Chapter 11

Unconstrained Optimization

The problem is to find the point to maximize or minimize an objective function.

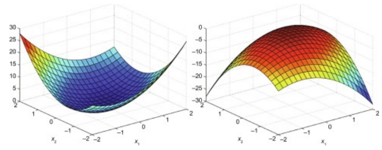


Fig. 11.1. Find the min (left) or max (right)

11.1 Two Choice Variables: Necessary Conditions

Theorem 11.1 (First-order necessary conditions). *A differentiable function $z = f(x, y)$ can have a max or min at an interior point (x_0, y_0) of its domain only if it is a critical point – that is, if the point (x_0, y_0) satisfies the two equations*

$$f'_x(x, y) = 0, \qquad f'_y(x, y) = 0.$$

These are known as first-order conditions (FOCs).

Example 11.1 (Profit-Maximization). Suppose the production function F is twice differentiable

$$Q = F(K, L),$$

where K, L are capital and labor. The input price is p, the capital rental cost is r, and the wage rate is w such that $p, r, w > 0$. The firm chooses the optimal input factors (K, L) to maximize the following profit

$$\max_{K,L} \pi = pF(K, L) - rK - wL$$

The FOCs are:

$$\pi'_K = pF'_K - r = 0,$$
$$\pi'_L = pF'_L - w = 0.$$

which implies the optimal choices of K^*, L^* as

$$F'_K(K^*, L^*) = r/p,$$
$$F'_L(K^*, L^*) = w/p.$$

This is a system of 2 equations of 2 unknowns, so there exists a unique solution for K^*, L^*.

Ex. 11.1 (Profit Maximizing Firm)**.** Solve the profit maximization problem for a firm with the following production function

$$Q = F(K, L) = 12K^{1/2}L^{1/4}$$

with $p = 1, w = 0.6, r = 1.2$.

[Answers: $K = L = 625$.]

Ex. 11.2 (Discriminating Monopolist)**.** A monopolist producing a SINGLE output has 2 types of customers. It produces Q_1 for customers of type 1, who are willing to pay with the price of $p_1 = 50 - 5Q_1$. The monopolist produces Q_2 for customers of type 2 under the price $p_2 = 100 - 10Q_2$. The cost of manufacturing is $90 + 20Q$. How should the monopolist produce for each market?

11.2 Two Choice Variables: Sufficient Conditions

The sufficient condition ensures that the optimal values found are indeed the maximizers or minimizers. In 1-variable calculus, you need to test the second-order derivatives. If $f''(x) \leq 0$, then the function is concave and your critical point is a MAX. For a 2-variable function, we need to rely on the Hessian matrix to test its concavity. The premise is still the same, the solution is indeed a MAX if the function is concave down, and MIN if the function is concave up.

Theorem 11.1. *Let f be a C^2 function defined in a convex set S in $\mathbb{R}^2$. Suppose (x^*, y^*) is an interior critical point, then the Hessian matrix of f at this critical point is*

$$D^2F(x^*, y^*) = \begin{pmatrix} \dfrac{\partial^2 F}{\partial x^2}(x^*, y^*) & \dfrac{\partial^2 F}{\partial x \partial y}(x^*, y^*) \\ \dfrac{\partial^2 F}{\partial y \partial x}(x^*, y^*) & \dfrac{\partial^2 F}{\partial y^2}(x^*, y^*) \end{pmatrix}$$

shorthand

$$D^2F(x^*, y^*) = \begin{pmatrix} f''_{xx} & f''_{yx} \\ f''_{xy} & f''_{yy} \end{pmatrix}$$

- *If the leading principal minors alternate in sign such that*

$$|f''_{xx}| = f''_{xx} \leq 0,$$
$$\begin{vmatrix} f''_{xx} & f''_{yx} \\ f''_{xy} & f''_{yy} \end{vmatrix} = f''_{xx}f''_{yy} - f''_{yx}f''_{xy} \geq 0$$

at the point (x^*, y^*), then the Hessian is negative definite, and (x^*, y^*) is a MAXIMUM point of f in S.

- *If the leading principal minors are all positive such that*

$$|f''_{xx}| = f''_{xx} \geq 0,$$
$$\begin{vmatrix} f''_{xx} & f''_{yx} \\ f''_{xy} & f''_{yy} \end{vmatrix} = f''_{xx}f''_{yy} - f''_{yx}f''_{xy} \geq 0$$

at the point (x^*, y^*), then the Hessian is negative definite, and (x^*, y^*) is a MINIMUM point of f in S.

Example 11.2. Show that we have indeed found a maximum in Ex.11.1.

For $K, L > 0$, We find that

$$\pi''_{KK} < 0, \qquad \det(D^2\pi) > 0$$

So the critical point is indeed the maximizer.

Ex. 11.3. A firm producing 3 items x, y, z has the following cost function:

$$\begin{aligned} C_1(x) &= 200 + \frac{1}{100}x^2, \\ C_2(y) &= 200 + y + \frac{1}{300}y^3, \\ C_3(z) &= 200 + 10z, \\ x + y + z &= 2000. \end{aligned}$$

1. Find the optimal values of x, y, z to minimize cost

$$\min_{x,y,z} C(x, y, z) = C_1(x) + C_2(y) + C_3(z)$$

2. Show that the values are indeed optimal.

Ex. 11.4. A firm producing 2 items, x units of A and y units of B has the following cost function:

$$C(x, y) = 2x^2 - 4xy + 4y^2 - 40x - 20y + 514$$

The unit price of A is \$24 and B is \$12.

1. Find the optimal values of x, y to maximize profit.
2. The firm is required to produce exactly 54 units of the combined products. What is the optimal production now?

Ex. 11.5. Profit function

$$\pi(x, y) = px + qy - \alpha x^2 - \beta y^2$$

Find x^*, y^* to maximize profits. Verify using the second-order conditions.

11.3 General Case

Say we want to find the solutions of n choice variables ($\mathbf{x} = (x_1, \ldots, x_n)$)

$$\max_{\mathbf{x}} F(\mathbf{x})$$

11.3.1 Necessary Conditions

This condition requires that the solution $\mathbf{x}^*$ must be a critical point of f, that is $f'(\mathbf{x}^*) = 0$. $\mathbf{x}^*$ will not be the endpoint of the interval under consideration, which means it lies in the INTERIOR of the domain of f.

Theorem 11.1. *Let $F : U \mapsto \mathbb{R}^1$ be a C^1 function defined on a subset U of $\mathbb{R}^n$. If $\mathbf{x}^*$ is a local max or min of F in U, and if $\mathbf{x}^*$ is an interior point of U, then*

$$\frac{\partial F}{\partial x_i}(\mathbf{x}^*) = 0 \text{ for } i = 1, \ldots, n$$

Basically, the FOCs to every choice variable must be 0. We can write the condition in the form of Jacobian

$$DF(\mathbf{x}^*) = \begin{pmatrix} \frac{\partial F}{\partial x_1}(\mathbf{x}^*) & \ldots & \frac{\partial F}{\partial x_n}(\mathbf{x}^*) \end{pmatrix} = \mathbf{0}$$

11.3.2 Sufficient Conditions

We need to use a condition on the second derivatives of F to determine whether the critical point is a max or a min. A C^2 function of n variables has n^2 second-order partial derivatives at each point in its domain. We combine them into a $n \times n$ matrix called the **Hessian** of F

$$D^2F(\mathbf{x}^*) = \begin{pmatrix} \frac{\partial^2 F}{\partial x_1^2}(\mathbf{x}^*) & \ldots & \frac{\partial^2 F}{\partial x_1 \partial x_n}(\mathbf{x}^*) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 F}{\partial x_n \partial x_1}(\mathbf{x}^*) & \ldots & \frac{\partial^2 F}{\partial x_n^2}(\mathbf{x}^*) \end{pmatrix}$$

The Hessian is always a symmetric matrix. Due to Young's theorem and the symmetry of the Hessian, you can also write the Hessian alternatively as

$$D^2F(\mathbf{x}^*) = \begin{pmatrix} \frac{\partial^2 F}{\partial x_1^2}(\mathbf{x}^*) & \ldots & \frac{\partial^2 F}{\partial x_n \partial x_1}(\mathbf{x}^*) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 F}{\partial x_1 \partial x_n}(\mathbf{x}^*) & \ldots & \frac{\partial^2 F}{\partial x_n^2}(\mathbf{x}^*) \end{pmatrix}$$

We follow the convention as demonstrated in the former. Whether the critical point is a min or max or neither depends on the definiteness of the Hessian matrix at that point.

Theorem 11.2. *Let $F : U \mapsto \mathbb{R}^1$ be a C^2 function whose domain is an open set U in $\mathbb{R}^n$. Suppose that $\mathbf{x}^*$ is a critical point of F, then*

1. *If the Hessian $D^2F(\mathbf{x}^*)$ is a NEGATIVE DEFINITE symmetric matrix, then $\mathbf{x}^*$ is a strict LOCAL MAX of F,*
2. *If the Hessian $D^2F(\mathbf{x}^*)$ is a POSITIVE DEFINITE symmetric matrix, then $\mathbf{x}^*$ is a strict LOCAL MIN of F,*
3. *If the Hessian $D^2F(\mathbf{x}^*)$ is INDEFINITE, then $\mathbf{x}^*$ is neither a local max nor a local min of F.*

How to test for definiteness of the Hessian?

Now, to determine the definiteness of the Hessian matrix, we rely on the following theorems.

Theorem 11.3 (Sufficient Conditions for a MAX)**.** *Let $F : U \mapsto \mathbb{R}^1$ be a C^2 function whose domain is an open set U in $\mathbb{R}^1$. Suppose that*

$$\frac{\partial F}{\partial x_i}(\mathbf{x}^*) = 0 \text{ for } i = 1, \ldots, n$$

and that the n leading principal minors of $D^2F(\mathbf{x}^*)$ alternate in sign

$$|F''_{x_1\ x_1}| < 0, \begin{vmatrix} F''_{x_1\ x_1} & F''_{x_2\ x_1} \\ F''_{x_1\ x_2} & F''_{x_2\ x_2} \end{vmatrix} > 0, \begin{vmatrix} F_{x_1\ x_1} & F_{x_2\ x_1} & F_{x_3\ x_1} \\ F_{x_1\ x_2} & F_{x_2\ x_2} & F_{x_3\ x_2} \\ F_{x_1\ x_3} & F_{x_2\ x_3} & F_{x_3\ x_3} \end{vmatrix} < 0, \ldots$$

at $\mathbf{x}^*$. Then $\mathbf{x}^*$ is a strict local max of F.

Theorem 11.4 (Sufficient Conditions for a MIN)**.** *Let $F : U \mapsto \mathbb{R}^1$ be a C^2 function whose domain is an open set U in $\mathbb{R}^1$. Suppose that*

$$\frac{\partial F}{\partial x_i}(\mathbf{x}^*) = 0 \text{ for } i = 1, \ldots, n$$

and that the n leading principal minors of $D^2F(\mathbf{x}^*)$ all positive

$$|F_{x_1\ x_1}| > 0, \begin{vmatrix} F_{x_1\ x_1} & F_{x_2\ x_1} \\ F_{x_1\ x_2} & F_{x_2\ x_2} \end{vmatrix} > 0, \begin{vmatrix} F_{x_1\ x_1} & F_{x_2\ x_1} & F_{x_3\ x_1} \\ F_{x_1\ x_2} & F_{x_2\ x_2} & F_{x_3\ x_2} \\ F_{x_1\ x_3} & F_{x_2\ x_3} & F_{x_3\ x_3} \end{vmatrix} > 0, \ldots$$

at $\mathbf{x}^*$. Then $\mathbf{x}^*$ is a strict local min of F.

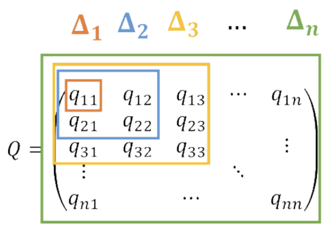


Fig. 11.2. Principal Minors

11.3.3 Sufficient Conditions: Eigenvalues

Another way to test for concavity using the Hessian matrix is to evaluate the Eigenvalues of the Hessian at critical points.

Theorem 11.5 (Eigenvalues Test for Sufficient Conditions)**.** *If the Hessian at a given point has* ***all positive eigenvalues****, it is said to be positive-definite, meaning the function is* ***concave up (convex)*** *at that point. If all the* ***eigenvalues are negative****, it is said to be a negative-definite, equivalent to* ***concave down****.*

When a random matrix $\mathbf{A}$ acts as a scalar multiplier on a vector $\mathbf{x}$, then that vector is called an eigenvector of $\mathbf{x}$. The value of the multiplier is known as an eigenvalue. For the purpose of analyzing Hessians, the eigenvectors are not important, but the eigenvalues are. Because the Hessian of an equation is a square matrix, its eigenvalues can be found. Because Hessians are also symmetric, they have a special property that their eigenvalues will always be real numbers.

Steps in finding the Eigenvectors and Eigenvalues

$$\begin{aligned} \mathbf{A}\mathbf{x} &= \lambda\mathbf{x} \\ \Rightarrow (\mathbf{A}-\lambda)\mathbf{x} &= 0 \\ \Rightarrow (\mathbf{A}-\lambda\mathbf{I})\mathbf{x} &= 0 \\ \iff \det(\mathbf{A}-\lambda\mathbf{I}) &= 0 \end{aligned}$$

Solving this expression gives us the Eigenvalues λ and the eigenvector $\mathbf{x}$.

Checks when you found the eigenvalues

1. (trace) the sum of all the eigenvalues will be the sum of the diagonal of $\mathbf{A}$
2. (determinants) the product of all the eigenvalues is the determinant

The eigenvectors tell you the directions that do not change during some linear transformation, while the eigenvalues tell you the scaling vector of these eigenvectors. As learned earlier, the gradient of a multivariable function can be regarded as the first derivative of a function, and the Hessian matrix is the second derivative.

If the eigenvalues of the Hessian matrix are **positive** at the critical point we found, it means that the gradient of a function at that point is increasing for any point outside the critical point, so the critical point should be a local MIN.

Similarly, if the eigenvalues of the Hessian matrix are negative, then the critical point attains a local MAX.

If the eigenvalues are mixed with 1 positive, and 1 negative, then we have a SADDLE point. The function is concave up in one direction and concave down in the other, about the critical point.

If the eigenvalues are zero, that means the determinant of the Hessian is zero, and we cannot conclude the nature of the critical point.

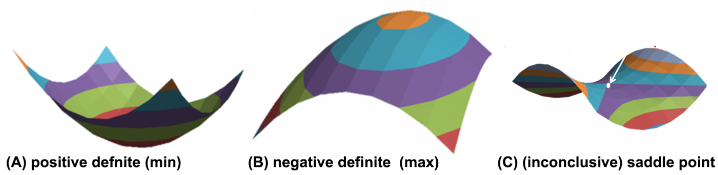


Fig. 11.3. The definiteness of the Hessian

Example 11.3. Suppose $\mathbf{A}$ is a square 2×2 matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Then finding the eigenvalues is to solve

$$\left| \begin{pmatrix} a & b \\ c & d \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \right| = \begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} = 0$$

And the eigenvalues λ are the solutions of

$$(a-\lambda)(d-\lambda) - bc = 0$$

Find the eigenvalues of

$$\begin{pmatrix} 2 & 3 \\ 5 & 4 \end{pmatrix}$$

We need to solve

$$\begin{vmatrix} 2-\lambda & 3 \\ 5 & 4-\lambda \end{vmatrix} = 0$$

which yields $\lambda = 7$ or -1. These are the eigenvalues. We can do a quick verification.

1. (trace) $\lambda_1 + \lambda_2 = 6 = 2 + 4$ (diagonal of $\mathbf{A}$)
2. (determinants) $\lambda_1 \times \lambda_2 = -7 = \det \mathbf{A} (= 2 \times 4 - 5 \times 3)$.

Ex. 11.6. Find the eigenvalues of

$$(a) \begin{pmatrix} 3 & 8 \\ 4 & -1 \end{pmatrix}, \quad (b) \begin{pmatrix} 2 & 6 \\ -1 & 3 \end{pmatrix}$$

Ex. 11.7. Find the eigenvalues and evaluate them at given points, and determine whether the matrix is negative-definite, positive-definite, or indefinite.

$$(a) \begin{pmatrix} 12x^2 & -1 \\ -1 & 2 \end{pmatrix} \quad \text{at } (3,1),$$
$$(b) \begin{pmatrix} 6x & 0 \\ 0 & 6y \end{pmatrix} \quad \text{at } (-1,2),$$
$$(c) \begin{pmatrix} -2y^2 & -4xy \\ -4xy & -2x^2, \end{pmatrix} \quad \text{at } (1,-1) \text{ and } (1,0)$$

Ex. 11.8. Determine the concavity of

$$f(x,y) = x^3 + 2y^3 - xy$$

at the following points
$(0,0),\ (3,3),\ (3,-3),\ (-3,3),\ (-3,-3)$.

Ex. 11.9. Find the critical points of the following functions, and use the eigenvalues calculations to determine the Hessian to determine whether the point is a min, max, or neither

$$(a)\ f(x,y) = 4x + 2y - x^2 - 3y^2,$$
$$(b)\ f(x,y) = x^4 + y^2 - xy.$$

11.4 Local Extreme Points

A (strict) local minimum point is defined in an obvious way, and it should also be clear what we mean by local maximum and minimum values, local extreme points, and local extreme values. Note how these definitions imply that a global extreme point is also a local extreme point; the converse is not true, of course.

These first-order conditions are necessary for a differentiable function to have a local extreme point. However, a critical point does not have to be a local extreme point. A

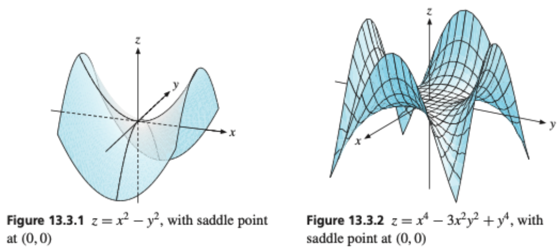


Figure 13.3.1 $z = x^2 - y^2$, with saddle point at $(0, 0)$

Figure 13.3.2 $z = x^4 - 3x^2y^2 + y^4$, with saddle point at $(0, 0)$

Fig. 11.4. Saddle Points

critical point (x_0, y_0) of f which, like point R in Fig. 11.4 is neither a local maximum nor a local minimum point, is called a saddle point of f. Hence: A saddle point (x_0, y_0) is a critical point with the property that there exist points (x, y) arbitrarily close to (x_0, y_0) with $f(x, y) < f(x_0, y_0)$, and there also exist such points with $f(x, y) > f(x_0, y_0)$.

The following theorem can be used to determine the nature of the critical points in most cases.

Theorem 11.1 (Second-order derivatives Test for Local Extrama)**.** *Suppose that $f(x, y)$ is a C^2 function in a domain S, and let (x_0, y_0) be an interior critical point of S, the Hessian is*

$$H = \begin{pmatrix} f''_{xx} & f''_{yx} \\ f''_{xy} & f''_{yy} \end{pmatrix}$$

and its determinant is $\det(H) = f''_{xx} f''_{yy} - f''^2_{xy}$, *then*

1. *if $f''_{xx} < 0$, and $\det(H) > 0$, then (x_0, y_0) is a strict local MAX.*
2. *if $f''_{xx} > 0$, and $\det(H) > 0$, then (x_0, y_0) is a strict local MIN.*
3. *if $\det(H) < 0$, then (x_0, y_0) is a saddle point.*
4. *if $\det(H) = 0$, then (x_0, y_0) could be a local max, local min, or a saddle point.*

Ex. 11.10. A function $f(x, y) = x^2 + 2xy^2 + 2y^2$.

1. Find its partial derivatives of first and second order.
2. Show that its critical points are $(0, 0), (-1, 1), (-1, -1)$, classify them.

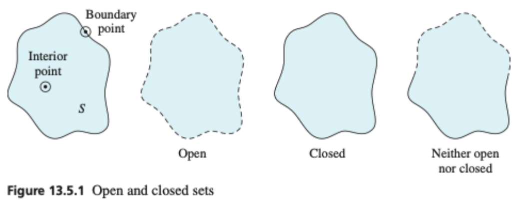


Figure 13.5.1 Open and closed sets

Fig. 11.5. Sets' Nature.

The analysis of local extreme points brings about the topic of a true extreme value. While we can show a point is a local MAX, there is no guarantee that it is truly the MAX of a given set S. Suppose the set S is bounded and closed, then we also need to evaluate the boundary points of S, and compare its value to the critical points to find the maximum/minimum on the set.

Theorem 11.2 (Extreme Value Theorem). *Suppose the function $f(x, y)$ is continuous throughout a* ***nonempty, closed, and bounded*** *set S in the plane. Then there exists both a point (a, b) in S where f has a minimum and a point (c, d) in S where it has a maximum—that is,*

$$f(a, b) \leq f(x, y) \leq f(c, d) \tag{11.1}$$

for all (x, y) in S.

This theorem helps prove the existence of an extreme value. Although we don't go so deep to prove the closedness/ boundedness of a set, it is still good to know about the theorem.

11.5 Comparative Statics & Envelope Theorem

Optimization problems in economics usually involve maximizing or minimizing functions that depend not only on endogenous variables one can choose but also on one or more exogenous parameters like prices, tax rates, income levels, etc. Although these parameters are held constant during the optimization, they vary according to the economic situation. For example, we may calculate a firm's profit-maximizing input and output quantities while treating the prices it faces as parameters. But then we may want to know how the optimal quantities respond to changes in those prices, or in whatever other exogenous parameters affect the problem we are considering.

Consider first the following simple problem. A function f depends on a single variable x as well as on a single parameter r. We wish to maximize $f(x, r)$ w.r.t. x while keeping r constant

$$\max_x f(x, r)$$

The value of x that maximizes f will usually depend on r, so we denote it by $x^*(r)$. Inserting $x^*(r)$ into $f(x, r)$, we obtain the **value function**:

$$f^*(r) = f(x^*(f), r)$$

What happens to the value function as r changes? Assuming that $f^*(r)$ is differentiable, the chain rule yields

$$\frac{df^*(r)}{dr} = \underbrace{f'_x(x^*)\frac{dx}{dr}}_{=0 \text{ in FOC}} + f'_r(x^*) \tag{11.2}$$

If f achieves a maximum at an interior point $x^*(r)$ in the domain of variation for x, then the FOC $f'_x(x^*(r), r) = 0$ is satisfied. It follows that

$$\frac{df^*(r)}{dr} = f'_r(x^*) \tag{11.3}$$

Note that when r is changed, then $f^*(r)$ changes for 2 reasons. First, a change in r changes the value of f^* directly because r is the second variable in $f(x, r)$. Second, a change in r changes the value of the function $x^*(r)$, and hence $f^*(x^*(r), r)$ is changed indirectly. (11.3) shows that the total effect is simply found by computing the partial derivative of $f^*(x^*(r), r)$ w.r.t. r, ignoring entirely the indirect effect of the dependence of x^* on r.

On further reflection, however, you may realize that the first-order condition for $x^*(r)$ to maximize f w.r.t. x implies that any small change in x, whether or not it is induced by a small change in r, must have a negligible effect on the value of f. The collection of all the value functions of $f^*(r)$ when r changes constitutes the **Envelope Theorem**.

Theorem 11.1 (Envelope Theorem)**.** *Assume that $f(\mathbf{x},\mathbf{r})$ is differentiable. If $f^*(\mathbf{r}) = \max_{\mathbf{x}} f(\mathbf{x},\mathbf{r})$ and if $\mathbf{x}^*(\mathbf{r})$ is the value of $\mathbf{x}$ that maximizes $f(\mathbf{x},\mathbf{r})$ then*

$$\frac{\partial f^*(\mathbf{r})}{\partial r_j} = \frac{\partial f(\mathbf{x}^*(\mathbf{r}),\mathbf{r})}{\partial r_j}$$

for $j = 1, \ldots, m$ provided that the partial derivative exists.

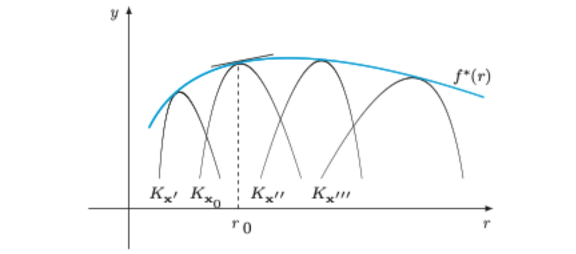


Figure 13.7.1 The curve $y = f^*(r)$ is the envelope of all the curves $y = f(\mathbf{x}, r)$

Fig. 11.6. Envelope shape

In Economics, these results are crucial for a comparative statics analysis of the steady state when we change some parameters. It also provides a shortcut.

Example 11.4. Consider the maximization problem

$$\max\ f(x) = -x^2 + 2ax + 4a^2$$

What is the effect of an increase in a on the maximal value of $f(x, a)$?

First, Find the critical point. Taking FOC yields $x^* = a$

1. (Direct Solution) Inserting to the objective function $f(a) = 5a^2$ and so the effect of a on f is $df/da = 10a$.

2. (Envelop Theorem)

$$f'(a) = \frac{\partial f}{\partial a} = 2x + 8a$$

evaluate at $x = a$ also yields $10a$.

11.6 Review Exercises

Ex. 11.11. For the following functions, find the critical points (FOC) and determine whether they are the local maximum, local minimum, saddle points, or undetermined using the Second-order conditions. [1]

$$
\begin{aligned}
&(a)\ f(x,y) = -2x^2 + 2xy - y^2 + 18x - 14y + 4\\
&(b)\ f(x,y) = 4x^2 - xy + y^2 - x^3,\\
&(c)\ f(x,y) = 3x^4 + 3x^2y - y^3,\\
&(d)\ f(x,y) = x^4 + x^2 - 6xy + 3y^2\\
&(e)\ f(x,y) = x^2 - 6xy + 2y^2 + 10x + 2y - 5\\
&(f)\ f(x,y) = xy^2 + x^3y - xy\\
&(g)\ f(x,y,z) = 2x^2 + y^2 + 4z^2 - x + 2z\\
&(h)\ f(x,y,z) = x^2 + 6xy + y^2 - 3yz + 4z^2 - 10x - 5y - 21z.
\end{aligned}
$$

Example 11.5. Suppose

$$f(x,y) = x^4 + y^2 - xy$$

The critical point is found by

$$
\begin{aligned}
(x) &: \frac{\partial f}{\partial x} = 4x^3 - y = 0 \iff y = 4x^3,\\
(y) &: \frac{\partial f}{\partial y} = 2y - x = 0 \iff y = x/2.
\end{aligned}
$$

Solving for x yields the following critical points

$$(x^*, y^*) = (0,0),\ (\frac{1}{2\sqrt{2}}, \frac{1}{4\sqrt{2}}),\ (-\frac{1}{2\sqrt{2}}, -\frac{1}{4\sqrt{2}})$$

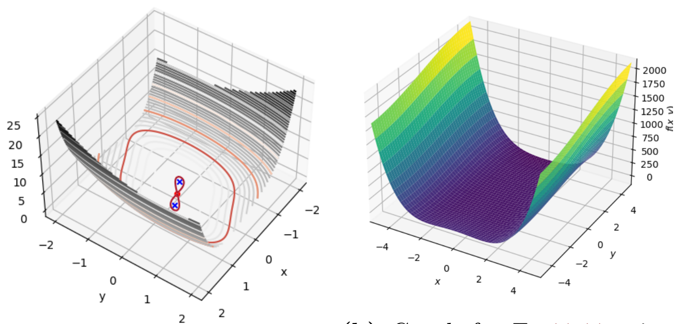
To verify the extremum, we evaluate the Hessian matrix at the critical points

$$H = \begin{pmatrix} 12x^2 & -1 \\ -1 & 2 \end{pmatrix}$$

The first principal minor is $12x^2$. The second principal minor is H itself where

$$|H| = 24x^2 - 1$$

Applying Theorem 11.1, we conclude that $(0,0)$ is a saddle point. The other 2 critical points $(\frac{1}{2\sqrt{2}}, \frac{1}{4\sqrt{2}})$, $(-\frac{1}{2\sqrt{2}}, -\frac{1}{4\sqrt{2}})$ are minima.



(a) Graph for Example 11.5

(b) Graph for Ex.11.11 using plot3d command

Fig. 11.7. Illustrations of using Python

[1] Source: http://www.columbia.edu/ md3405/Unconstrained_Optimization_Solutions.pdf and https://s3.wp.wsu.edu/uploads/sites/289/2018/09/Problem-set-1-unconstrained-opt-f-2018-ans-key.pdf.

Chapter 12

Constrained Optimization

Let's start with a simple 2 variables optimization problem with 1 equality constraint.

$$\max_{x_1, x_2} f(x_1, x_2) \tag{12.1}$$

$$\textbf{s.t.} \quad h(x_1, x_2) = c. \tag{12.2}$$

Eq (12.1) is called the OBJECTIVE function that we want to maximize or minimize. However, it is now "constrained" by (12.2), which is also called the ADMISSIBLE SET. Thus, when we want to perform optimization, we need to make sure the solutions satisfy the admissible set as well.

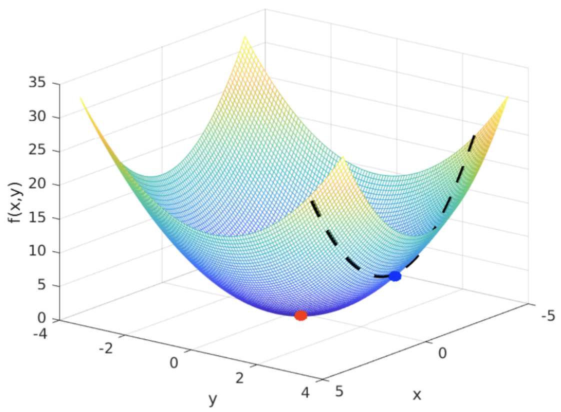


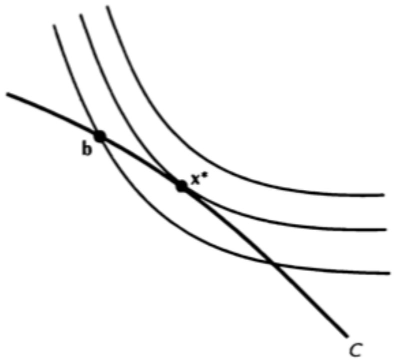
Fig. 12.1. Illustration of a minimization problem with constraint

As illustrated, the function above is convex, which produces a MIN. Without constraints, the solutions can be the red dot. However, if a constraint is introduced, the optimal now is the blue dot, which is on both the admissible set and the objective function.

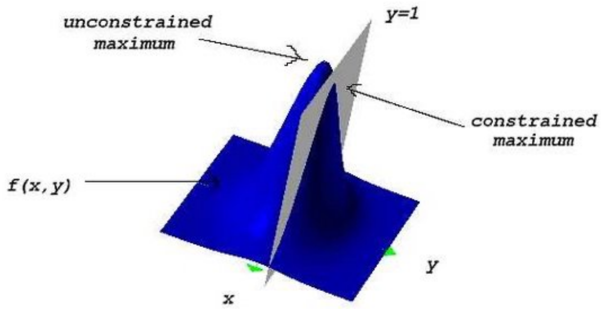
12.1 Equality Constraints (Lagrangian)

12.1.1 First-order Necessary Conditions

Geometrically, we have the constraint set C in the x_1x_2-plane and the level curves of the objective function f. The maximizer is the highest level curve of f to touch (be tangent to) C.



(a) At the cosntrained max x^*, the highest level curve of f is tangent to the constraint set C



(b) Constrained vs Unconstrained Maximum

Fig. 12.2. Visualization of a Constrained Optimum.

The fact that the level curve of f is tangent to the constraint set C at the constrained maximizer $\mathbf{x}^*$ means that the slope of the level set of f equals the slope of the constraint curve C at $\mathbf{x}^*$.

Using the Implicit Function Theorem, the slope of the level set at $\mathbf{x}^*$ is

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

The slope of the constraint set at $\mathbf{x}^*$ is

$$-\frac{\frac{\partial h}{\partial x_1}(\mathbf{x}^*)}{\frac{\partial h}{\partial x_2}(\mathbf{x}^*)}.$$

These two slopes are equal at $\mathbf{x}^*$ implies

$$\frac{\frac{\partial f}{\partial x_1}(\mathbf{x}^*)}{\frac{\partial f}{\partial x_2}(\mathbf{x}^*)} = \frac{\frac{\partial h}{\partial x_1}(\mathbf{x}^*)}{\frac{\partial h}{\partial x_2}(\mathbf{x}^*)}.$$

Rearranging

$$\frac{\frac{\partial f}{\partial x_1}(\mathbf{x}^*)}{\frac{\partial h}{\partial x_1}(\mathbf{x}^*)} = \frac{\frac{\partial f}{\partial x_2}(\mathbf{x}^*)}{\frac{\partial h}{\partial x_2}(\mathbf{x}^*)}. \tag{12.3}$$

To avoid working with (possibly) zero denominators, let λ denote the common value of the 2 quotients

$$\frac{\frac{\partial f}{\partial x_1}(\mathbf{x}^*)}{\frac{\partial h}{\partial x_1}(\mathbf{x}^*)} = \frac{\frac{\partial f}{\partial x_2}(\mathbf{x}^*)}{\frac{\partial h}{\partial x_2}(\mathbf{x}^*)} = \lambda.$$

Rewrite the above as the 2 equations

$$\frac{\partial f}{\partial x_1}(\mathbf{x}^*) - \lambda\frac{\partial h}{\partial x_1}(\mathbf{x}^*) = 0,$$
$$\frac{\partial f}{\partial x_2}(\mathbf{x}^*) - \lambda\frac{\partial h}{\partial x_2}(\mathbf{x}^*) = 0.$$

To solve for 3 unknowns (x_1, x_2, λ) we need 3 equations. Aside from the above 2, the third equation is the budget constraint.

$$
\begin{aligned}
&\frac{\partial f}{\partial x_1}(\mathbf{x}^*) - \lambda \frac{\partial h}{\partial x_1}(\mathbf{x}^*) = 0, \\
&\frac{\partial f}{\partial x_2}(\mathbf{x}^*) - \lambda \frac{\partial h}{\partial x_2}(\mathbf{x}^*) = 0, \\
&h(x_1, x_2) - c = 0.
\end{aligned}
$$

The above system can be written in a convenient way – the **Lagrangian**

$$\mathcal{L}(x_1, x_2, \lambda) = f(x_1, x_2) - \lambda\,[h(x_1, x_2) - c]$$

and solve for $\mathbf{x}^*$ by computing $\partial\mathcal{L}/\partial x_1 = \partial\mathcal{L}/\partial x_2 = \partial\mathcal{L}/\partial\lambda = 0$, which gives us exactly the above results.

This technique will not work if both $\partial h/\partial x_1$ and $\partial h/\partial x_2$ were zero at the maximizer $\mathbf{x}^*$ in Eq. (12.3). For this reason, we need to assume that $\partial h/\partial x_1$ or $\partial h/\partial x_2$ (or both) is not zero at the constrained maximizer. This is called a **constraint qualification**. If the constraint is LINEAR, this constraint is automatically satisfied.

Theorem 12.1. *Let f and h be C^1 functions of 2 variables. Suppose that $\mathbf{x}^* = (x_1^*, x_2^*)$ is a solution of the problem*

$$
\begin{aligned}
&\max f(x_1, x_2) \\
&s.t.\ h(x_1, x_2) = c
\end{aligned}
$$

Suppose further that (x_1^*, x_2^*) is not a critical point of h. Then, there is a real number λ^* such that $(x_1^*, x_2^*, \lambda^*)$ is a critical point of the Lagrangian function

$$\mathcal{L}(x_1, x_2, \lambda) \equiv f(x_1, x_2) - \lambda[h(x_1, x_2) - c].$$

In other words, at $(x_1^*, x_2^*, \lambda^*)$ we can obtain the First-order conditions

$$\frac{\partial\mathcal{L}}{\partial x_1} = \frac{\partial\mathcal{L}}{\partial x_2} = \frac{\partial\mathcal{L}}{\partial\lambda} = 0.$$

or

$$\nabla\mathcal{L}(x_1, x_2, \lambda) = 0$$

The above theorem provides us with a cookbook method to solve a constrained optimization problem with 2 variables and 1 equality constraint. λ is interpreted as the Lagrangian Multiplier.

Example 12.1. The problem is

$$
\begin{aligned}
&\max_{x,y}\ f(x, y) = xy \\
&s.t.\ 2x + y = 100
\end{aligned}
$$

Method 1: Lagrangian Method

The Lagrangian is

$$\mathcal{L} = xy - \lambda(2x + y - 100)$$

The FOCs:

$$\mathcal{L}'_x = y - 2\lambda = 0,$$
$$\mathcal{L}'_y = x - \lambda = 0,$$
$$\mathcal{L}'_\lambda = 2x + y - 100 = 0,$$

which yields the solution $(x^*, y^*) = (25, 50)$.

Method 2: Substitution Method or "Naive" Method

We can turn the constrained optimization problem into an unconstrained problem. From the constraint, we have $y = 100 - x$, the problem becomes

$$\max_x x(100 - 2x)$$

The FOC is

$$100 - 4x = 0,$$

which also gives $x^* = 25, y^* = 50$.

Ex. 12.1. Use the Lagrangian method for these problems
(a) $\max_{x,y} \quad f(x, y) = x^{1/2}y^{1/3} \quad \text{s.t.} \quad 2x + 4y = m$
(b) $\max_{x,y} \quad f(x, y) = \log(x) + \beta \log(y) \quad \text{s.t.} \quad x + y = c$
(c) $\max_{x,y} \quad f(x, y) = \log(x) + \gamma \log(1 - y) \quad \text{s.t.} \quad (1 - \tau)x + c = yw$

12.1.2 Sufficient Conditions

The Lagrangian method gives the necessary conditions for the local solution of constrained optimization problems. To confirm that we have found the solution, however, a more careful check is needed.

Theorem 12.2. *Suppose (x_0, y_0) is a critical point for the Lagrangian $\mathcal{L}$*

- *If the Lagrangian is concave, then (x_0, y_0) is a MAX*
- *If the Lagrangian is convex, then (x_0, y_0) is a MIN*

To show concave, show that $\mathcal{L}' > 0, \mathcal{L}'' < 0$. To show convex, show that $\mathcal{L}' > 0, \mathcal{L}'' > 0$.

Normally, for a linear constraint, this condition is neatly verified (sometimes, you don't need to even provide the sufficient conditions if the constraint is linear). However, when the constraint is nonlinear, typically the case when there is a product of the choice variables presented, or the constraint is nonlinear, then we may need to rely on the Hessian.

Intuitively, the second-order test for a constrained problem

- should involve the definiteness of some Hessian matrix, BUT
- should only concern with the directions along the constraint set,

which will involve the notion of a **bordered Hessian matrix**. This is the most common method but cumbersome to prove and will be presented in a general case. We provide here only the sufficient conditions for a problem of **2 variables and 1 constraint**, which is the most common.

Theorem 12.3. *Let f, h be C^2 functions on $\mathbb{R}^2$. Consider the problem*

$$\max_{x,y} f(x,y) \quad \textbf{\textit{s.t.}} \;\; h(x,y) = c \;\; \text{for } c \in C_h \text{(constraint set)}$$

The Lagrangian is

$$\mathcal{L}(x,y,\lambda) = f(x,y) - \lambda(h(x,y) - c)$$

Suppose that (x^*, y^*, λ^*) satisfies the following FOCs

$$\mathcal{L}'_x = \mathcal{L}'_y = \mathcal{L}'_\lambda = 0 \quad \text{at } (x^*, y^*, \lambda^*)$$

and the bordered Hessian matrix is

$$H = \begin{pmatrix} 0 & \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} \\ \frac{\partial h}{\partial x} & \frac{\partial^2 \mathcal{L}}{\partial x^2} & \frac{\partial^2 \mathcal{L}}{\partial x \partial y} \\ \frac{\partial h}{\partial y} & \frac{\partial^2 \mathcal{L}}{\partial y \partial x} & \frac{\partial^2 \mathcal{L}}{\partial y^2} \end{pmatrix}$$

1. *if $\det(H) > 0$ at (x^*, y^*), then (x^*, y^*) is the local MAX of f on C_h.*
2. *if $\det(H) < 0$ at (x^*, y^*), then (x^*, y^*) is the local MIN of f on C_h.*

Proof. The idea is to convey the original problem into a problem of unconstrained one-variable optimization. The assumptions imply that

$$\frac{\partial h}{\partial x}(x^*, y^*) \neq 0,$$
$$\frac{\partial h}{\partial y}(x^*, y^*) \neq 0$$

Then, by the IFT, the constraint set C_h can be considered as a graph of a C^1 function

$$y = \phi(x) \;\; \text{around } (x^*, y^*) \tag{12.4}$$

and thus one can write the constraint as

$$h(x, \phi(x)) = C \text{ for } x \text{ near } x^* \tag{12.5}$$

Differentiating (12.5) wrt x yields

$$\frac{\partial h}{\partial x}(x, \phi(x)) + \frac{\partial h}{\partial y}(x, \phi(y)) \cdot \phi'(x) = 0 \tag{12.6}$$

rearrange

$$\phi'(x) = -\frac{\frac{\partial h}{\partial x}(x, \phi(x))}{\frac{\partial h}{\partial y}(x, \phi(x))} \tag{12.7}$$

Let

$$F(x) \equiv f(x, \phi(x)) \tag{12.8}$$

be f evaluated on C_h a function of 1 unconstrained variable. The problem is simplified to an unconstrained maximization problem of 1 variable. We will test the first and second-order of only this function.

FOC:

$$F'(x) = \frac{\partial f}{\partial x}(x, \phi(x)) + \frac{\partial f}{\partial y}(x, \phi(x)) \cdot \phi'(x) \tag{12.9}$$

Since (12.6) is zero, we can multiply (12.6) by $-\lambda$ and add to (12.9) without resulting in any changes, by evaluating both at $x = x^*$, we have

$$\begin{aligned}
F'(x^*) &= \left(\frac{\partial f}{\partial x}(x^*, y^*) - \lambda^* \frac{\partial h}{\partial x}(x^*, y^*)\right) + \phi'(x) \left(\frac{\partial f}{\partial y}(x^*, y^*) - \lambda^* \frac{\partial h}{\partial y}(x^*, y^*)\right) \\
&= \frac{\partial \mathcal{L}}{\partial x}(x^*, y^*) + \phi'(x^*) \frac{\partial \mathcal{L}}{\partial y}(x^*, y^*) \\
&= 0 \quad \text{by the original FOC of this Theorem}
\end{aligned}$$

Now, we take another derivative of $F(x)$ at x^*, setting $y^* = \phi'(x^*)$:

$$\begin{aligned}
F''(x^*) &= \frac{\partial}{\partial x}\left(\frac{\partial F}{\partial x}\right) \\
&= \frac{\partial}{\partial x}\left(\frac{\partial \mathcal{L}}{\partial x}(x^*, y^*) + \phi'(x^*) \frac{\partial \mathcal{L}}{\partial y}(x^*, y^*)\right) \\
&= \frac{\partial^2 \mathcal{L}}{\partial x^2} + \frac{\partial^2 \mathcal{L}}{\partial x \partial y}\frac{\partial y}{\partial x} + \phi'(x)\left[\frac{\partial^2 \mathcal{L}}{\partial y \partial x} + \frac{\partial^2 \mathcal{L}}{\partial y^2}\frac{\partial y}{\partial x}\right] + \frac{\partial \mathcal{L}}{\partial y}\phi''(x) \\
&= \frac{\partial^2 \mathcal{L}}{\partial x^2} + 2\frac{\partial^2 \mathcal{L}}{\partial x \partial y}\phi'(x^*) + \frac{\partial^2 \mathcal{L}}{\partial y^2}\phi'(x^*)^2 \quad \left(\text{since } \frac{\partial \mathcal{L}}{\partial y} = 0 \text{ due to FOC}\right) \\
&= \frac{\partial^2 \mathcal{L}}{\partial x^2} + 2\frac{\partial^2 \mathcal{L}}{\partial x \partial y}\left(-\frac{\frac{\partial h}{\partial x}}{\frac{\partial h}{\partial y}}\right) + \frac{\partial^2 \mathcal{L}}{\partial y^2}\left(-\frac{\frac{\partial h}{\partial x}}{\frac{\partial h}{\partial y}}\right)^2 \\
&= \frac{1}{\left(\frac{\partial h}{\partial y}\right)^2}\left[\frac{\partial^2 \mathcal{L}}{\partial x^2}\left(\frac{\partial h}{\partial y}\right)^2 - 2\frac{\partial^2 \mathcal{L}}{\partial x \partial y}\frac{\partial h}{\partial x}\frac{\partial h}{\partial y} + \frac{\partial^2 \mathcal{L}}{\partial y^2}\left(\frac{\partial h}{\partial x}\right)^2\right] \\
&= \frac{1}{\left(\frac{\partial h}{\partial y}\right)^2}(-|H|)
\end{aligned} \tag{12.10}$$

To see why we got (12.10), take the determinant of the bordered Hessian

$$\begin{aligned}
|H| &= 0\begin{vmatrix} \frac{\partial^2 \mathcal{L}}{\partial x^2} & \frac{\partial^2 \mathcal{L}}{\partial x \partial y} \\ \frac{\partial^2 \mathcal{L}}{\partial y \partial x} & \frac{\partial^2 \mathcal{L}}{\partial y^2} \end{vmatrix} - \frac{\partial h}{\partial x}\begin{vmatrix} \frac{\partial h}{\partial x} & \frac{\partial^2 \mathcal{L}}{\partial x \partial y} \\ \frac{\partial h}{\partial y} & \frac{\partial^2 \mathcal{L}}{\partial y^2} \end{vmatrix} + \frac{\partial h}{\partial y}\begin{vmatrix} \frac{\partial h}{\partial x} & \frac{\partial^2 \mathcal{L}}{\partial x^2} \\ \frac{\partial h}{\partial y} & \frac{\partial^2 \mathcal{L}}{\partial y \partial x} \end{vmatrix} \\
&= -\frac{\partial h}{\partial x}\left(\frac{\partial h}{\partial x}\frac{\partial^2 \mathcal{L}}{\partial y^2} - \frac{\partial h}{\partial y}\frac{\partial^2 \mathcal{L}}{\partial x \partial y}\right) + \frac{\partial h}{\partial y}\left(\frac{\partial h}{\partial x}\frac{\partial^2 \mathcal{L}}{\partial y \partial x} - \frac{\partial h}{\partial y}\frac{\partial^2 \mathcal{L}}{\partial x^2}\right) \\
&= -\left(\frac{\partial h}{\partial x}\right)^2\frac{\partial^2 \mathcal{L}}{\partial y^2} + 2\frac{\partial^2 \mathcal{L}}{\partial x \partial y}\frac{\partial h}{\partial x}\frac{\partial h}{\partial y} - \left(\frac{\partial h}{\partial y}\right)^2\frac{\partial^2 \mathcal{L}}{\partial x^2},
\end{aligned}$$

Thus,

1. If we have $|H| > 0$, then (12.10) becomes negative, and we have a MAX.
2. If we have $|H| < 0$, then (12.10) becomes positive, and we have a MIN.

■

Ex. 12.2. Find the maximum and verify with SOC (p.464 of Simon and Blume (1994))

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

12.1.3 General Case of n choice variables and k constraints

Move on to the general case. Here, we borrow the lecture note from Tornike Kadeishvili [(1)]. For a detailed exposition, see Simon and Blume, 1994 (Chapter 19.3)

Let us consider an optimization problem for n variables with k constraints such that

$$\max_{\mathbf{x}}(\min) f \underbrace{(x_1, \ldots, x_n)}_{\mathbf{x}}$$
$$s.t. \ h_i(\mathbf{x}) = c_i \quad \text{for} \quad i = 1, \ldots, k.$$

NDCQ

Before we proceed, there is one caveat. For the constraint optimization to work, the constraints have to pass a **Nondegenrate Constraint Qualification (NDQC)** where

$$rank \begin{pmatrix} \frac{\partial h_1}{\partial x_1} & \cdots & \frac{\partial h_1}{\partial x_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial h_k}{\partial x_1} & \cdots & \frac{\partial h_k}{\partial x_n} \end{pmatrix} = m.$$

In other words, the gradient vectors $Dh_1(x^*), \ldots, Dh_m(x^*)$ are linear independent in $\mathbb{R}^n$.

Assume that NDQC is satisfied. There will be k Lagrangian multipliers λ_i for $i = 1, \ldots, k$. The Lagrangian is

$$\mathcal{L}(\mathbf{x}, \lambda_i) = f(\mathbf{x}) - \sum_{i=1}^{k} \lambda_i (h_i(\mathbf{x}) - c_i)$$

As usual, the FOC is just

$$\frac{\partial \mathcal{L}}{\partial x_i} = 0 \quad \text{for } i = 1, \ldots, n$$

For sufficient conditions, we need to use the notion of Bordered Hessian Matrix. The construction of such a matrix is

$$H = \left(\begin{array}{ccc|ccc} 0 & \ldots & 0 & B_{11} & \ldots & B_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & B_{m1} & \ldots & B_{mn} \\ \hline B_{11} & \ldots & B_{m1} & a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ B_{1n} & \ldots & B_{mn} & a_{1n} & \ldots & a_{nn} \end{array}\right)$$

[(1)] https://rmi.tsu.ge/ kade/LecturesT.Kadeishvili/MathEconomics/Term4/

in short, it looks like this

$$H = \begin{pmatrix} 0 & \mathbf{B} \\ \mathbf{B}^T & \text{Ł} \end{pmatrix}$$

where $\mathbf{B}$ is the matrix of derivatives of constraints h_i wrt to $\mathbf{x}$, and Ł is the matrix of second-order derivatives of the Lagrangian $\mathcal{L}$ wrt to $\mathbf{x}$.

In our case, it looks like this

$$H = \left(\begin{array}{ccc|ccc} 0 & \dots & 0 & \frac{\partial h_1}{\partial x_1} & \dots & \frac{\partial h_1}{\partial x_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \frac{\partial h_k}{\partial x_1} & \dots & \frac{\partial h_k}{\partial x_n} \\ \hline \frac{\partial h_1}{\partial x_1} & \dots & \frac{\partial h_k}{\partial x_1} & \frac{\partial^2 \mathcal{L}}{\partial x_1^2} & \dots & \frac{\partial^2 \mathcal{L}}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_1}{\partial x_n} & \dots & \frac{\partial h_k}{\partial x_1} & \frac{\partial^2 \mathcal{L}}{\partial x_n \partial x_1} & \dots & \frac{\partial^2 \mathcal{L}}{\partial x_n^2} \end{array}\right)$$

This $(k+n) \times (k+n)$ matrix has $k+n$ leading principal minors (the biggest one is H itself). The first m matrices $H_1, \dots, H_k$ are zero matrices. The next $k-1$ matrices $H_{k+1}, \dots, H_{2k-1}$ have zero determinant.

The determinant of the next minor H_{2k} is $\pm \det(H')^2$ where H' is the upper $k \times k$ minor of H after block of zeros, so $\det H_{2k}$ does not contain information about f.

And only the determinants of the last $n-k$ leading principal minors

$$H_{2k+1}, H_{2k+2}, \dots, H_{2k+(n-k)=k+n} \equiv H$$

carry information about both, the objective function f and the constraints h_i. Exactly these minors are essential for the following sufficient condition for constraint optimization.

Theorem 12.4 (Second-order Conditions)**.** *Suppose* $\mathbf{x}^*$ *satisfies the FOCs.*

1. *For the bordered Hessian matrix H, the last $n-k$ leading principal minors*

$$H_{2k+1}, H_{2k+2}, \dots, H_{2k+(n-k)=k+n} \equiv H$$

evaluated at the critical point ALTERNATE in signs where the last minor $H_{n+k} \equiv H$ has the sign as $(-1)^n$, then $\mathbf{x}^*$ is a LOCAL MAX.

2. *For the bordered Hessian matrix H, the last $n-k$ leading principal minors*

$$H_{2k+1}, H_{2k+2}, \dots, H_{2k+(n-k)=k+n} \equiv H$$

evaluated at the critical point have the SAME sign where the last minor $H_{n+k} \equiv H$ has the sign as $(-1)^k$, then $\mathbf{x}^*$ is a LOCAL MIN.

which can be summarized as

	H_{2k+1}	H_{2k+2}	$\dots$	H_{k+n-1}	$H_{k+n} \equiv H$
max	$(-1)^{k+1}$	$(-1)^{k+2}$	$\dots$	$(-1)^{n-1}$	$(-1)^n$
min	$(-1)^k$	$(-1)^k$	$\dots$	$(-1)^k$	$(-1)^k$

The case of 3 choice variables and 1 constraint

In this case, $n = 3, k = 1$ so the second-order conditions check is to check the last 2 principal minors.

$$H = \left(\begin{array}{c|ccc} 0 & dh/dx & dh/dy & dh/dz \\ \hline dh/dx & d^2L/dx^2 & d^2L/dydx & d^2L/dzdx \\ dh/dy & d^2L/dxdy & d^2L/dy^2 & d^2L/dzdy \\ dh/dz & d^2L/dxdz & d^2L/dydz & d^2L/dz^2 \end{array}\right)$$

1. For the problem to be a max, the sign of $\det H_4 \equiv \det H$ must be $(-1)^n = (-1)^3 < 0$, and the second-last principal minors $\det H_3$ must > 0.
2. For the problem to be a min, the sign of the last 2 principal minors must be negative as $(-1)^1 < 0$.

Example 12.2. Find extremum of

$$\begin{aligned} F(x, y) &= xy \\ s.t.\ h(x, y) &= x + y = 6. \end{aligned}$$

The Lagrangian is

$$L(x, y) = xy - \lambda(x + y - 6)$$

The FOCs are

$$\begin{aligned} (x) &: \frac{\partial L}{\partial x} = y - \lambda = 0, \\ (y) &: \frac{\partial L}{\partial y} = x - \lambda = 0, \\ (\lambda) &: \frac{\partial L}{\partial \lambda} = x + y - 6 = 0, \end{aligned}$$

which gives

$$x^* = y^* = 3,\ \lambda = 3$$

To tell the nature of its extremum, we test the second-order conditions. The bordered Hessian is

$$H = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

We have $n = 2, k = 1$ so we have to check the $n - k = 1$ last leading principal minors, a.k.a, H itself. Calculation shows that $\det H = 2 > 0$ has the same sign with $(-1)^n = (-1)^2 > 0$ so our critical point is a MAX.

Example 12.3. Find extremum of

$$\begin{aligned} F(x, y, z) &= x^2 + y^2 + z^2 \\ s.t.\ h_1(x, y, z) &= 3x + y + z = 5, \\ h_2(x, y, z) &= x + y + z = 1 \end{aligned}$$

The Lagrangian is

$$L(x, y) = x^2 + y^2 + z^2 - \lambda_1(3x + y + z - 5) - \lambda_2(x + y + z - 1)$$

The FOCs are

$$
\begin{aligned}
(x): \frac{\partial L}{\partial x} &= 2x - 3\lambda_1 - \lambda_2 = 0,\\
(y): \frac{\partial L}{\partial y} &= 2y - \lambda_1 - \lambda_2 = 0,\\
(z): \frac{\partial L}{\partial z} &= 2z - \lambda_1 - \lambda_2 = 0,\\
(\lambda_1): \frac{\partial L}{\partial \lambda_1} &= 3x + y + z - 5 = 0,\\
(\lambda_2): \frac{\partial L}{\partial \lambda_2} &= x + y + z - 1 = 0
\end{aligned}
$$

which gives

$$x^* = 2, \quad y^* = -1/2, \quad z^* = -1/2, \quad \lambda_1 = 5/2, \quad \lambda_2 = -7/2$$

To tell the nature of its extremum, we test the second-order conditions. The bordered Hessian is

$$
H = \left(\begin{array}{cc|ccc}
0 & 0 & 3 & 1 & 1\\
0 & 0 & 1 & 1 & 1\\
\hline
3 & 1 & 2 & 0 & 0\\
1 & 1 & 0 & 2 & 0\\
1 & 1 & 0 & 0 & 2
\end{array}\right)
$$

We have $n = 3, k = 2$ so we have to check the $n - k = 1$ last leading principal minors, a.k.a, H itself. Calculation shows that $\det H = 16 > 0$ has the same sign with $(-1)^k = (-1)^2 > 0$ so our critical point is a MIN.

Ex. 12.3. Find the extremum

$$
\begin{aligned}
f(x, y, z) &= x^2y^2z^2,\\
s.t. \ \ x^2 + y^2 + z^2 &= 3
\end{aligned}
$$

12.1.4 Review Exercises

Ex. 12.4. Find the extremum of f wrt h. Verify whether they are a min or max (Chiang, 1984).

$$
\begin{aligned}
&(a)\ f(x, y) = xy, && h(x, y) = x + 2y = 2;\\
&(b)\ f(x, y) = x(y + 4), && h(x, y) = x + y = 8;\\
&(c)\ f(x, y) = x - 3y - xy, && h(x, y) = x + y = 6;\\
&(d)\ f(x, y) = 7 - y - x^2, && h(x, y) = x + y = 0;\\
&(e)\ f(x, y) = 5x - 3y, && h(x, y) = x^2 + y^2 = 136
\end{aligned}
$$

Ex. 12.5. Some problems with 3 choice variables and 1 or more constraint. (p. 421 of Simon and Blume (1994)).

(a) Minimize $x_1^2 + x_2^2 + x_3^2$ subject to $x_1 + x_2 + x_3 = 1$.

(b) Maximize $x_1x_2x_3$ subject to $x_1 + x_2 + x_3 = 10$.

(c) Maximize $x_1^2 + x_2^2 + x_3^2$ subject to $x_1 + x_2 + x_3 = 5$.

(d) Maximize $yz + xz$ subject to $y^2 + z^2 = 1$ and $xz = 3$.

(e) Maximize xyz subject to $x^2 + y^2 = 1$ and $x + z = 1$

Ex. 12.6. Some problems with utility functions subject to budget constraints.

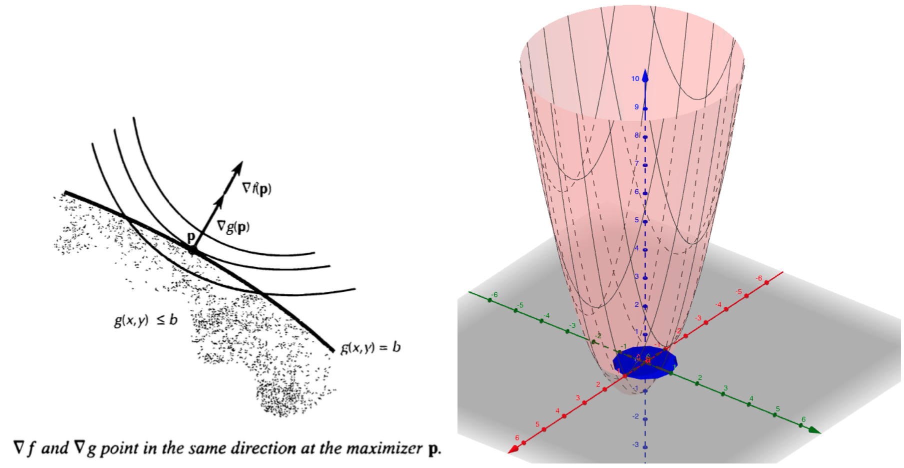
(a) Maximize $u(x_1, x_2, x_3) = x_1^{0.5} x_2^{0.3} x_3^{0.2}$ subject to $p_1x_1 + p_2x_2 + p_3x_3 = I$ and $p_1x_1 + p_2x_2 = 0.8I$.

(b) Maximize $u(x_1, x_2, x_3) = \frac{x_1^{0.5} x_2^{0.3}}{x_3^{0.2}}$ subject to $p_1x_1 + p_2x_2 + p_3x_3 = I$ and $p_1x_1 + p_3x_3 = 0.9I$.

(c) Maximize $u(x_1, x_2, x_3) = \ln(x_1) + 2\ln(x_2) + 3\ln(x_3)$ subject to $p_1x_1 + p_2x_2 + p_3x_3 = I$ and $p_1x_1 + p_2x_2 = 0.8I$.

12.2 Inequality Constraints (KKT)

The inequality constraint problem can be visualized as follows.



(a) FOC where ∇g and ∇f point in the same direction at the max p

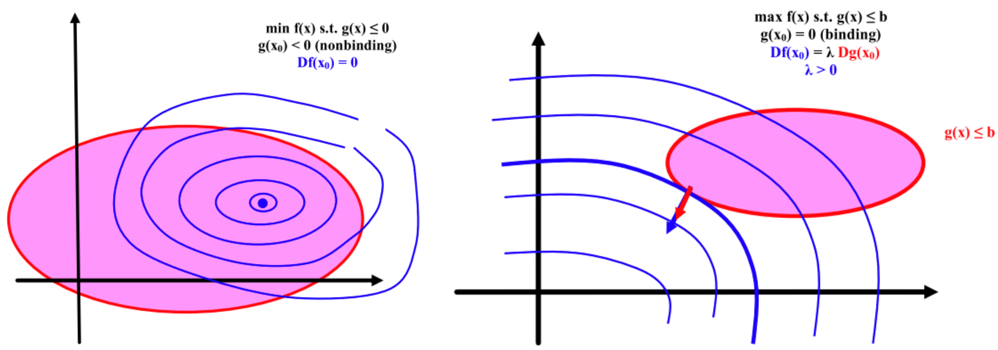
(b) admissible set now includes all points inside the blue circle.

Fig. 12.3. Inequality Constraint Visualized

Say we want to

$$\max f(x,y) \ s.t. \ g(x,y) \le b$$

The constraint may or may not be binding at the optimum.



(a) Constraint Non-Binding

(b) Constraint Binding

Fig. 12.4. Optimal point Binding vs Non-binding cases

12.2.1 KKT First-order Conditions for MAX

In this branch of problems, the constraint has inequality signs.

$$\max f(x,y) \quad s.t. \quad g(x,y) \le c.$$

We solve this problem by employing the cookbook method called KKT conditions (Karush-Kuhn-Tucker).

Theorem 12.1 (The KKT Conditions for MAX). *Suppose we have 2 choice variables and 1 inequality constraint.*

$$\max f(x,y) \ s.t. \ g(x,y) \leq c$$

1. *Construct the Lagrangian*

$$\mathcal{L}(x,y) = f(x,y) - \lambda(g(x,y) - c)$$

2. *FOCs*

$$\begin{aligned}
&\mathcal{L}'_x = f'_x - \lambda g'_x = 0, \\
&\mathcal{L}'_y = f'_y - \lambda g'_y = 0, \\
&\lambda \cdot (g(x,y) - c) = 0, \\
&\lambda \geq 0, \\
&g(x,y) \leq c
\end{aligned}$$

3. *Complimentary slackness condition*

$$\begin{aligned}
&\lambda > 0, \ \textit{the constraint binds so that } g(x,y) = c \\
&\lambda = 0, \ \textit{the constraint does not bind so that } g(x,y) < c
\end{aligned}$$

4. *For a minimum problem, the FOCs are the same, except that* $\lambda \leq 0$.

The two inequalities $\lambda \geq 0$ and $g(x,y) \leq c$ are complementary in the sense that at most one can be "slack" – that is, at most one can hold with inequality. Equivalently, at least one must be an equality. Failure to observe that it is possible to have both $\lambda = 0$ and $g(x,y) = c$ in the complementary slackness condition is the most common error when solving nonlinear programming problems.

Theorem 12.2 (Sufficient Conditions). *Suppose* (x_0, y_0) *satisfies the KKT conditions, then consider the Lagrangian*

$$\mathcal{L}(x,y) = f(x,y) - \lambda(g(x,y) - c)$$

If the Lagrangian is concave, then (x_0, y_0) *is the MAX. If the Lagrangian is convex, then* (x_0, y_0) *is the MIN.*

Example 12.4. Solve the problem

$$\begin{aligned}
\max f(x,y) &= x^2 + y^2 + y - 1, \\
s.t. g(x,y) &= x^2 + y^2 \leq 1
\end{aligned}$$

The Lagrangian is

$$\mathcal{L} = x^2 + y^2 + y - 1 - \lambda(x^2 + y^2 - 1)$$

The FOCs are

$$\begin{aligned}
&\mathcal{L}'_x = 2x - 2\lambda x = 0 \iff 2x(1-\lambda) = 0,\\
&\mathcal{L}'_y = 2y + 1 - 2\lambda y = 0 \iff 2y(1-\lambda) + 1 = 0,\\
&\lambda(x^2 + y^2 - 1) = 0,\\
&\lambda \geq 0 \quad (*),\\
&x^2 + y^2 \leq 1 \quad (**)
\end{aligned}$$

The last condition implies that

$$\begin{cases} \lambda > 0, x^2 + y^2 = 1 \quad \text{or} \\ \lambda = 0, x^2 + y^2 < 1 \end{cases} \tag{12.11}$$

(i) Suppose the constraint binds such that

$$x^2 + y^2 = 1$$

From $\mathcal{L}'_x = 0$, that means $x = 0$, and so $y = \pm 1$. Try $y = 1$, then $\mathcal{L}'_y = 0$ implies $\lambda = 3/2$, which satisfies 12.11, so $(0, 1)$ is a candidate for optimality. Next, try $y = -1$ then $\mathcal{L}'_y = 0$ implies $\lambda = 1/2$, which satisfies 12.11, so $(0, -1)$ is another candidate for optimality.

(ii) Suppose the constraint does not bind (or slacks) such that

$$x^2 + y^2 < 1$$

Then, $\lambda = 0$, the FOC $\mathcal{L}'_x = 0$ implies that $x = 0$, and therefore, $-1 < y < 1$. Since $\lambda = 0$, the FOC $\mathcal{L}'_y = 0$ yields $y = -1/2$. Since these values satisfy 12.11, $(0, -1/2)$ is another candidate.

In summary, the 3 candidates for optimality are

$$f(0, 1) = 1, \quad f(0, -1) = -1, \quad f(0, -1/2) = -5/4.$$

Since the function is continuous over a closed, bounded set, by the Extreme Value Theorem, a solution exists. It is easy to see from the value functions that $(x, y) = (0, 1)$ is the solution to maximize f.

Ex. 12.7. Use the KKT conditions to find solutions to

$$\begin{aligned}
&(a)\ \max\ -x^2 - y^2 \quad s.t.\ \ x - 3y \leq 10,\\
&(b)\ \max\ x^{1/2} + \sqrt{y} \quad s.t.\ \ px + qy \leq m,\\
&(c)\ \max\ 4 - \frac{1}{2}x^2 - 4y^2 \quad s.t.\ \ 6x - 4y \leq a
\end{aligned}$$

Write $V(a)$ as the value function, verify that $V'(a) = \lambda$.

$$\begin{aligned}
&(d)\ \max\ x^2 + 2y^2 - x \quad s.t.\ \ x^2 + y^2 \leq 1,\\
&(e)\ \max\ f(x, y) = 2 - (x-1)^2 - e^{y^2} \quad s.t.\ \ x^2 + y^2 \leq 1/2
\end{aligned}$$

Let $f^*(a)$ as the value function, verify that $df^*(a)/da = \lambda$.

12.2.2 KKT First-order Conditions for MIN

For a minimization problem, you have 3 options

1. **Flip the sign of the objective function**, then we will turn a Minimization problem into a Maximization problem and its FOCs follow suit.
2. Keep the constraints as is (where all constraints are $\leq$), and the FOCs are the same as the MAXIMIZATION problem **except that** $\lambda \leq 0$.
3. Flip the **signs of the constraints so that they have the form** $\geq$, then the FOCs are the same as the MAXIMIZATION problem where $\lambda \geq 0$.

It is easier to follow the last option.

Example 12.5.

$$\begin{aligned} \min f(x,y) &= 2y - x^2 \\ s.t.\ x^2 + y^2 &\leq 1 \end{aligned}$$

Rewrite the problem as

$$\begin{aligned} \min f(x,y) &= 2y - x^2 \\ s.t.\ \ -x^2 - y^2 &\geq -1 \end{aligned}$$

The Lagrangian is

$$L(x,y,\lambda) = 2y - x^2 - \lambda(-x^2 - y^2 + 1)$$

FOCs:

$$\begin{aligned} &(i)\ \frac{\partial L}{\partial x} = 0 \iff -2x + 2\lambda x = 0 \\ &(ii)\ \frac{\partial L}{\partial y} = 0 \iff 2 + 2\lambda y = 0 \\ &(iii)\ \lambda \cdot (-x^2 - y^2 + 1) = 0, \\ &(iv)\ \lambda \geq 0 \text{ (if } \lambda > 0\text{, constraint binds).} \end{aligned}$$

From $(i), (ii)$, we can derive $\lambda = 1, y = -1$. Since $\lambda > 0$, the constraint binds and we have $x^2 + y^2 = 1$. Since $y = -1$, we have $x = 0$, as the optimum.

12.2.3 Multiple Inequality Constraints

Consider an optimization problem of n choice variables and m inequality constraints

$$\begin{aligned} &\max f\underbrace{(x_1, \ldots, x_n)}_{\mathbf{x}} \\ &s.t. g_1(\mathbf{x}) \leq c_1, \\ &\quad \ldots, \\ &\quad g_m(\mathbf{x}) \leq c_m \end{aligned} \tag{12.12}$$

Theorem 12.3 (KKT Formulation). *Steps in solving the problem* (12.12)

1. *Write down the Lagrangian*

$$\mathcal{L}(\mathbf{x}) = f(\mathbf{x}) - \sum_{j=1}^{m} \lambda_j (g_j(\mathbf{x}) - c_j)$$

2. *FOCs:*

$$\frac{\partial \mathcal{L}}{\partial x_i} = 0,$$

   for each $i = 1, \dots, n$.

3. *Complementary slackness*

$$\lambda_j \geq 0, g_j(\mathbf{x}) = c_j \ \textbf{\textit{or}}\ \lambda_j = 0, g_j(\mathbf{x}) < c_j$$

   for $j = 1, \dots, m$. *Can also be summarized as*

$$\lambda_j \cdot g_j(\mathbf{x}) = 0.$$

4. *Find all* $\mathbf{x} = (x_1, \dots, x_n)$ *associated with their* $\lambda_1, \dots, \lambda_m$ *that satisfy FOCs and the complementary slackness. These are the solution candidates, and at least 1 of them solves the problem if it has a solution.*

Example 12.6. The problem is

$$\max\ x + 3y - 4e^{-x-y}$$
$$s.t. \begin{cases} 2 - x \geq 2y \\ x - 1 \leq -y \end{cases}$$

Write the problem as

$$\max\ x + 3y - 4e^{-x-y}$$
$$s.t. \begin{cases} x + 2y \leq 2 \\ x - 1 \leq -y \end{cases}$$

The Lagrangian is

$$\mathcal{L}(x, y) = x + 3y - 4e^{-x-y} - \lambda_1(x + 2y - 2) - \lambda_2(x + y - 1)$$

KKT conditions

$$\begin{aligned}
&(i)\ \mathcal{L}'_x = 1 + 4e^{-x-y} - \lambda_1 - \lambda_2 = 0,\\
&(ii)\ \mathcal{L}'_y = 3 + 4e^{-x-y} - 2\lambda_1 - \lambda_2 = 0,\\
&\lambda_1 \cdot (x + 2y - 2) = 0,\\
&\lambda_2 \cdot (x - 1 + y) = 0,\\
&(iii)\ \lambda_1 \geq 0\ (\lambda_1 = 0 \iff x + 2y < 2),\\
&(iv)\ \lambda_2 \geq 0\ (\lambda_2 = 0 \iff x + y < 1)
\end{aligned}$$

Since $\mathcal{L}$ is concave (prove this by evaluating the Hessian matrix), the KKT conditions are both necessary and sufficient for optimality.

From (ii), (i), we get $\lambda_1 = 2 > 0$, thus making (iii) binds such that $x + 2y = 2$. Suppose $\lambda_2 = 0$, there is a contradiction. Suppose $\lambda_2 > 0$, from (iv) we deduce $x+y = 1$. Using (i) and (ii), we can find $\lambda_2 = e^{-1}(4 - e) > 0$. Thus, the solution is

$$(x^*, y^*, \lambda_1, \lambda_2) = (0,\ 1, 2\ ,\ e^{-1}.(4 - e))$$

A general method for finding all candidates for optimality in a nonlinear programming problem with two constraints can be formulated as follows: First, examine the case where both constraints bind. Next, examine the two cases where only one constraint binds. Finally, examine the case where neither constraint binds. In each case, find all vectors $\mathbf{x}$, with associated non-negative values of the Lagrange multipliers, that satisfy all the relevant conditions—if any do. Then calculate the value of the objective function for these values of $\mathbf{x}$, and retain that $\mathbf{x}$ with the highest values. Except for perverse problems, this procedure will find the optimum.

12.2.4 SOC: The Bordered Hessian

For completeness, we present here the general SOC for an inequality constrained optimization problem. The bordered Hessian matrix now must be partitioned between binding and non-binding constraints.

For a maximization problem of a C^2 function of n variables

$$f(x_1, \dots, x_n) = f(\mathbf{x})$$

with k equality constraints denoted by h and m inequality constraints denoted by g such that they define the constraint set:

$$\begin{aligned} &g_1(\mathbf{x}) \le b_1, \dots, g_m(\mathbf{x}) \le b_m, \\ &h_1(\mathbf{x}) = c_1, \dots, h_k(\mathbf{x}) = c_k \end{aligned}$$

Form the Lagrangian

$$L(x_1, \dots, x_n, \lambda_1, \dots, \lambda_m, \mu_1, \dots, \mu_k) = f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i (g_i(\mathbf{x}) - b_i) - \sum_{j=1}^{k} \mu_j (h_j(\mathbf{x}) - c_j)$$

The FOCs:

$$\begin{aligned} &\frac{\partial L}{\partial x_1} = \cdots = \frac{\partial L}{\partial x_n} = 0, \\ &h_j(\mathbf{x}) = c_j \quad \text{for } j = 1, \dots, k, \\ &\begin{cases} \lambda_i \ge 0, \\ \lambda_i (g_i(\mathbf{x}) - b_i) = 0 \end{cases} \quad \text{(complimentary slackness)} \end{aligned}$$

Let $g_1, \dots, g_e$ be binding constraints at $\mathbf{x}^*$, and $g_{e+1}, \dots, g_m$ are non-binding constraints. The Bordered Hessian only concerns the binding constraints and thus be written as

$$H = \left(\begin{array}{cccccc|ccc}
0 & \dots & 0 & 0 & \dots & 0 & \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_1}{\partial x_n} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & \dots & 0 & 0 & \dots & 0 & \frac{\partial g_e}{\partial x_1} & \dots & \frac{\partial g_e}{\partial x_n} \\
0 & \dots & 0 & 0 & \dots & 0 & \frac{\partial h_1}{\partial x_1} & \dots & \frac{\partial h_1}{\partial x_n} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & \dots & 0 & 0 & \dots & 0 & \frac{\partial h_k}{\partial x_1} & \dots & \frac{\partial h_k}{\partial x_n} \\
\hline
\frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_e}{\partial x_1} & \frac{\partial h_1}{\partial x_1} & \dots & \frac{\partial h_k}{\partial x_1} & \frac{\partial^2 \mathcal{L}}{\partial x_1^2} & \dots & \frac{\partial^2 \mathcal{L}}{\partial x_1 \partial x_n} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\frac{\partial g_1}{\partial x_n} & \dots & \frac{\partial g_e}{\partial x_n} & \frac{\partial h_1}{\partial x_n} & \dots & \frac{\partial h_k}{\partial x_n} & \frac{\partial^2 \mathcal{L}}{\partial x_n \partial x_1} & \dots & \frac{\partial^2 \mathcal{L}}{\partial x_n^2}
\end{array}\right)$$

1. If the last $n-(e+k)$ leading principal minors alternate in sign with the sign of the determinant of the LARGEST matrix has the same sign of $(-1)^n$, then it is a MAX.
2. If the last $n-(e+k)$ leading principal minors all have the same sign as $(-1)^{e+k}$, then it is a MIN.

12.3 Nonnegativity Constraints

Most oftentimes, in economics, we want to restrict the choice variables to take nonnegative values.

Theorem 12.1 (Reduced KKT conditions for Nonnegativity)**.** *Consider the problem*

$$\begin{aligned} \max\ & f(x,y) \\ s.t.\ & g(x,y) \leq c, \\ & x \geq 0, \\ & y \geq 0 \end{aligned}$$

Rewrite the problem to

$$\begin{aligned} \max & f(x,y) \\ s.t. & g(x,y) \leq c, \\ & -x \leq 0, \\ & -y \leq 0 \end{aligned}$$

The Lagrangian is [a]

$$\mathcal{L}(x,y) = f(x,y) - \lambda[g(x,y)-c] - \mu_1(-x) - \mu_1(-y)$$

The KKT conditions

$$\begin{aligned}
&(i)\ \mathcal{L}'_x = f'_x - \lambda g'_x + \mu_1 = 0\\
&(ii)\ \mathcal{L}'_y = f'_y - \lambda g'_y + \mu_2 = 0\\
&(iii)\ \lambda \geq 0,\ \textit{with}\ \lambda = 0\ \ \textit{if}\ \ g(x,y) < c\\
&(iv)\ \mu_1 \geq 0,\ \textit{with}\ \mu_1 = 0\ \ \textit{if}\ \ x > 0\\
&(v)\ \mu_2 \geq 0,\ \textit{with}\ \mu_2 = 0\ \ \textit{if}\ \ y > 0
\end{aligned}$$

Combining (i) and (iv) yields

$$f'_x - \lambda g'_x \leq 0,\ \textbf{\textit{with equality if}}\ x > 0$$

Combining (ii) and (v) yields

$$f'_y - \lambda g'_y \leq 0,\ \textbf{\textit{with equality if}}\ y > 0$$

So the KKT conditions are reduced to just

$$\begin{aligned}
&f'_x - \lambda g'_x \leq 0,\ \textbf{\textit{with equality if}}\ x > 0\\
&f'_y - \lambda g'_y \leq 0,\ \textbf{\textit{with equality if}}\ y > 0\\
&\lambda \geq 0,\ \textit{with}\ \lambda = 0\ \ \textit{if}\ \ g(x,y) < c
\end{aligned}$$

[a]tips: You should denote the Lagrangian multipliers for the main constraint by λ and nonnegativity constraints by μ for easier handling.

Although the reduced form is introduced, it might be easier to just verify all multipliers with slackness.

Example 12.7. Consider the utility maximization problem where there are 2 goods x, y, price of good x is p and price of good y is normalized to 1, the budget is m. Find the optimal x, y.

$$\begin{aligned}
\max\ & x + \ln(1+y)\\
s.t.\ & px + y \leq m,\\
& x \geq 0\\
& y \geq 0.
\end{aligned}$$

Solutions: The Lagrangian is

$$L = x + \ln(1+y) - \lambda(px + y - m)$$

Assume the solution (x^*, y^*) exists, it must satisfy the following KKT conditions

$$\begin{aligned}
&(i) && L'_x = 1 - p\lambda \leq 0,\ \textbf{with}\ 1 - p\lambda = 0 \iff x^* > 0,\\
&(ii) && L'_y = \frac{1}{1+y^*} - \lambda \leq 0\ \textbf{with}\ \frac{1}{1+y^*} = 0 \iff y^* > 0\\
&(iii) && \lambda \cdot (px^* + y^* - m) = 0,\\
&(iv) && \lambda \geq 0, px^* + y^* \leq m
\end{aligned}$$

The objective function is concave in (x, y), the constraint is linear, thus the Lagrangian is concave, so the FOC is also sufficient.

Observe from condition (i) that λ cannot be zero, then condition (iii) implies that $\lambda > 0$ and the constraint binds such that

$$(iv)\ px^* + y^* = m$$

Regarding which constraints $x \geq 0, y \geq 0$ bind, we need to consider 4 cases

1. $x^* = 0, y^* = 0$: Since $m > 0$, this is impossible.
2. $x^* > 0, y^* = 0$: From (ii), we get $\lambda \geq 1$, then (i) implies that

   $$p = \frac{1}{\lambda} \leq 1$$

   Then from (iv), we have

   $$\begin{aligned} x^* &= m/p, \\ \lambda &= 1/p \end{aligned}$$

   if $0 < p < 1$.
3. $x^* = 0, y^* > 0$: By (iv), we have

   $$y^* = m$$

   Then (ii) yields

   $$\lambda = \frac{1}{1+y^*} = \frac{1}{1+m}$$

   Then from (i) we get the condition for this is that

   $$p \geq m + 1$$

4. $x^* > 0, y^* > 0$: With equality in both (i) and (ii), we have

   $$\lambda = 1/p = 1/(1 + y^*)$$

   It follows that

   $$\begin{aligned} &y^* = p - 1, \\ &p > 1 \text{ (because } y^* > 0) \end{aligned}$$

   Equation (iv) yields

   $$\begin{aligned} &x^* = \frac{m + 1 - p}{p}, \\ &p < m + 1 \text{ (because } x^* > 0) \end{aligned}$$

In summary

1. If $0 < p \leq 1$, then $(x^*, y^*) = (m/p, 0)$ with $\lambda = 1/p$
2. if $1 < p < m + 1$, then $(x^*, y^*) = (\frac{m+1-p}{p}, p - 1)$ with $\lambda = 1/p$
3. if $p \geq m + 1$, then $(x^*, y^*) = (0, m)$ with $\lambda = 1/(1 + m)$

In the 2 extreme cases (1) and (3), it is optimal to spend everything on only the cheaper good – x in case (1) and y in case (3).

Ex. 12.8. Solve Example 12.7 using the formal way (without employing the reduced form).

Example 12.8. We deal with a minimum optimization problem with nonnegativity constraints from Example 12.5. We formalize it as

$$\begin{aligned}
&\min f(x,y) = 2y - x^2,\\
&s.t. x^2 + y^2 \leq 1,\\
&\quad x \geq 0,\\
&\quad y \geq 0.
\end{aligned}$$

Rewrite it as

$$\begin{aligned}
&\min f(x,y) = 2y - x^2,\\
&s.t. - x^2 - y^2 \geq -1,\\
&\quad x \geq 0,\\
&\quad y \geq 0.
\end{aligned}$$

The Lagrangian

$$L(x, y, \lambda, \mu_1, \mu_2) = 2y - x^2 - \lambda(-x^2 - y^2 + 1) - \mu_1 x - \mu_2 y$$

KKT FOCs are

$$\begin{aligned}
&(i)\ \frac{\partial L}{\partial x} = 0 \iff -2x + 2\lambda x - \mu_1 = 0\\
&(ii)\ \frac{\partial L}{\partial y} = 0 \iff 2 + 2\lambda y - \mu_2 = 0\\
&(iii)\ \lambda \cdot (-x^2 - y^2 + 1) = 0,\\
&(iv)\ \mu_1 x = 0,\\
&(v)\ \mu_2 y = 0,\\
&(a)\ \lambda \geq 0 \text{ (if } \lambda > 0\text{, constraint binds)},\\
&(b)\ \mu_1 \geq 0 \text{ (if } \mu_1 = 0, x > 0\text{, otherwise } x = 0 \text{)},\\
&(c)\ \mu_2 \geq 0 \text{ (if } \mu_2 = 0, y > 0\text{, otherwise } y = 0 \text{)}
\end{aligned}$$

From $(i), (ii)$ we have

$$\begin{aligned}
&(i)\ 2x + \mu_1 = 2\lambda x,\\
&(ii)\ 2 + 2\lambda y = \mu_2
\end{aligned}$$

Since λ, y are nonnegative, from (ii) we see that $\mu_2 > 0$. Because of that, from (c), we conclude that $y = 0$, and therefore $\mu_2 = 2$.

If $x = 0$, then $\mu_1 = 0$ from (i), which contradicts (b).

If $x > 0$, then $\mu_1 = 0$ from (b), which implies that $\lambda = 1$ from (i) where $\mu_1 = 2x(\lambda - 1)$. Since $\lambda > 0$, the constraint binds (iii) such that $x^2 + y^2 = 1$. And since $y = 0, x > 0$, we conclude that $x = 1$.

Therefore, the solution is $x^* = 1, y^* = 0$, which differs from Example 12.5.

12.4 The Meaning of the Lagrangian Multiplier

Consider again the problem

$$\max(\min)\ f(x,y) \quad s.t. \quad g(x,y) = c$$

The FOCs

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial x_1} &= f'_x - \lambda g'_x = 0 \iff f'_x = \lambda g' x, \\
\frac{\partial \mathcal{L}}{\partial x_y} &= f'_y - \lambda g'_y = 0 \iff f'_y = \lambda g'_y
\end{aligned}$$

Suppose that x^*, y^* are the values of x, y that solve the problem. In general, x^* and y^* depend on c, so we write $x^* = x(c)$ and $y^* = y(c)$. We assume that these solutions are differentiable functions of c. The associated value of $f(x, y)$ is then also a function of c, with

$$f^*(c) = f(x^*(c),\ y^*(c))$$

Here $f^*(c)$ is called the (optimal) value function for the problem. Of course, the associated value of the Lagrange multiplier also depends on c, in general, so we write $\lambda(c)$. Taking the total differentiation of the value function yields

$$\begin{aligned}
df^*(c) &= f'_x(x^*, y^*)dx^* + f_y(x^*, y^*)dy^* \\
&= \lambda(c)g'_x(x^*, y^*)dx^* + \lambda(c)g'_y(x^*, y^*)dy^* \\
&= \lambda(c)\underbrace{\left(g'_x(x^*, y^*)dx^* + g'_y(x^*, y^*)dy^*\right)}_{dg(x^*, y^*)}
\end{aligned}$$

By identity $g(x, y) = c$ so $dg \equiv dc$, which implies that

$$df^*(c) = \lambda(c)\ dc$$

Thus, we have a remarkable result that

$$\frac{df^*(c)}{dc} = \lambda(c)$$

Thus, the Lagrange multiplier $\lambda = \lambda(c)$ is the rate at which the optimal value of the objective function changes with respect to changes in the constraint constant c. In particular, if dc is a small change in c, then

$$f^*(c + dc) - f^*(c) \approx \lambda(c)dc$$

In economic applications, c often denotes the available stock of some resource, and $f(x, y)$ denotes utility or profit. Then $\lambda(c)dc$ measures the approximate change in utility or profit that can be obtained from dc units more (or $-dc$ unit less when $dc < 0$). That is why economists call λ a **shadow price** of the resource constraint.

Notice the KKT problems, when the constraint is not binding, the Lagrangian multiplier $\lambda = 0$, or the shadow price in this case is 0. What it means is that the maximal value $f(x^*, y^*)$ does not not change when c changes a little.

12.5 Review Exercises

Ex. 12.9. The following problems are borrowed from Kadeishvili[2]
Compare the solutions from different optimization problems.

1. Maximization problem of
   - $f(x, y) = 10 - x^2 - y^2$
   - $f(x, y) = 10 - x^2 - y^2$ s.t. $h(x, y) = 2x^2 + y^2 = 2$.
   - $f(x, y) = 10 - x^2 - y^2$ s.t. $h(x, y) = 2x^2 + y^2 \leq 2$.
   - $f(x, y) = 10 - x^2 - y^2$ s.t. $h(x, y) = 2x^2 + y^2 \geq 2$.
2. Minimization problem of
   - $f(x, y) = 5 + x^2 + y^2$
   - $f(x, y) = 5 + x^2 + y^2$ s.t. $h(x, y) = 2x^2 + y^2 = 2$
   - $f(x, y) = 5 + x^2 + y^2$ s.t. $h(x, y) = 2x^2 + y^2 \leq 2$
   - $f(x, y) = 5 + x^2 + y^2$ s.t. $h(x, y) = 2x^2 + y^2 \geq 2$

Ex. 12.10. Solve the following problems with KKT necessary conditions and verify the solutions using sufficient conditions.

1. $\max_{x,y} f(x, y) = xy$ subject to $x^2 + y^2 \leq 1$.
2. $\max_{x,y} f(x, y) = x^2 + 3y^2$ subject to $2x + y \leq 4$.
3. $\max_{x,y} f(x, y) = x^2 + 2xy + y^2$ subject to $x + y \leq 2$.
4. $\max_{x,y} f(x, y) = 2x^2 + y^2$ subject to $x + y \leq 3$.
5. $\max_{x,y} f(x, y) = x^3 + 3xy^2$ subject to $x^2 + y^2 \leq 4$.
6. $\max_{x,y} f(x, y) = x^2 + 2xy + y^2$ subject to $x + y \geq 2$.
7. $\max_{x,y} f(x, y) = x^3 + 3xy^2$ subject to $x + y \leq 3$.
8. $\max_{x,y} f(x, y) = x^2 + 2y^2$ subject to $x + 2y \leq 3$.
9. $\max_{x,y} f(x, y) = x^3 + 2xy + y^3$ subject to $x + y \leq 2$.
10. $\max_{x,y} f(x, y) = 2x^2 + 3y^2$ subject to $x^2 + y^2 \leq 1/2$.

Ex. 12.11. Solve the problems for 3 variables using KKT

1. $\min_{x,y,z} f(x, y, z) = -x^3 + y^2 - 2xz^2$ subject to $2x + y^2 + z - 5 = 0$, $5x - y^2 - z \geq 2$, and x, y, z all ≥ 0. Is (1,0,3) a solution?
2. $\min_{x,y,z} = x^2 + y^2 + z^2$ subject to $-x + y - z \geq -10$, and $x + y + 4z \geq 20$. Find all solutions.

Ex. 12.12. Solve the following problems in Simon and Blume (1994).

1. $\max f(x, y) = x^2 + y^2$ s.t. $2x + y \leq 2$, and $x > 0, y > 0$.
2. $\max f(x, y) = 2y^2 - x$ s.t. $x^2 + y^2 \leq 1$, and $x \geq 0, y \geq 0$.
3. $\max f(x, y) = 3xy - x^3$ s.t. $2x - y = -5$, $5x + 2y \geq 37$, and $x \geq 0, y \geq 0$.

---

[2] https://rmi.tsu.ge/ kade/LecturesT.Kadeishvili/MathEconomics/Term4/Week3Lagrangeineq.pdf.