x_1} = 0, \ \frac{\partial L}{\partial x_2} = 0 \text{ and } \frac{\partial L}{\partial \lambda} = 0.$$

Let $D^2_{(\lambda,x,y)} L^*$ denote the Hessian of the Lagrangian at (x^*, y^*, λ^*)

$$\overline{H} = D^2_{(\lambda,x,y)} L^* = \begin{pmatrix} 0 & g_x & g_y \\ g_x & L_{xx} & L_{xy} \\ g_y & L_{yx} & L_{yy} \end{pmatrix}.$$

1. *If $|\overline{H}| > 0$, then (x^*, y^*) is a strict local maximizer of f on $\mathcal{D}$.*
2. *If $|\overline{H}| < 0$, then (x^*, y^*) is a strict local minimizer of f on $\mathcal{D}$.*

- Next we consider the more general problem of maximizing or minimizing $f : \mathbb{R}^n \to \mathbb{R}$ over the set
$$\mathcal{D} = U \cap \{x \mid g(x) = 0\},$$
where $g : \mathbb{R}^n \to \mathbb{R}^k$, and U is an open subset of $\mathbb{R}^n$. We will assume that f and g are both C^2 functions.
- We form the Lagrangian $L(x, \lambda) = f(x) + \sum_{i=1}^k \lambda_i g_i(x)$.
- The second derivative $D^2_x L(x, \lambda)$ of $L(\cdot, \lambda)$ with respect to the x variables is the $n \times n$ matrix defined by
$$D^2_x L(x, \lambda) = D^2 f(x) + \sum_{i=1}^k \lambda_i D^2 g_i(x).$$
- Since f and g are C^2 functions, so is $L(\cdot, \lambda)$ for any given $\lambda \in \mathbb{R}^k$. Thus, $D^2_x L(x, \lambda)$ is a symmetric matrix and defines a quadratic form on $\mathbb{R}^n$.

Theorem 4. *Suppose there exist points $x^* \in \mathcal{D}$ and $\lambda^* \in \mathbb{R}^k$ such that $\operatorname{rank} Dg(x^*) = k$ and $Df(x^*) + \sum_{i=1}^k \lambda_i^* Dg_i(x^*) = 0$. Define*

$$\mathcal{Z}(x^*) = \{z \in \mathbb{R}^n \mid Dg(x^*)z = 0\}$$

and let $D^2_x L^*$ denote $D^2_x L(x^*, \lambda^*) = D^2 f(x^*) + \sum_{i=1}^k \lambda_i^* D^2 g_i(x^*)$.

1. *If f has a local maximum at x^*, then $z^T (D^2_x L^*) z \leq 0$ for all $z \in \mathcal{Z}(x^*)$.*
2. *If f has a local minimum at x^*, then $z^T (D^2_x L^*) z \geq 0$ for all $z \in \mathcal{Z}(x^*)$.*

3. *If $z^T(D_x^2 L^*)z < 0$ for all $z \in \mathcal{Z}(x^*)$ with $z \neq 0$, then x^* is a strict local maximizer of f on $\mathcal{D}$.*

4. *If $z^T(D_x^2 L^*)z > 0$ for all $z \in \mathcal{Z}(x^*)$ with $z \neq 0$, then x^* is a strict local minimizer of f on $\mathcal{D}$.*

- Note the similarity between the conditions of the theorem and the corresponding theorem for unconstrained maximization problems. There are two important differences.
  - Here the second order conditions are stated in terms of the second derivatives of the Lagrangean instead of the function f.
  - The properties of the quadratic form $D_x^2 L(x^*, \lambda^*)$ are only required to hold on a subset of $\mathbb{R}^n$ defined by $\mathcal{Z}(x^*)$.
- How do we verify the definiteness of $D_x^2 L(x^*, \lambda^*)$ on the constraint set $\mathcal{Z}$?
- We form a $(k+n) \times (k+n)$ matrix consisting of the Hessian of the Lagrangean with respect to x bordered by the Jacobian of the constraint functions g. This is sometimes called a *bordered Hessian*.
- It is the Hessian of the Lagrangean with respect to both λ and x.
- We will denote this matrix by $\overline{H}$:

$$\begin{aligned}
\overline{H} &= D^2 L_{(\lambda,x)}(x^*, \lambda^*) \\
&= \begin{pmatrix} 0 & Dg(x^*) \\ Dg(x^*) & D_x^2 L(x^*, \lambda^*) \end{pmatrix} \\
&= \left(\begin{array}{ccc|ccc}
0 & \cdots & 0 & \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & \frac{\partial g_k}{\partial x_1} & \cdots & \frac{\partial g_k}{\partial x_n} \\
\hline
\frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_k}{\partial x_1} & \frac{\partial^2 L}{\partial x_1^2} & \cdots & \frac{\partial^2 L}{\partial x_n \partial x_1} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\frac{\partial g_1}{\partial x_n} & \cdots & \frac{\partial g_k}{\partial x_n} & \frac{\partial^2 L}{\partial x_1 \partial x_n} & \cdots & \frac{\partial^2 L}{\partial x_n^2}
\end{array} \right).
\end{aligned}$$

- If the determinant of $\overline{H}$ has the same sign as $(-1)^n$, and the *last* $n-k$ leading principal minors of $\overline{H}$ *alternate* in sign, then the condition in part (3) of the theorem holds and x^* is a local maximizer.
- If the last $n-k$ leading principal minors of $\overline{H}$ have the same sign as $(-1)^k$, then the condition in part (4) holds and x^* is a local minimizer.
- If both of the above conditions on $\overline{H}$ are violated by *nonzero* leading principal minors, then (1) and (2) do not hold and x^* is neither a local maximizer nor a local minimizer.

- We summarize these results in the following theorem.

Theorem 5. *Suppose there exist points $x^* \in \mathcal{D}$ and $\lambda^* \in \mathbb{R}^k$ such that* $\operatorname{rank} Dg(x^*) = k$ *and* $Df(x^*) + \sum_{i=1}^{k} \lambda_i^* Dg_i(x^*) = 0$. *Consider the bordered Hessian given above. Let $\overline{H}_r$ denote the rth order leading principal submatrix of $\overline{H}$.*

1. *If $(-1)^{r-k}|\overline{H}_r| > 0$ for all $r = 2k+1, \dots, n+k$, then x^* is a strict local maximizer of f on $\mathcal{D}$.*
2. *If $(-1)^{k}|\overline{H}_r| > 0$ for all $r = 2k+1, \dots, n+k$, then x^* is a strict local minimizer of f on $\mathcal{D}$.*
3. *If either of the above conditions is violated by nonzero leading principal minors, then x^* is neither a local maximizer nor a local minimizer.*

Using the Theorem of Lagrange.

- We now describe a "cookbook" procedure for using the theorem of Lagrange to solve a maximization problem.
- Consider an equality constrained optimization problem of the form

$$\text{Maximize } f(x) \text{ subject to } x \in \mathcal{D} = U \cap \{x \mid g(x) = 0\},$$

  where $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}^k$ are C^1 functions and U is an open subset of $\mathbb{R}^n$.

1. Set up a function $L : \mathcal{D} \times \mathbb{R}^k \to \mathbb{R}$, called the *Lagrangean* defined by

$$L(x, \lambda) = f(x) + \sum_{i=1}^{k} \lambda_i g_i(x).$$

   The vector $\lambda = (\lambda_1, \dots, \lambda_k) \in \mathbb{R}^k$ is called the vector of *Lagrange multipliers*.

2. Find the set of all critical points of L for which $x \in U$ i.e. all points (x, λ) at which $DL(x, \lambda) = 0$ and $x \in U$. Since $x \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}^k$, this condition results in a system of $n + k$ equations in $n + k$ unknowns:

$$\begin{aligned} \frac{\partial L}{\partial x_j}(x, \lambda) &= 0, & j &= 1, \dots, n \\ \frac{\partial L}{\partial \lambda_i}(x, \lambda) &= 0, & i &= 1, \dots, k. \end{aligned}$$

   Let M be the set of solutions to these equations for which $x \in U$:

$$M = \{(x, \lambda) \mid x \in U \text{ and } DL(x, \lambda) = 0\}.$$

3. Finally, evaluate f at each point x in the set

$$\{x \in \mathbb{R}^n \mid \text{there is some } \lambda \text{ such that } (x, \lambda) \in M\}.$$

   The values of x which maximize f over this set are also usually solutions to the equality constrained maximization problem.

Theorem 6. *Suppose the following two conditions hold.*

1. *A global optimum x^* exists to the given problem.*
2. *The constraint qualification is met at x^*.*

Then there exists a λ^* such that (x^*, λ^*) is a critical point of L.

- Under the two conditions above, the Lagrangean method wil be successful in finding the optimum x^*.
- This result also explains why the Lagrangean method usually works in practice.
  - The existence of a solution is usually not a problem (check using Weierstrass theorem) and neither is the constraint qualification.
  - Although, in general, it is not possible to verify that the constraint qualification holds beforehand, it is often the case that the constraint qualification holds everywhere on the feasible set $\mathcal{D}$.
  - In particular, if there is a single linear constraint and two choice variables, the constraint qualification will hold at all $x \in \mathcal{D}$.
- Unfortunately if the conditions of theorem (6) fail to hold, the procedure can also fail to identify global optima.
  - First, if a global optimum exists but the constraint qualification is not met at the optimum, then the optimum will not be found among the set of critical points.
  - Second, even if the constraint qualification holds everywhere on $\mathcal{D}$, the procedure can fail simply because no global optimum exists.
- Most problems in economic theory involve inequality rather than equality constraints.
- However, *under suitable conditions*, it is possible to reduce inequality constrained problems to *equivalent* equality constrained problems. Then the theorem of Lagrange can be applied.

Example 4. Consider a utility maximization problem in which a consumer consumes two goods.

- The consumer's utility from consuming amount x_i of commodity $i = 1, 2$, is given by $u(x_1, x_2) = x_1 x_2$.

- The consumer has an income $I > 0$, and the price of commodity i is $p_i > 0$.
- Thus, the problem is to solve

$$\max\{x_1 x_2 \mid I - p_1 x_1 - p_2 x_2 \geq 0, x_1 \geq 0, x_2 \geq 0\}$$

- We will proceed in three steps.

1. We begin by reducing the utility maximization problem to an equality constrained problem.

- First note that the budget set

$$\mathcal{B}(p, I) = \{(x_1, x_2) \mid I - p_1 x_1 - p_2 x_2 \geq 0, x_1 \geq 0, x_2 \geq 0\}$$

  is a compact set and the utility function is continuous on this set.
- Thus, by the Weierstrass theorem, a solution (x_1^*, x_2^*) does exist to the given maximization problem.
- Now, if either $x_1 = 0$ or $x_2 = 0$, then $u(x_1, x_2) = 0$.
- However, the consumption point $(\overline{x}_1, \overline{x}_2) = (I/2p_1, I/2p_2)$, which divides income equally between the two commodities is feasible, and satisfies $u(\overline{x}_1, \overline{x}_2) = \overline{x}_1 \overline{x}_2 > 0$.
- Since any solution (x_1^*, x_2^*) must satisfy $u(x_1^*, x_2^*) \geq u(\overline{x}_1, \overline{x}_2)$, it follows that any solution must satisfy $x_i^* > 0$, $i = 1, 2$.
- Furthermore, any solution must satisfy the budget constraint with equality, or total utility could be increased.
- Thus, we can see that (x_1^*, x_2^*) is a solution to the original problem iff it is a solution to the problem

$$\max\{x_1 x_2 \mid I - p_1 x_1 - p_2 x_2 = 0, x_1 > 0, x_2 > 0\}$$

- The constraint set of this reduced problem, which we will denote by $\mathcal{B}^*(p, I)$, can be written as

$$\mathcal{B}^*(p, I) = \mathbb{R}^2_{++} \cap \{(x_1, x_2) \mid I - p_1 x_1 - p_2 x_2 = 0\}$$

  and by setting $U = \mathbb{R}^2_{++}$ and $g(x_1, x_2) = I - p_1 x_1 - p_2 x_2$ we can use the theorem of Lagrange.

2. Next we obtain the critical points of the Lagrangean.

- We first set up the Lagrangean

$$L(x_1, x_2, \lambda) = x_1 x_2 + \lambda(I - p_1 x_1 - p_2 x_2).$$

- The critical points of L are the solutions $(x_1^*, x_2^*, \lambda) \in \mathbb{R}^2_{++} \times \mathbb{R}$ to

$$\frac{\partial L}{\partial x_1} = x_2 - \lambda p_1 = 0$$
$$\frac{\partial L}{\partial x_2} = x_1 - \lambda p_2 = 0$$
$$\frac{\partial L}{\partial \lambda} = I - p_1 x_1 - p_2 x_2 = 0$$

- If $\lambda = 0$, this system of equations has no solution, since then we need $x_1 = x_2 = 0$ from the first two equations.

- So suppose $\lambda \neq 0$. From the first two equations, we then find $\lambda = x_1/p_1 = x_2/p_2$, so that $x_1 = p_2 x_2/p_1$. Using this in the third equation, we obtain the unique solution to the set of equations: $x_1^* = I/2p_1$, $x_2^* = I/2p_2$ and $\lambda^* = I/2p_1 p_2$.

3. Now we classify the critical points of the Lagrangean.

- To classify the single critical point of L we will apply the second order conditions to check that (x_1^*, x_2^*) is a strict local maximum of u on $\mathcal{B}^*(p, I)$.

- First, note that $Dg(x_1^*, x_2^*) = (-p_1 \;\; -p_2)$, so we have

$$\mathcal{Z}(x^*) = \{z \in \mathbb{R}^2 \mid Dg(x^*)z = 0\} = \left\{ z \in \mathbb{R}^2 \mid z_1 = -\frac{p_2 z_2}{p_1} \right\}.$$

- Define $D_x^2 L^* = D^2 u(x^*) + \lambda^* D^2 g(x^*)$. Then we have

$$D_x^2 L^* = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \lambda^* \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

  So that for any $z \in \mathbb{R}^2$, we have $z^T (D_x^2 L^*) z = 2 z_1 z_2$. Thus, for any $z \in \mathcal{Z}(x^*)$ with $z \neq 0$, we have $z^T (D_x^2 L^*) z = -2 p_2 z_2^2 / p_1 < 0$.

- Alternatively we can check the determinant of $D^2_{(\lambda,x)}(x^*)$.

$$|D^2_{(\lambda,x)} L(x^*, \lambda^*)| = \begin{vmatrix} 0 & -p_1 & -p_2 \\ -p_1 & 0 & 1 \\ -p_2 & 1 & 0 \end{vmatrix} = 2 p_1 p_2 > 0$$

- Both methods show that (x_1^*, x_2^*) satisfies the second order conditions for a strict local maximizer of u on $\mathcal{B}^*(p, I)$.

- We can actually show the stronger result that (x_1^*, x_2^*) is a *global* maxmizer on $\mathcal{B}^*(p, I)$, by showing that the conditions of theorem (6) hold.

- First, note that a global maximum exists by the argument in step 1.

- Next, note that the single constraint $g(x_1, x_2)$ satisfies $Dg(x_1, x_2) = (-p_1 \; p_2) \neq 0$ everywhere on $\mathcal{B}^*(p, I)$. Hence $\operatorname{rank} Dg(x_1, x_2) = 1$ at all $(x_1, x_2) \in \mathcal{B}^*(p, I)$ and the constraint qualification holds at the global maximum.

- Therefore, by theorem (6), the global maximum must be a critical point of the Lagrangean. Since there is only one critical point (x_1^*, x_2^*), this must be the problem's global maximum. ⧫