18. Constrained Optimization I: First Order Conditions

The typical problem we face in economics involves optimization under constraints. From supply and demand alone we have: maximize utility, subject to a budget constraint and non-negativity constraints; minimize cost, subject to a quantity constraint; minimize expenditure, subject to a utility constraint; maximize profit, subject to constraints on production. And those are just the basic supply and demand related problems. Then there are other types of constrained optimization ranging from finding Pareto optima given resource and technology constraints to setting up incentive schemes subject to participation constraints.

The generic constrained optimization problem involves a thing to be optimized, the *objective function*, and one or more *constraint functions* used to define the constraints.

It often looks something like this.

$$\begin{aligned} &\max_{\mathbf{x}} \; f(\mathbf{x}) \\ &\text{s.t. } g_i(\mathbf{x}) \le b_i, \text{ for } i = 1, \ldots, k \\ &\qquad h_j(\mathbf{x}) = c_j, \text{ for } j = 1, \ldots, \ell \end{aligned}$$

Here f is the *objective function*, the equations $g_i(\mathbf{x}) \le b_i$ are referred to as *inequality constraints*, while the equations $h_j(\mathbf{x}) = c_j$ are called *equality constraints*.[1]

[1] The letters “s.t.” can be read as “subject to”, “such that”, or “so that”.

18.1 Two Optimization Problems of Economics

Sometimes a problem will have only one kind of constraint. An example is the following consumer's problem defining the *indirect utility function* $v(\mathbf{p}, m)$. It only has inequality constraints.

$$
\begin{aligned}
v(\mathbf{p}, m) = \max_{\mathbf{x}} \; & u(\mathbf{x}) \\
\text{s.t. } & \mathbf{p}\cdot\mathbf{x} \le m \\
& \mathbf{x} \ge \mathbf{0}.
\end{aligned}
$$

Here $\mathbf{p}$ is the price vector, u is the utility function, and m is income. All the constraints here are *inequality constraints*. The constraints $x_i \ge 0$ ($\mathbf{x} \ge \mathbf{0}$) are known as *non-negativity constraints*. The solutions $\mathbf{x}(\mathbf{p}, m)$ are the *Marshallian demands*.

Another example is the firm's cost minimization problem which defines the *cost function* $c(\boldsymbol{w}, q)$.

$$
\begin{aligned}
c(\boldsymbol{w}, q) = \min_{\boldsymbol{z}} \; & \boldsymbol{w}\cdot\boldsymbol{z} \\
\text{s.t. } & f(\boldsymbol{z}) \ge q, \boldsymbol{z} \ge \mathbf{0}.
\end{aligned}
$$

Here q is the amount produced, $\boldsymbol{w}$ is the vector of factor prices, and f is the production function. The solutions are the *conditional factor demands*, $\boldsymbol{z}(\boldsymbol{w}, q)$.

18.2 A Simple Consumer's Problem

We start by examining a simple consumer's problem with two goods and a single equality constraint, the budget constraint. This consumer's problem is

$$\max_{x} u(x_1, x_2)$$
$$\text{s.t. } p_1x_1 + p_2x_2 = m.$$

We're dropping the usual non-negativity constraints in the interest of simplifying the problem.

Consider the geometry of the solution. As we teach in undergrad micro, the indifference curve must be tangent to the budget line at the utility maximum. That is, the slopes of the two curves must be the same. This is illustrated in Figure 18.3.1. Notice that we have extended the budget line because we are not imposing any non-negativity constraints.

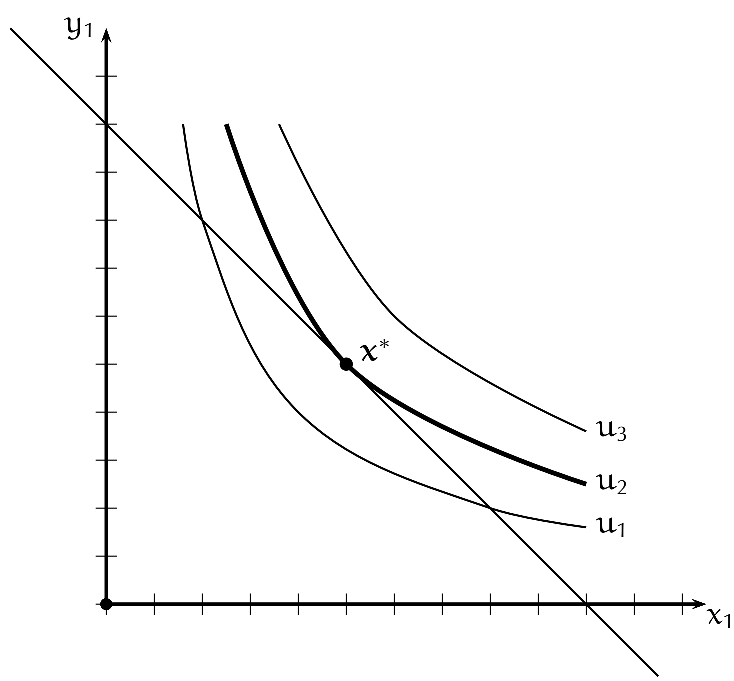

Figure 18.2.1: Three indifference curves are shown in the diagram. Indifference curve u_2 is the highest indifference curve the consumer can afford. This happens at x^* where the indifference curve is tangent to the budget line.

18.3 Solution to the Simple Consumer's Problems

Aligning the tangents for the budget line and indifference curve can be accomplished by making sure they have the same slope. Both slopes are negative. The absolute slope of the budget line is the relative price p_1/p_2 while the absolute slope of the indifference curve is the marginal rate of substitution, $\text{MRS}_{12} = (\partial u/\partial x_1)/(\partial u/\partial x_2)$. It follows that at the utility maximum $\boldsymbol{x}^*$,

$$\frac{p_1}{p_2} = \frac{\partial u/\partial x_1}{\partial u/\partial x_2}.$$

Another way to think about this is that we are lining up the tangent spaces of both the budget constraint and the optimal indifference curve. Both the budget line and the indifference curves are level sets of functions. As such, their tangent spaces are the null spaces of their derivatives.

The tangents $\boldsymbol{v}$ to the budget line obey $(p_1, p_2)\boldsymbol{v} = 0$ and the tangents $\boldsymbol{w}$ to the optimal indifference curve obey $[Du(\boldsymbol{x}^*)]\boldsymbol{w} = 0$. Since this is in $\mathbb{R}^2$, the normal vectors $\nabla u(\boldsymbol{x}^*)$ and $\mathbf{p} = (p_1, p_2)^{\mathsf{T}}$ must be collinear. Thus $\nabla u(\boldsymbol{x}^*) = \mu\mathbf{p}$ for some $\mu \in \mathbb{R}$.

We could alternatively write the problem as maximizing u over the line $-\mathbf{p}\cdot\boldsymbol{x} = -m$. It is the same set, but has a different derivative. We could use $\alpha\mathbf{p}\cdot\boldsymbol{x} = \alpha m$ for any $\alpha \neq 0$. Each α gives a different derivative to the constraint.

We don't have the freedom to alter the direction of $\nabla u(\boldsymbol{x}^*)$ because it points in the direction of maximum increase of u. Reversing the direction would reverse consumer preferences. What was better would become worse and vice-versa.

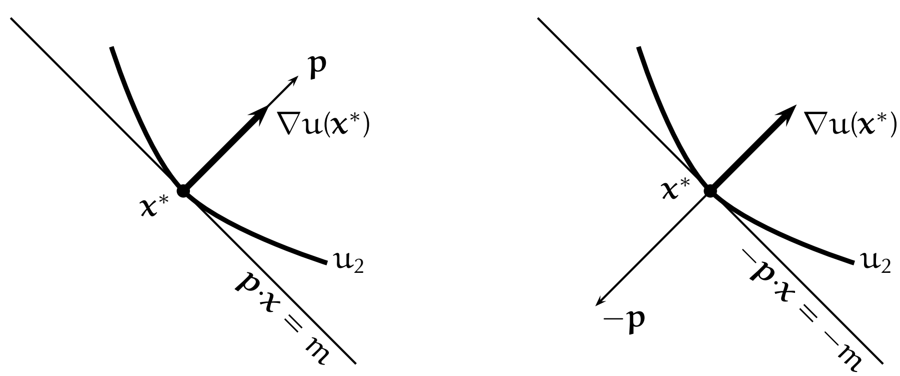

Figure 18.3.1: Here $\boldsymbol{x}^*$ is the consumer's utility maximum. At the maximum, the vectors $\mathbf{p}$ and $\nabla u(\boldsymbol{x}^*)$ are collinear, with the heavier arrow being $\nabla u(\boldsymbol{x}^*)$. With an equality constraint, it doesn't matter whether the vectors point in the same or opposite directions. It will matter for inequality constraints. The right panel writes the budget constraint as $-\mathbf{p}\cdot\boldsymbol{x} = -m$. Either way, there is a $\mu \in \mathbb{R}$ with $\nabla u(\boldsymbol{x}^*) = \mu\mathbf{p}$.

18.4 More on the Simple Consumer's Problem

If the tangents (or gradients) don't align, the choice $\mathbf{x}^*$ does not maximize utility. In that case there is a $\mathbf{v}$ with $\mathbf{p}{\cdot}\mathbf{v} = 0$ and $\nabla u(\mathbf{x}^*){\cdot}\mathbf{v} > 0$. Travelling along the tangent to the budget constraint (the budget constraint itself) in the direction $\mathbf{v}$ will increase utility.

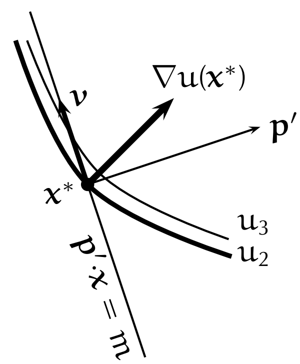

Figure 18.4.1: Here the prices have been changed to $\mathbf{p}'$ so that the indifference curve and budget line are no longer tangent. With the new prices, we can move along the budget line in the direction $\mathbf{v}$ to increase utility above u_2 and even u_3.

The situation is similar when $\mathbf{x}^*$ is a minimum.

18.5 Optimization Under Equality Constraints

We can use the tangent spaces to find a necessary condition for constrained optimization when the constraints are equality constraints.

Consider the problem

$$\max_{\mathbf{x}} f(\mathbf{x})$$
$$\text{s.t. } h_j(\mathbf{x}) = c_j, \text{ for } i = 1, \ldots, \ell.$$

That is, we are attempting to maximize $f(\mathbf{x})$ under the constraints that $\mathbf{h}(\mathbf{x}^*) = \mathbf{c}$. The key result is the Tangent Space Theorem.[2]

Tangent Space Theorem. *Let* $U \subset \mathbb{R}^m$ *and* $f\colon U \to \mathbb{R}$, $\mathbf{h}\colon \mathbb{R}^m \to \mathbb{R}^\ell$ *be* $\mathcal{C}^1$ *functions. Suppose* $\mathbf{x}^*$ *either maximizes or minimizes* $f(\mathbf{x})$ *over the set* $M = \{\mathbf{x} : \mathbf{h}(\mathbf{x}) = \mathbf{c}\}$ *with* $Df(\mathbf{x}^*) \neq \mathbf{0}$ *and* $\operatorname{rank} D\mathbf{h}(\mathbf{x}^*) = \ell < m$. *Then the tangent space of the differentiable manifold* F, $T_{\mathbf{x}^*}F = \{\mathbf{x} : f(\mathbf{x}) = f(\mathbf{x}^*)\}$ *at* $\mathbf{x}^*$ *contains in the tangent space at* $\mathbf{x}^*$ *of the differentiable manifold* M. *Moreover, there are unique* μ_i^*, $i = 1, \ldots, \ell$ *with* $\sum_{i=1}^{\ell} \mu_i^* Dh_j(\mathbf{x}^*) = Df(\mathbf{x}^*)$.

In our simple consumer's problem, $m = 2$ and $\ell = 1$, so both the indifference curve F and the budget constraint M are one-dimensional manifolds. The tangent spaces were also one-dimensional, and so had to coincide. Here the tangent space of F is $(m-1)$-dimensional and the tangent space of M is $(m-\ell)$-dimensional.

Because the dimensions are different if $\ell > 1$, the tangent spaces cannot then coincide. The best we can do in that line is for the smaller space $T_{\mathbf{x}^*}M$ to be contained in the larger space $T_{\mathbf{x}^*}F$.

The proof shows that if movement in a direction $\mathbf{v}$ is allowed by the constraints, and if $\mathbf{v}$ is not in $T_{\mathbf{x}^*}F$, then the objective f can be increased. This means that at the maximum, the constraints can only allows moves in directions in $T_{\mathbf{x}^*}F$.

The condition that $\operatorname{rank} D\mathbf{h}(\mathbf{x}^*) = \ell$ is called the *non-degenerate constraint qualification condition (NDCQ) at* $\mathbf{x}^*$. It ensures that M is a regular $(m-\ell)$-manifold by Theorem 15.18.1, which lets us find its tangent space.

Finally, the numbers $\mu_1^*, \ldots, \mu_\ell^*$ are called *Lagrange multipliers*.

---

[2] The name Tangent Space Theorem is not in general use. We do refer to it a number of times and I thought a name more informative than a number.

18.6 Proof of the Tangent Space Theorem

Tangent Space Theorem. *Let $U \subset \mathbb{R}^m$ and $f\colon U \to \mathbb{R}$, $\mathbf{h}\colon \mathbb{R}^m \to \mathbb{R}^\ell$ be $\mathcal{C}^1$ functions. Suppose $\mathbf{x}^*$ either maximizes or minimizes $f(\mathbf{x})$ over the set $M = \{\mathbf{x} : \mathbf{h}(\mathbf{x}) = \mathbf{c}\}$ with $Df(\mathbf{x}^*) \neq \mathbf{0}$ and $\operatorname{rank} D\mathbf{h}(\mathbf{x}^*) = \ell < m$. Then the tangent space of the differentiable manifold F, $T_{\mathbf{x}^*}F = \{\mathbf{x} : f(\mathbf{x}) = f(\mathbf{x}^*)\}$ at $\mathbf{x}^*$ contains in the tangent space at $\mathbf{x}^*$ of the differentiable manifold M. Moreover, there are unique μ_i^*, $i = 1, \ldots, \ell$ with $\sum_{i=1}^{\ell} \mu_i^* Dh_j(\mathbf{x}^*) = Df(\mathbf{x}^*)$.*

Proof. Suppose $\mathbf{x}^*$ is a maximum. Because $\operatorname{rank} D\mathbf{h}(\mathbf{x}^*) = \ell$, there is a neighborhood U of $\mathbf{x}^*$ where $\operatorname{rank} D\mathbf{h}(\mathbf{x}) = \ell$. By Theorem 15.18.1, we can define a chart (U, φ) from $U \cap M$ to $\mathbb{R}^{m-\ell}$. It follows that the tangent space $T_{\mathbf{x}^*}M = \ker D\mathbf{h}(\mathbf{x}^*)$ is $(m-\ell)$-dimensional.

Similarly, F is a $(m-1)$-dimensional differentiable manifold, with an $(n-1)$-dimensional tangent space $T_{\mathbf{x}^*}F$.

By way of contradiction, suppose there is $\boldsymbol{v} \in T_{\mathbf{x}^*}M$ with $\boldsymbol{v} \notin T_{\mathbf{x}^*}F$. Then $D\mathbf{h}(\mathbf{x}^*)\boldsymbol{v} = \mathbf{0}$ and $Df(\mathbf{x}^*)\boldsymbol{v} \neq 0$. The latter is a real number, so if $Df(\mathbf{x}^*)\boldsymbol{v} < 0$, we may replace $\boldsymbol{v}$ by $-\boldsymbol{v}$, ensuring $Df(\mathbf{x}^*)\boldsymbol{v} > 0$.

Take a curve $\mathbf{x}(t)$ from $(-1, 1)$ to M with $\mathbf{x}(0) = \mathbf{x}^*$ and $\mathbf{x}'(0) = \boldsymbol{v} \in T_{\mathbf{x}^*}M$. This is possible because the tangent space $T_{\mathbf{x}^*}M$ can also be defined as the set of tangents at $\mathbf{x}^*$ of curves in M through $\mathbf{x}^*$.

Use the continuity of $Df(\mathbf{x})$ to pick $\varepsilon > 0$ so that $Df(\mathbf{x})\boldsymbol{v} > 0$ for $\mathbf{x} \in B_\varepsilon(\mathbf{x}^*) \subset U \cap M$. By the Mean Value Theorem applied to $\phi(t) = f\bigl(\mathbf{x}(t)\bigr)$ and the Chain Rule

$$f\bigl(\mathbf{x}(t)\bigr) = f(\mathbf{x}^*) + Df_{\mathbf{c}(t)}\bigl(\mathbf{x}'(t)\bigr)$$

for some $\mathbf{c}(t) \in \ell\bigl(\mathbf{x}(t), \mathbf{x}^*\bigr)$. For t small enough that $\mathbf{x}(t) \in B_\varepsilon(\mathbf{x}^*)$, we have

$$f\bigl(\mathbf{x}(t)\bigr) = f(\mathbf{x}^*) + Df_{\mathbf{c}(t)}\boldsymbol{v} > f(\mathbf{x}^*),$$

showing that $\mathbf{x}^*$ is not a maximum, **contradicting our hypothesis**. This establishes that $T_{\mathbf{x}^*}M \subset T_{\mathbf{x}^*}F$.

Now for any pair of sets in $\mathbb{R}^m$, $A \subset B$ implies $B^\perp \subset A^\perp$, where $A^\perp$ denotes the *orthogonal complement*, the set of vectors perpendicular to every vector in A.

Applying this to the tangent spaces, we find $(T_{\mathbf{x}^*}F)^\perp \subset (T_{\mathbf{x}^*}M)^\perp$. This implies $Df(\mathbf{x}^*) \in (T_{\mathbf{x}^*}M)^\perp$. Since Df is in the span of $D\mathbf{h}$, there are μ_i^*, $i = 1, \ldots, \ell$, with

$$Df(\mathbf{x}^*) = \sum_{i=1}^{\ell} \mu_i^* Dh_j(\mathbf{x}^*).$$

Since $\operatorname{rank} D\mathbf{h}(\mathbf{x}^*) = \ell$, the Dh_j are linearly independent, implying that the μ_i^* are unique. ∎

18.7 The Lagrangian

The first order conditions for an optimum are usually written using the *Lagrangian function*,

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\mu}) = f(\mathbf{x}) - \boldsymbol{\mu}^T\big(\mathbf{h}(\mathbf{x}) - \mathbf{c}\big) = f(\mathbf{x}) - \sum_{i=1}^{\ell} \mu_i\big(h_j(\mathbf{x}) - c_j\big).$$

This allows us to rewrite the key conclusions of the Tangent Space Theorem as follows:

Theorem 18.7.1. *Let $U \subset \mathbb{R}^m$ and $f\colon U \to \mathbb{R}$, $\mathbf{h}\colon \mathbb{R}^m \to \mathbb{R}^\ell$ be $\mathcal{C}^1$ functions. Suppose that $\mathbf{x}^*$ solves*

$$\begin{aligned} &\max_{\mathbf{x}}\ f(\mathbf{x}) \\ &\quad \textit{s.t. } h_j(\mathbf{x}) = c_j, \ \textit{for } i = 1, \dots, \ell \end{aligned}$$

or

$$\begin{aligned} &\min \mathbf{x}\ f(\mathbf{x}) \\ &\quad \textit{s.t. } h_j(\mathbf{x}) = c_j, \ \textit{for } i = 1, \dots, \ell \end{aligned}$$

and that $\operatorname{rank} D\mathbf{h}(\mathbf{x}^*) = \ell$ *holds, the non-degenerate constraint qualification (NDCQ). Then there are unique multipliers $\boldsymbol{\mu}^*$ such that $(\mathbf{x}^*, \boldsymbol{\mu}^*)$ is a critical point of the Lagrangian $\mathcal{L}$:*

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\mu}) = f(\mathbf{x}) - \boldsymbol{\mu}^T\big(\mathbf{h}(\mathbf{x}) - \mathbf{c}\big).$$

That is,

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x_i}(\mathbf{x}^*, \boldsymbol{\mu}^*) &= 0 \qquad \textit{for } i = 1, \dots, m \\ \frac{\partial \mathcal{L}}{\partial \mu_j}(\mathbf{x}^*, \boldsymbol{\mu}^*) &= 0 \qquad \textit{for } j = 1, \dots, \ell. \end{aligned}$$

Proof. If $Df(\mathbf{x}^*) \neq \mathbf{0}$, this follows immediately from the Tangent Space Theorem.

If $Df(\mathbf{x}^*) = \mathbf{0}$, the fact that $\operatorname{rank} D\mathbf{h}(\mathbf{x}^*) = \ell$ means that the ℓ rows of $D\mathbf{h}$ are linearly independent, so $\boldsymbol{\mu}^* = \mathbf{0}$ is the unique vector with $Df(\mathbf{x}^*) = (\boldsymbol{\mu}^*)^T D\mathbf{h}(\mathbf{x}^*)$. ∎

18.8 Solving a Simple Consumer's Problem, Part I

Consider the consumer's problem

$$v(\mathbf{p}, 100) = \max_{\mathbf{x}} u(\mathbf{x})$$
$$\text{s.t. } \mathbf{p}{\cdot}\mathbf{x} = 100$$

where $\mathbf{p} \in \mathbb{R}^3_{++}$ and

$$u(\mathbf{x}) = \begin{cases} x_1^{1/6} x_2^{1/2} x_3^{1/3} & \text{when } x_1, x_2, x_3 \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

In $m > 0$, any solution other than zero must have every $x_i > 0$. Theorem 29.20.1 then ensures that this problem has a solution.

In many economic problems, we will make assumptions that have an impact on optimization via the Lagrangian. Here, $Dh = \mathbf{p} \gg \mathbf{0}$. The NDCQ condition is satisfied.

Now we form the Lagrangian

$$\mathcal{L}(x_1, x_2, x_3, \mu_1) = x_1^{1/6} x_2^{1/2} x_3^{1/3} - \mu(\mathbf{p}{\cdot}\mathbf{x} - 100)$$

Critical points of the Lagrangian must obey

$$\frac{u(\mathbf{x})}{6x_1} = \mu p_1, \quad \frac{u(\mathbf{x})}{2x_2} = \mu p_2, \quad \frac{u(\mathbf{x})}{3x_3} = \mu p_3, \quad \mathbf{p}{\cdot}\mathbf{x} = 100$$

Another important feature of this problem is that income is positive. With positive price, that means that the budget line contains points that are strictly positive. For a Cobb-Douglas utility function, as we have here, this means that the maximum utility is positive and that $\mathbf{x}^* \gg \mathbf{0}$. As a result $Du(\mathbf{x}^*) \gg \mathbf{0}$. It follows from the first order equations above that $\mu^* > 0$.

18.9 Solving a Simple Consumer's Problem, Part II

We can rewrite the first three equations by dividing them in pairs, eliminating both μ and $u(\mathbf{x})$. Thus

$$\frac{x_2}{3x_1} = \frac{p_1}{p_2}, \quad \frac{x_3}{2x_1} = \frac{p_1}{p_3}, \quad \frac{3x_3}{2x_2} = \frac{p_2}{p_3}$$

The third equation is redundant, leaving us with

$$3p_1x_1 = p_2x_2 \text{ and } 2p_1x_1 = p_3x_3.$$

Substituting in our remaining equation, the budget constraint, we find

$$\mathbf{p}\cdot\mathbf{x} = p_1x_1 + 3p_1x_1 + 2p_1x_1 = 6p_1x_1 = 100.$$

Then

$$x_1 = \frac{100}{6p_1}, \quad x_2 = \frac{100}{2p_2}, \quad x_3 = \frac{100}{3p_3}$$

This implies the indirect utility function is

$$v(\mathbf{p}, 100) = u(\mathbf{x}^*) = \frac{100}{(6p_1)^{1/6}(2p_2)^{1/2}(3p_3)^{1/3}}$$

The multiplier μ^* is also easily calculated.

$$\mu^* = \frac{u(\mathbf{x}^*)}{100} = \frac{1}{(6p_1)^{1/6}(2p_2)^{1/2}(3p_3)^{1/3}}.$$

◀

How to Attack Such Problems. The basic steps used here were:

1. Rewrite the first order conditions as $\mathrm{MRS}_{ij} = p_i/p_j$ to eliminate the multiplier μ.
2. Express spending on each good in terms of spending on good one.
3. Substitute into the budget constraint so that everything is in terms of good one.
4. Solve for x_1, then substitute back to solve for the other x_j.

This often suffices to solve the problem, provided the equations involve are tractable.

18.10 Inequality Constraints: Non-Negative Multiplier

11/3/20

Although our simple consumer's problem in $\mathbb{R}^2$ involved only a single equality constraint, that is not typical. The consumer's problem in $\mathbb{R}^2$ usually involves three inequality constraints—two non-negativity constraints and the budget constraint. Other economics problems, such as the firm's cost minimization problem, or the consumer's expenditure minimization problem also use inequality constraints.

It will be helpful to distinguish cases where a particular constraint matters and where it does not. We say that a constraint $g(\mathbf{x}) \leq b$ *binds at* $\mathbf{x}^*$ if $g(\mathbf{x}^*) = b$. Otherwise the constraint is *non-binding*.

18.11 A Single Inequality Constraint

Let's start by investigating the case of a single inequality constraint. We will write the maximization problem in the following form:

$$\max_{x} u(x)$$
$$\text{s.t. } g(x) \leq b$$

Figure 18.11.1 illustrates two possibilities that we need to consider. It shows that the sign of the multiplier matters when we have an inequality constraint. Both Du and Dg must point in the same direction, otherwise we find ourselves minimizing utility over the constraint set as in the right panel of Figure 18.11.1.

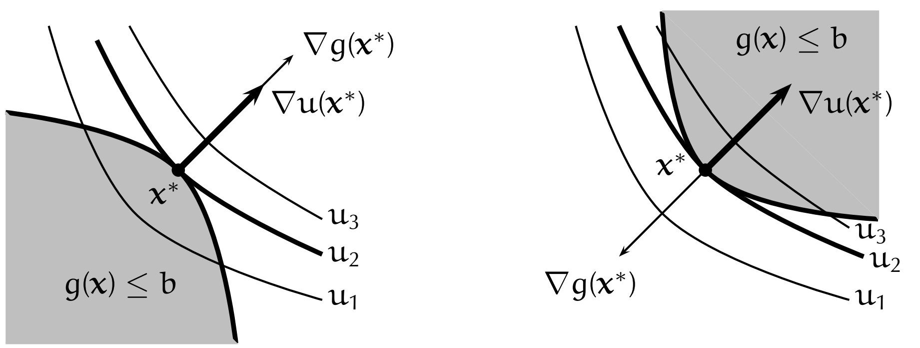

Figure 18.11.1: Here x^* is the consumer's utility maximum with $u(x^*) = u_2$. At the maximum, the vectors $\nabla g(x^*)$ and $\nabla u(x^*)$ are collinear, with the heavier arrow being $\nabla u(x^*)$. With an inequality constraint, like we have here, the two vectors must point in the same direction. Then $\nabla u(x^*) = \lambda \nabla g(x^*)$ for some $\lambda \geq 0$.

If $\nabla g(x^*)$ pointed in the opposite direction, it would mean that the region above u_2 would have lower values of g, as shown in the right panel. In that case utility is not maximized at x^*, since higher indifference curves such as u_3 can be attained.

Another way to think about it is that when $\nabla u(x^*) = \lambda \nabla g(x^*)$, a small move in the $\nabla u(x^*)$ direction will reduce g, and move into the interior of the constraint set while increasing utility. As in the proof of the Tangent Space Theorem, the Mean Value Theorem can be used to show this.

18.12 Inequality Constraints: Complementary Slackness

The Tangent Space Theorem can be easily modified to ensure that the multiplier is non-negative. However, there is another issue that might arise. The point $\mathbf{x}^*$ could be in the interior of the constraint set, where $g(\mathbf{x}^*) < b$. The Tangent Space Theorem does not apply there. However, $\mathbf{x}^*$ is then an interior point, so $Du(\mathbf{x}^*) = \mathbf{0}$.

This can be interpreted as the multiplier being zero, just as we did in the equality constraint case when $Df(\mathbf{x}^*) = \mathbf{0}$.

There is a useful condition that lets up package this up using the Lagrangian framework. We impose the *complementary slackness condition* that $\lambda\big(g(\mathbf{x}) - b\big) = 0$. If the constraint $g(\mathbf{x}) \le b$ binds, the complementary slackness condition tells us nothing. It's already zero. If the constraint doesn't bind, so $g(\mathbf{x}) - b < 0$, it forces the corresponding multiplier to be zero. This trick also works if we have several inequality constraints.

18.13 Maximization with Complementary Slackness

We sum up our discussion of complementary slackness in the following theorem.

Theorem 18.13.1. *Let* $U \subset \mathbb{R}^m$ *and suppose* $f\colon U \to \mathbb{R}$ *and* $\mathbf{g}\colon U \to \mathbb{R}^k$ *are* $\mathcal{C}^1$ *functions. Suppose that* $\mathbf{x}^*$ *solves*

$$\begin{aligned} &\max_{\mathbf{x}} f(\mathbf{x}) \\ &\textit{s.t.}\;\; \hat{g}_i(\mathbf{x}) \le b_i, \;\textit{for } i \textit{ where } \hat{g}_i \textit{ is defined}. \end{aligned}$$

and that $\hat{k}$ *constraints bind at* $\mathbf{x}^*$. *Let* $\hat{\mathbf{g}}$ *be the vector of binding inequality constraints at* $\mathbf{x}^*$ *and* $\hat{\mathbf{b}}$ *the corresponding constants. Form the Lagrangian* $\mathcal{L}$*:*

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) - \boldsymbol{\lambda}^{\mathsf{T}}\big(\hat{\mathbf{g}}(\mathbf{x}) - \hat{\mathbf{b}}\big).$$

If rank $D\hat{\mathbf{g}}(\mathbf{x}^*) = \hat{k} < m$ *holds (NDCQ), then there are multipliers* $\boldsymbol{\lambda}^*$ *such that*

(a) *The pair* $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ *is a critical point of the Lagrangian:*

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x_i}(\mathbf{x}^*, \boldsymbol{\lambda}^*) &= 0 \qquad \textit{for } i = 1, \dots, m \\ \frac{\partial \mathcal{L}}{\partial \lambda_j}(\mathbf{x}^*, \boldsymbol{\lambda}^*) h_j(\mathbf{x}^*) - c_j &= 0 \qquad \textit{for } j = 1, \dots, k, \end{aligned}$$

(b) *The complementary slackness conditions hold:*

$$\lambda_i^*\big[g_i(\mathbf{x}^*) - b_i\big] = 0 \qquad \textit{for } i = 1, \dots, k$$

(c) *The multipliers are non-negative:*

$$\lambda_1^* \ge 0, \dots, \lambda_k^* \ge 0,$$

(d) *The constraints are satisfied:* $\mathbf{g}(\mathbf{x}^*) \le \mathbf{b}$.

18.14 Failure of Constraint Qualification I

Now that we have a new tool, inequality constraints, you might be tempted to view an equality constraint as two inequality constraints. For example, you can write

$$p_1x_1 + p_2x_2 = m$$

as

$$\begin{aligned} p_1x_1 + p_2x_2 &\le m \\ -p_1x_1 - p_2x_2 &\le -m. \end{aligned}$$

This doesn't work. It runs afoul of the NDCQ. When both bind, as they must if both are obeyed, we have

$$D\hat{g} = \begin{pmatrix} p_1 & p_2 \\ -p_1 & -p_2 \end{pmatrix}.$$

This has rank 1 when it needs rank 2. In this case, the failure of constraint qualification is minor. If you do this with Cobb-Douglas utility, you won't be able to uniquely determine the two multipliers. However, you can determine their difference.

18.15 Solving a Cobb-Douglas Consumer's Problem I

The key to understanding how to use Theorem 18.13.1 is that complementary slackness conditions are a way of checking all the possible cases where some, but not all, constraints bind. Let's look at a concrete problem in $\mathbb{R}^2$ to see how this works.

$$\begin{aligned} &\max_{\mathbf{x}} x_1^{1/3} x_2^{2/3} \\ &\text{s.t. } p_1 x_1 + p_2 x_2 \le m \\ &\quad x_1 \ge 0, x_2 \ge 0. \end{aligned}$$

Where $p_1, p_2, m > 0$.

We form the Lagrangian by first rewriting the non-negativity constraints in the proper form: $-x_1 \le 0$, $-x_2 \le 0$. The Lagrangian is

$$\begin{aligned} \mathcal{L} &= u(\mathbf{x}) - \lambda_0(p_1 x_1 + p_2 x_2 - m) - \lambda_1(-x_1) - \lambda_2(-x_2) \\ &= u(\mathbf{x}) - \lambda_0(p_1 x_1 + p_2 x_2 - m) + \lambda_1 x_1 + \lambda_2 x_2. \end{aligned}$$

Non-negativity constraints always yield terms of the form $+\lambda_i x_i$.

18.16 Solving a Cobb-Douglas Consumer's Problem II

We now turn to constraint qualification. Consider

$$D\mathbf{g} = \begin{pmatrix} p_1 & p_2 \\ -1 & 0 \\ 0 & -1 \end{pmatrix}$$

Now if only constraint i binds, we must have $Dg_i(\mathbf{x}^*) \neq \mathbf{0}$, and we do. If two constraints bind, the matrix $D\hat{\mathbf{g}}(\mathbf{x}^*)$ obtained by deleting the non-binding row must be invertible, which it is. It is not possible for all three constraints to bind at once as we would have $p_1(0) + p_2(0) = m > 0$. The NDCQ condition is satisfied.

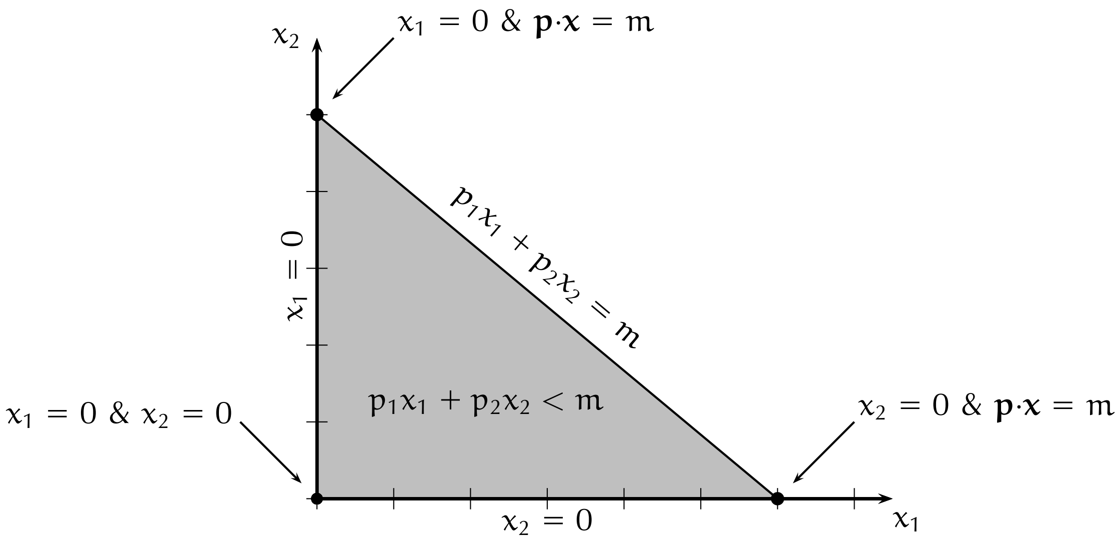

Figure 18.16.1: Without complementary slackness, we would have to check separately for maxima eight ways: at each of the three corner points, in the relative interior of each of the three boundary segments of the budget set, and in the interior of the budget set.

18.17 Solving a Cobb-Douglas Consumer's Problem III

Now differentiate $\mathcal{L}$ with respect to x_1 and x_2.

$$
\begin{aligned}
0 &= \frac{\partial \mathcal{L}}{\partial x_1} = \frac{1}{3}\left(\frac{x_2}{x_1}\right)^{2/3} - \lambda_0 p_1 + \lambda_1 \\
0 &= \frac{\partial \mathcal{L}}{\partial x_2} = \frac{2}{3}\left(\frac{x_1}{x_2}\right)^{1/3} - \lambda_0 p_2 + \lambda_2.
\end{aligned}
$$

With $m > 0$, we know that positive utility is possible, so $x_1 > 0$ and $x_2 > 0$. The complementary slackness conditions are $\lambda_1 x_1 = 0$ and $\lambda_2 x_2 = 0$ which imply $\lambda_1^* = \lambda_2^* = 0$.

We can now rewrite the first order conditions:

$$
\begin{aligned}
\lambda_0 p_1 &= \frac{1}{3}\left(\frac{x_2}{x_1}\right)^{2/3} \\
\lambda_0 p_2 &= \frac{2}{3}\left(\frac{x_1}{x_2}\right)^{1/3}
\end{aligned}
\tag{18.17.1}
$$

Because the right-hand side must be positive at $\mathbf{x}^*$, $\lambda_0 > 0$. Complementary slackness now yields $p_1 x_1 + p_2 x_2 = m$. The budget constraint must bind. At this point, we have reduced the possible locations of solutions to one region, the relative interior of the budget frontier.

Next we eliminate λ_0 from equation (18.17.1) by dividing the top line by the bottom line.

$$\frac{p_1}{p_2} = \frac{1}{2}\frac{x_2}{x_1}$$

implying $p_1 x_1 = \frac{1}{2} p_2 x_2$. We substitute in the budget constraint to find $3p_1 x_1 = m$, so

$$x_1^* = \frac{m}{3p_1} \quad \text{and} \quad x_2^* = \frac{2m}{3p_2}.$$

It follows that

$$\lambda_0^* = \frac{2^{2/3}}{3p_1^{1/3} p_2^{2/3}}.$$

18.18 Complementary Slackness Gone Wild: I

We consider the problem of maximizing a linear utility function in $\mathbb{R}^3_+$. One thing that's different about this problem is that it is purely an exercise in the use of complementary slackness.

$$\begin{aligned} \max_{\mathbf{x}} u(\mathbf{x}) &= a_1 x_1 + a_2 x_2 + a_3 x_3 \\ \text{s.t. } & p_1 x_1 + p_2 x_2 + p_3 x_3 \le m \\ & x_1 \ge 0, x_2 \ge 0, x_3 \ge 0. \end{aligned}$$

Where $\mathbf{p} \gg \mathbf{0}$, $m > 0$, and each $a_i > 0$. Theorem 29.20.1 guarantees that the problem has a solution.

We next check constraint qualification. Here

$$Dg = \begin{pmatrix} p_1 & p_2 & p_3 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

It is impossible for all four constraints to bind. If they did, we would have $x_1 = x_2 = x_3 = 0$ and $0 = \mathbf{p}\cdot\mathbf{x} = m > 0$. Any collection of 1, 2, or 3 rows of Dg is linearly independent, so constraint qualification (NDCQ) is satisfied.

18.19 Complementary Slackness Gone Wild: II

The Lagrangian is

$$\mathcal{L} = a_1x_1 + a_2x_2 + a_3x_3 - \lambda_0(\mathbf{p}\cdot\mathbf{x} - m) + \lambda_1x_1 + \lambda_2x_2 + \lambda_3x_3.$$

The first order conditions are

$$\begin{aligned} 0 &= \frac{\partial\mathcal{L}}{\partial x_1} = a_1 - \lambda_0p_1 + \lambda_1 \\ 0 &= \frac{\partial\mathcal{L}}{\partial x_2} = a_2 - \lambda_0p_2 + \lambda_2 \\ 0 &= \frac{\partial\mathcal{L}}{\partial x_3} = a_3 - \lambda_0p_3 + \lambda_3 \end{aligned}$$

which we rewrite as

$$\begin{aligned} \lambda_0p_1 &= a_1 + \lambda_1 \\ \lambda_0p_2 &= a_2 + \lambda_2 \\ \lambda_0p_3 &= a_3 + \lambda_3. \end{aligned}$$

At this point, you may be wondering how to proceed with such first order conditions. What can we do? The solution is complementary slackness.

18.20 Complementary Slackness Gone Wild: III

To recapitulate, we have the following first order conditions:

$$\begin{aligned}\lambda_0 p_1 &= a_1 + \lambda_1\\ \lambda_0 p_2 &= a_2 + \lambda_2\\ \lambda_0 p_3 &= a_3 + \lambda_3.\end{aligned}\tag{18.20.2}$$

Now $\lambda_1 \geq 0$, so $\lambda_0 p_0 \geq a_1 > 0$. This implies $\lambda_0 > 0$. Complementary slackness tells us that $\lambda_0(\mathbf{p}{\cdot}\boldsymbol{x} - m) = 0$. Since $\lambda_0 > 0$, that means $\mathbf{p}{\cdot}\boldsymbol{x} = m > 0$. The consumer spends their entire budget.

Dividing each line of equation (18.20.2) by p_i tells us

$$\begin{aligned}\lambda_0 &= a_1/p_1 + \lambda_1/p_1 \geq a_1/p_1\\ \lambda_0 &= a_2/p_2 + \lambda_2/p_2 \geq a_2/p_2\\ \lambda_0 &= a_3/p_3 + \lambda_3/p_3 \geq a_3/p_3.\end{aligned}\tag{18.20.3}$$

Then

$$\lambda_0 \geq \max_{i=1,2,3}\{a_i/p_i\}.$$

If $\lambda_0 > \max_{i=1,2,3}\{a_i/p_i\}$, then each $\lambda_i > 0$ for $i = 1,2,3$. Complementary slackness requires that $\lambda_i x_i = 0$ for $i = 1,2,3$, implying that $x_1 = x_2 = x_3 = 0$. But this is impossible because we know $\mathbf{p}{\cdot}\boldsymbol{x} = m > 0$. At least one x_i must be positive, and for that i, $\lambda_0 = a_i/p_i$. So we can write

$$\lambda_0 = \max_{i=1,2,3}\left\{\frac{a_i}{p_i}\right\}.$$

18.21 Complementary Slackness Gone Wild: IV

We are now prepared to find the solution.

If $a_i/p_i < \lambda_0$, $\lambda_i > 0$ and so $x_i = 0$. The only goods that are consumed are those for which

$$\frac{a_h}{p_h} = \max_{i=1,2,3}\left\{\frac{a_i}{p_i}\right\}.$$

There is at least one such good. There may be 2 or 3, depending on the a_i's and p_i's. If goods one and two are both consumed and good three is not consumed, then

$$u(\mathbf{x}) = a_1x_1 + a_2x_2 = \lambda_0p_1x_1 + \lambda_0p_2x_2 = \lambda_0m.$$

Utility doesn't depend on how consumption is distributed between the goods that are consumed.

That principle applies regardless of which goods are consumed. Therefore the solution to the problem is that our consumer consumes only goods h obeying $a_h/p_h = \max_i\{a_i/p_i\}$, and the any distribution of consumption over those goods that spends all of the available income m is optimal.

18.22 Example: Quasi-linear Utility I

Now consider maximizing the quasi-linear utility function $u(x, y) = x + \sqrt{y}$. We must solve:

$$\begin{aligned} \max_{\mathbf{x}} u(\mathbf{x}) &= x + \sqrt{y} \\ \text{s.t. } p_x x + p_y y &\leq m \\ x \geq 0, y &\geq 0. \end{aligned}$$

As in the last two problems, and most standard consumer's problems, the NDCQ condition is satisfied. The Lagrangian is

$$\mathcal{L} = x + \sqrt{y} - \lambda_0(p_x x + p_y y - m) + \lambda_x x + \lambda_y y$$

yielding first order conditions

$$\begin{aligned} 0 &= \frac{\partial \mathcal{L}}{\partial x} = 1 - \lambda_0 p_x + \lambda_x \\ 0 &= \frac{\partial \mathcal{L}}{\partial y} = -\frac{1}{2\sqrt{y}} + \lambda_y. \end{aligned} \tag{18.22.4}$$

18.23 Example: Quasi-linear Utility II

We rewrite equation (18.22.4) as

$$\begin{aligned}\lambda_0 p_x &= 1 + \lambda_x \\ \lambda_0 p_y &= \frac{1}{2\sqrt{y}} + \lambda_y.\end{aligned} \tag{18.23.5}$$

The top line of equation (18.23.5) implies $\lambda_0 \geq 1/p_x > 0$. By complementary slackness, the budget constraint must bind: $p_x x + p_y y = m$. This is also the usual result in standard consumer's problems.

There are now three cases to consider, depending on which constraints bind. We organize these based on the variables. This is usually better than trying to organize based on the multipliers because when a multiplier is zero, that constraint may still bind. The values of the multipliers do not translate nicely into information about which constraints bind.

The cases are (I) $x > 0$ and $y = 0$, (II) $x = 0$ and $y > 0$, and (III) $x > 0$ and $y > 0$. Notice how we have reduced the original eight cases to three.

18.24 Example: Quasi-linear Utility III

Our first order conditions are

$$\begin{aligned}\lambda_0 p_x &= 1 + \lambda_x \\ \lambda_0 p_y &= \frac{1}{2\sqrt{y}} + \lambda_y.\end{aligned} \tag{18.23.5}$$

Case I: Here $x > 0$ and $y = 0$. The complementary slackness condition $\lambda_x x = 0$ tells us that $\lambda_x = 0$. By the top line of equation (18.23.5), $\lambda_0 = 1/p_x$. The second equation of equation (18.23.5) becomes

$$p_y/p_x = \frac{1}{2\sqrt{y}} + \lambda_y.$$

But this is impossible to satisfy since $y = 0$. Case I is out.

Case II: Here $x = 0$ and $y > 0$. The complementary slackness condition $\lambda_y y = 0$ implies $\lambda_y = 0$, so the bottom line of equation (18.23.5) becomes

$$\lambda_0 p_y = \frac{1}{2\sqrt{y}}$$

The budget constraint tells us $y = m/p_y$. It follows that $\lambda_0 = 1/2\sqrt{p_y m}$. By the top line of equation (18.23.5),

$$\frac{p_x}{2\sqrt{p_y m}} = 1 + \lambda_x \geq 1.$$

This solution works if $p_x^2 \geq 4 p_y m$.

18.25 Example: Quasi-linear Utility IV

Case III: Here $x, y > 0$. By complementary slackness $\lambda_x = \lambda_y = 0$, so the first order conditions (18.23.5) are

$$\begin{aligned} \lambda_0 p_x &= 1 \\ \lambda_0 p_y &= \frac{1}{2\sqrt{y}}. \end{aligned}$$

Then $\lambda_0 = 1/p_x$ and $p_y/p_x = 1/2\sqrt{y}$. It follows that $y = p_x^2/4p_y^2$. Using the budget constraint, we find $p_x x = m - p_y y = m - p_x^2/4p_y$. The non-negativity constraint for x requires $m - p_x^2/4p_y \geq 0$ or equivalently, $4p_y m \geq p_x^2$.

Collecting the results together, we find

$$(x, y) = \begin{cases} \left(0, \frac{m}{p_y}\right) & \text{when } p_x^2 \geq 4p_y m \\ \left(\frac{m}{p_x} - \frac{p_x}{4p_y^2}, \frac{p_x^2}{4p_y}\right) & \text{when } p_x^2 \leq 4p_y m. \end{cases}$$

When $p_x^2 = 4p_y m$, the two cases agree.

18.26 Failure of Constraint Qualification II

In some cases, the failure of constraint qualification is quite serious, serious enough that it prevents us from solving the problem as advertised. Consider this example borrowed from section 19.5 of Simon and Blume.

$$\max_{(x,y)} x$$
$$\text{s.t. } x^3 + y^2 = 0.$$

Suppose we blindly charge ahead, setting up the Lagrangian

$$\mathcal{L} = x - \mu(x^3 + y^2).$$

The critical points of the Lagrangian solve

$$1 = 3\mu x^2, \quad 0 = 2\mu y, \quad \text{and} \quad x^3 + y^2 = 0.$$

There are no solutions. If $y = 0$, $x = 0$ by the constraint, so $1 = 3\mu x^2 = 0$. If $\mu = 0$, we again have $1 = 0$. But $0 = 2\mu y$ so one of μ or y must be zero.

So what went wrong. Constraint qualification fails because $Dh = (3x^2, 2y)$ has a critical point, $(x, y) = (0, 0)$. This is a **problem** because the maximum is also at $(0, 0)$.

How do we know? The constraint says $x^3 = -y^2$, so $-x^3 \le 0$, implying $x \le 0$. Of course, we can maximize $x \le 0$ by setting $x = 0$, which is feasible. All we have to do is set $y = 0$, landing us on the critical point of the constraint.

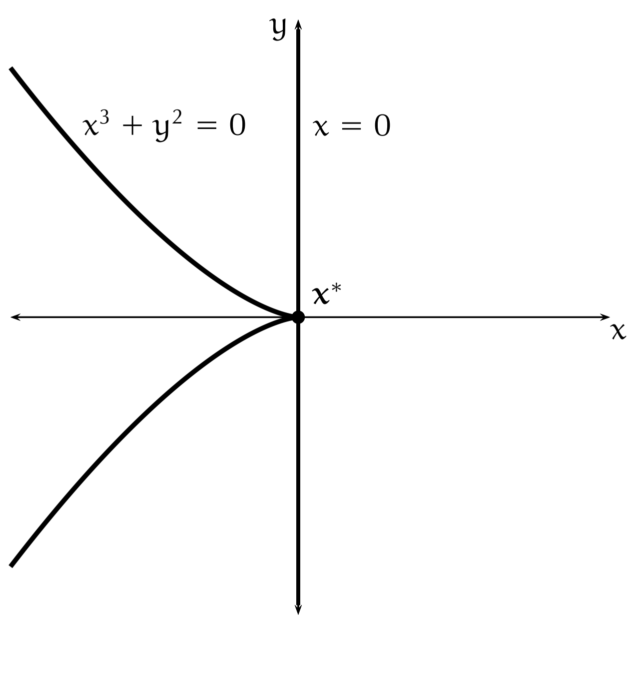

Figure 18.26.1: The graph of the constraint is illustrated in the diagram. It's pretty obvious from the graph that the maximum value of x on the constraint is 0 and occurs at $\mathbf{x}^* = (0, 0)$.

In Chapter 19, we will see that this type of problem can be solved using a modified Lagrangian.

18.27 Mixed Constraints

Finally, we state a combined theorem incorporating both equality and inequality constraints.

Theorem 18.27.1. *Let* $U \subset \mathbb{R}^m$ *and suppose* $f\colon U \to \mathbb{R}$, $\mathbf{g}: U \to \mathbb{R}^k$, *and* $\mathbf{h}\colon \mathbb{R}^m \to \mathbb{R}^\ell$ *are* $\mathcal{C}^1$ *functions. Suppose that* $\mathbf{x}^*$ *solves*

$$\begin{aligned} &\max_{\mathbf{x}}\ f(\mathbf{x}) \\ &\textit{s.t.}\ \ g_i(\mathbf{x}) \le b_i, \ \textit{for } i = 1, \dots, k \\ &\qquad h_j(\mathbf{x}) = c_j, \ \textit{for } i = 1, \dots, \ell \end{aligned}$$

and that $\hat{k}$ *inequality constraints bind at* $\mathbf{x}^*$. *Let* $\hat{\mathbf{g}}$ *be the vector of binding inequality constraints at* $\mathbf{x}^*$ *and* $\hat{\mathbf{b}}$ *the corresponding constants. Form the Lagrangian:*

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\mathbf{x}) - \boldsymbol{\lambda}^{\mathbf{T}}\big(\hat{\mathbf{g}}(\mathbf{x}) - \mathbf{b}\big) - \boldsymbol{\mu}^{\mathbf{T}}\big(\mathbf{h}(\mathbf{x}) - \mathbf{c}\big).$$

Set

$$\mathbf{G}(\mathbf{x}) = \begin{pmatrix} \hat{\mathbf{g}}(\mathbf{x}) \\ \mathbf{h}(\mathbf{x}) \end{pmatrix}$$

If $\operatorname{rank} D\mathbf{G}(\mathbf{x}^*) = \hat{k} + \ell < m$ *holds (NDCQ), then there are multipliers* $\boldsymbol{\lambda}^*$ *and* $\boldsymbol{\mu}^*$ *such that*

(a) *The triplet* $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ *is a critical point of the Lagrangian:*

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x_i}(\mathbf{x}^*, \boldsymbol{\lambda}^*) &= 0 && \textit{for } i = 1, \dots, m \\ \frac{\partial \mathcal{L}}{\partial \lambda_i}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) = \hat{\mathbf{g}}(\mathbf{x}) - \hat{\mathbf{b}} &= 0 && \textit{for } i = 1, \dots, k \textit{ and binding}, \\ \frac{\partial \mathcal{L}}{\partial \mu_j}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) = h_j(\mathbf{x}) - c_j &= 0 && \textit{for } j = 1, \dots, \ell, \end{aligned}$$

(b) *The complementary slackness conditions hold:*

$$\lambda_i^*\big[g_i(\mathbf{x}^*) - b_i\big] = 0 \qquad \textit{for all } i = 1, \dots, k$$

(c) *The multipliers are non-negative:*

$$\lambda_1^* \ge 0, \dots, \lambda_k^* \ge 0,$$

(d) *The constraints are satisfied:* $\mathbf{g}(\mathbf{x}^*) \le \mathbf{b}$ *and* $\mathbf{h}(\mathbf{x}^*) = \mathbf{c}$.

Notice that the constraint on $\mathbf{h}$ appears twice, in both parts (a) and (d).

18.28 Example with Mixed Constraints I

Consider the problem

$$\begin{aligned} \max_{x}\ & f(x,y) = x^2 - y^2 \\ \text{s.t. } & x \geq 0, y \geq 0 \\ & x^2 + y^2 = 4. \end{aligned}$$

The constraint set is compact, and the function is continuous, so it will have both maxima and minima by the Weierstrass Theorem.

The matrix $D\hat{\mathbf{g}}$ will be a submatrix of

$$\begin{pmatrix} 2x & 2y \\ -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Notice that any 2×2 submatrix has rank 2 and all rows are non-negative because $x^2 + y^2 = 4$. The NDCQ condition is satisfied.

Form the Lagrangian

$$\mathcal{L} = x^2 - y^2 - \mu(x^2 + y^2 - 4) + \lambda_x x + \lambda_y y.$$

The first order conditions are

$$\begin{aligned} 0 &= \frac{\partial \mathcal{L}}{\partial x} = 2x - 2x\mu + \lambda_x \\ 0 &= \frac{\partial \mathcal{L}}{\partial y} = -2y - 2y\mu + \lambda_y \\ 0 &= \frac{\partial \mathcal{L}}{\partial \mu} = -x^2 - y^2 + 4 \end{aligned}$$

The last one is the equality constraint. The solution must also obey the two complementary slackness conditions

$$\lambda_x x = 0, \quad \lambda_y y = 0$$

and the non-negativity constraints

$$x \geq 0, \quad y \geq 0, \quad \lambda \geq 0.$$

18.29 Example with Mixed Constraints II

We rewrite the first order conditions

$$\begin{aligned} 2x + \lambda_x &= 2\mu x \\ 2y + 2\mu y &= \lambda_y \end{aligned}$$

Multiply the first equation by x and the second by y, obtaining

$$\begin{aligned} 2x^2 + \lambda_x x &= 2\mu x^2 \\ 2y^2 + 2\mu y^2 &= \lambda_y y \end{aligned}$$

Apply complementary slackness to both: $\lambda_x x = 0$ and $\lambda_y y = 0$. Then

$$\begin{aligned} 2x^2 &= 2\mu x^2 \\ 2y^2 + 2\mu y^2 &= 0. \end{aligned}$$

From the first equation, if $x > 0$ then $\mu = 1$, while according to the second equation, if $y > 0$, $\mu = -1$. At least one of x and y must be zero. They can't both be zero due to the constraint $x^2 + y^2 = 4$.

That gives us two cases to consider: (1) $x > 0$, $y = 0$, and (2) $y > 0$, $x = 0$.

Case I: $x > 0$ and $y = 0$. Here $\mu = +1$, $\lambda_x = 0$ and $\lambda_y = 0$. Since $x^2 + y^2 = 4$ and $x \geq 0$, this solution is $(2, 0)$ with $f(4, 0) = 4$.

Case II: $x = 0$ and $y > 0$. Here $\mu = -1$, $\lambda_x = 0$ and $\lambda_y = 0$. Since $x^2 + y^2 = 4$ and $y \geq 0$, this solution is $(0, 2)$ with $(0, 2) = -4$.

Based on these results, $(2, 0)$ is the maximum and $(0, 2)$ is the minimum.

18.30 Minimization with Inequality Constraints

When we only had equality constraints, the same conditions found critical points for both maxima and minima. That is no longer true when there are inequality constraints as the sign of the associated multipliers depends on whether we are maximizing or minimizing.

There are various ways to handle minimization problems with inequality constraints. One of the more obvious is to maximize $-f$ instead of minimizing f. That yields Lagrangian

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = -f(\mathbf{x}) - \boldsymbol{\lambda}^T\big(\mathbf{g}(\mathbf{x}) - \mathbf{b}\big) - \boldsymbol{\mu}^T\big(\mathbf{h}(\mathbf{x}) - \mathbf{c}\big).$$

Multiplying by -1, we obtain

$$f(\mathbf{x}) + \boldsymbol{\lambda}^T\big(\mathbf{g}(\mathbf{x}) - \mathbf{b}\big) + \boldsymbol{\mu}^T\big(\mathbf{h}(\mathbf{x}) - \mathbf{c}\big).$$

Since the sign of $\boldsymbol{\mu}$ doesn't matter, this is equivalent to using the Lagrangian

$$\mathcal{L}_1(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\mathbf{x}) + \boldsymbol{\lambda}^T\big(\mathbf{g}(\mathbf{x}) - \mathbf{b}\big) - \boldsymbol{\mu}^T\big(\mathbf{h}(\mathbf{x}) - \mathbf{c}\big).$$

where $\boldsymbol{\lambda} \geq \mathbf{0}$ or using

$$\mathcal{L}_2(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\mathbf{x}) - \boldsymbol{\lambda}^T\big(\mathbf{g}(\mathbf{x}) - \mathbf{b}\big) - \boldsymbol{\mu}^T\big(\mathbf{h}(\mathbf{x}) - \mathbf{c}\big).$$

with $\boldsymbol{\lambda} \leq \mathbf{0}$. Yet another method is the one favored in the book, writing the inequality constraints the opposite way, setting $\mathbf{g}' = -\mathbf{g}$ and $\mathbf{b}' = -\mathbf{b}$, yielding constraints $\mathbf{g}'(\mathbf{x}) \geq \mathbf{b}'$ instead of $\mathbf{g}(\mathbf{x}) \leq \mathbf{b}$. We will state the result in this form, which is more natural when thinking about duality.

It doesn't really matter which form of Lagrangian you use. Just be sure not to mix them!

18.31 Minimization with Mixed Constraints

We state a version Theorem 18.27.1 that works for minimization problems.

Theorem 18.31.1. *Let* $U \subset \mathbb{R}^m$ *and suppose* $f\colon U \to \mathbb{R}$, $\mathbf{g}\colon U \to \mathbb{R}^k$, *and* $\mathbf{h}\colon \mathbb{R}^m \to \mathbb{R}^\ell$ *are* $\mathcal{C}^1$ *functions. Suppose that* $\mathbf{x}^*$ *solves*

$$\begin{aligned} &\min_{\mathbf{x}} f(\mathbf{x}) \\ &\text{s.t. } g_i(\mathbf{x}) \geq b_i, \text{ for } i = 1, \dots, k \\ &\qquad h_j(\mathbf{x}) = c_j, \text{ for } i = 1, \dots, \ell \end{aligned}$$

and that $\hat{k}$ *constraints bind at* $\mathbf{x}^*$. *Let* $\hat{\mathbf{g}}$ *be the vector of binding inequality constraints at* $\mathbf{x}^*$ *and* $\hat{\mathbf{b}}$ *the corresponding constants. Form the Lagrangian* $\mathcal{L}$:

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\mathbf{x}) - \boldsymbol{\lambda}^T\big(\mathbf{g}(\mathbf{x}) - \hat{\mathbf{b}}\big) - \boldsymbol{\mu}^T\big(\mathbf{h}(\mathbf{x}) - \mathbf{c}\big).$$

If

$$\operatorname{rank} D\begin{pmatrix} \hat{\mathbf{g}}(\mathbf{x}^*) \\ \mathbf{h}(\mathbf{x}^*) \end{pmatrix} = \hat{k} + \ell < m$$

holds (NDCQ), then there are multipliers $\boldsymbol{\lambda}^*$ *and* $\boldsymbol{\mu}^*$ *such that*

(a) *The triplet* $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ *is a critical point of the Lagrangian:*

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x_i}(\mathbf{x}^*, \boldsymbol{\lambda}^*) &= 0 && \text{for } i = 1, \dots, m \\ \frac{\partial \mathcal{L}}{\partial \lambda_i}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) = g_i(\mathbf{x}) - b_i &= 0 && \text{for binding } i = 1, \dots, k, \\ \frac{\partial \mathcal{L}}{\partial \mu_j}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) = h_j(\mathbf{x}^*) - c_j &= 0 && \text{for } j = 1, \dots, \ell, \end{aligned}$$

(b) *The complementary slackness conditions hold:*

$$\lambda_i^*\big[g_i(\mathbf{x}^*) - b_i\big] = 0 \qquad \text{for all } i = 1, \dots, k$$

(c) *The multipliers are non-negative:*

$$\lambda_1^* \geq 0, \dots, \lambda_k^* \geq 0,$$

(d) *The constraints are satisfied:* $\mathbf{g}(\mathbf{x}^*) \geq \mathbf{b}$ *and* $\mathbf{h}(\mathbf{x}^*) = \mathbf{c}$.

November 4, 2020