Mathematical Methods for Economic Dynamics (ECON 696)

Lecture 4: Multivariate Calculus II (Ch. 2) and Static Optimization (Ch. 3)

Characterization of Quasiconcavity via First Derivatives (Ch. 2.5)

Theorem: Let $f : S \to \mathbb{R}$ be a C^1 function. Then $f(\cdot)$ is quasiconcave on S if and only if for all $x, x^0 \in S$,

$$f(x) \geq f(x^0) \Rightarrow \nabla f(x^0) \cdot (x - x^0) = \sum_{i=1}^{n} \frac{\partial f(x^0)}{\partial x_i}(x_i - x_i^0) \geq 0.$$

Proof: We skip the proof. ■

The content of the above theorem is that for any quasiconcave function $f(\cdot)$ and any pair of points x and x^0 with $f(x) \geq f(x^0)$, the gradient vector $\nabla f(x^0)$ and the vector $(x - x^0)$ must form an acute angle.

A Determinant Criterion for Quasiconcavity (Ch. 2.5)

Theorem: Let $S \subseteq \mathbb{R}^2$ be an open, convex set and $f : S \to \mathbb{R}$ be a C^2 function. Define the **bordered Hessian determinant**

$$B_2(x,y) = \begin{vmatrix} 0 & f_1'(x,y) & f_2'(x,y) \\ f_1'(x,y) & f_{11}''(x,y) & f_{12}''(x,y) \\ f_2'(x,y) & f_{21}''(x,y) & f_{22}''(x,y) \end{vmatrix}.$$

1. A necessary condition for f to be **quasiconcave** in S is that $B_2(x,y) \geq 0$ for all $(x,y) \in S$.

2. A sufficient condition for f to be **strictly quasiconcave** in S is that $f_1'(x,y) \neq 0$ and $B_2(x,y) > 0$ for all $(x,y) \in S$.

Proof: We omit the proof. ■

We move on to the general case. Define the bordered Hessian determinants

$$B_r(\mathbf{x}) = \begin{vmatrix} 0 & f_1'(\mathbf{x}) & \cdots & f_r'(\mathbf{x}) \\ f_1'(\mathbf{x}) & f_{11}''(\mathbf{x}) & \cdots & f_{1r}''(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ f_r'(\mathbf{x}) & f_{r1}''(\mathbf{x}) & \cdots & f_{rr}''(\mathbf{x}) \end{vmatrix}.$$

for $r = 1, \ldots, n$.

Theorem: Let $S \subseteq \mathbb{R}^n$ be an open, convex set and $f : S \to \mathbb{R}$ be a C^2 function. Then,

1. A necessary condition for f to be **quasiconcave** is that $(-1)^r B_r(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in S$ and all $r = 1, \ldots, n$.

2. A sufficient condition for f to be **strictly quasiconcave** is that $(-1)^r B_r(\mathbf{x}) > 0$ for all $\mathbf{x} \in S$ and all $r = 1, \ldots, n$.

Proof: We omit the proof. ∎

Unconstrained Optimization (Ch. 3.1, 3.2)

An **optimization problem** is one where the values of a given function $f : \mathbb{R}^n \to \mathbb{R}$ are to be maximized or minimized over a given set $S \subseteq \mathbb{R}^n$.

This $f(\cdot)$ is called the **objective function** and the set S is called the **constraint set**.

Objectives of Optimization Theory

(1) we identify a set of conditions on $f(\cdot)$ and S under which the **existence** of the solutions to optimization problem is guaranteed. This is already achieved by Weierstrass (the extreme value) theorem.

(2) we obtain various **receipes** for finding optimal points.

Extreme Points (Ch. 3.1)

Suppose that the point $x^* = (x_1^*, \ldots, x_n^*)$ belongs to S and

$$f(x^*) \geq f(x) \text{ for all } x \in S \ (*)$$

Then, x^* is called a (global) **maximum point** for $f(\cdot)$ in S and $f(x^*)$ is called the **maximum value**.

If the inequality $(*)$ is strict for all $x \neq x^*$, then x^* is a **strict** maximum point for $f(\cdot)$ in S.

We can define **minimum point** and **minimum value** by reversing the inequality sign in $(*)$.

As collective names, we use **extreme points** and **extreme values** to indicate both maxima or minima.

A Necessary Condition for Extreme Points

Theorem: Let $f(\cdot)$ be defined on a set S in $\mathbb{R}^n$ and let $x^* = (x_1^*, \ldots, x_n^*)$ be an interior point in S at which $f(\cdot)$ has partial derivatives. A necessary condition for x^* to be an extreme point for f is that x^* is a **stationary point** for $f(\cdot)$ – that is, it satisfies the equations

$$\nabla f(x) = \mathbf{0} \Longleftrightarrow \frac{\partial f(x)}{\partial x_i} = 0, \text{ for } i = 1, \ldots, n$$

Proof: Suppose, on the contrary, that x^* is a maximum point but not a stationary point for $f(\cdot)$. Then, there is no loss of generality to assume that there exists at least i such that $\partial f(x^*)/\partial x_i > 0$. Define $x^{**} = (x_1^*, \ldots, x_i^* + \varepsilon, \ldots, x_n^*)$. Since x^* is an interior point in S, one can make sure that $x^{**} \in S$ by choosing $\varepsilon > 0$ sufficiently small. Then,

$$f(x^{**}) \approx f(x^*) + \nabla f(x) \cdot (0, \ldots, 0, \underbrace{\varepsilon}_{i}, 0, \ldots, 0) > f(x^*).$$

However, this contradicts the hypothesis that x^* is a maximum point for $f(\cdot)$. ■

Theorem: Suppose that the function $f(\cdot)$ is defined in a convex set $S \subseteq \mathbb{R}^n$ and let x^* be an **interior** point of S. Assume that $f(\cdot)$ is C^1 in a ball around x^*.

1. If $f(\cdot)$ is concave in S, then x^* is a (global) maximum point for $f(\cdot)$ in S if and only if x^* is a stationary point for $f(\cdot)$.

2. $f(\cdot)$ is convex in S, then x^* is a (global) minimum point for $f(\cdot)$ in S if and only if x^* is a stationary point for $f(\cdot)$.

Proof: We focus on the first part of the theorem. The second part follows once we take into account that $-f$ is concave. ($\Rightarrow$) This follows from the previous theorem.

($\Leftarrow$) Suppose that x^* is a stationary point for $f(\cdot)$ and that $f(\cdot)$ is concave. We use the following characterization of concave functions:

Theorem: If $f : S \to \mathbb{R}$ is concave, for any $x, x^{'} \in S$,

$$f(x^{'}) - f(x) \leq \nabla f(x) \cdot (x^{'} - x).$$

Setting $x^{'} = x$ and $x = x^*$,

$$f(x) - f(x^*) \leq \nabla f(x^*) \cdot (x - x^*) = 0 \ (\because \nabla f(x^*) = 0)$$

Thus, we have $f(x) \leq f(x^*)$ for any $x \in S$ as desired. ■

Local Extreme Points (Ch. 3.2)

$x^* \in S$ is a **local maximum point** of $f(\cdot)$ in S if there exists an $\varepsilon > 0$ such that $f(x) \leq f(x^*)$ for all $x \in B_\varepsilon(x^*) \cap S$.

If x^* is the unique local maximum point for $f(\cdot)$, then it is a **strict local maximum point** for $f(\cdot)$ in S.

A **(strict) local minimum point** is defined in the obvious way, and the meaning of the following terms should be clear: **local maximum and minimum values, local extreme points**, and **local extreme values**.

A stationary point x^* of $f(\cdot)$ that is neither a local maximum point nor a local minimum point is called a **saddle point** of $f(\cdot)$.

Theorem (Sufficient Conditions for Local Extreme Points): Suppose that $f(x) = f(x_1, \ldots, x_n)$ is defined on a set $S \subseteq \mathbb{R}^n$ and that $x^* \in S$ is an interior stationary point. Assume also that $f(\cdot)$ is C^2 in an open ball around x^*. Then,

1. $D^2 f(x^*)$ is positive definite $\Rightarrow x^*$ is a local minimum point.

2. $D^2 f(x^*)$ is negative definite $\Rightarrow x^*$ is a local maximum point.

Proof: We skip the proof. ■

The next lemma establishes a sufficient condition for saddle points.

Lemma: If x^* is an interior stationary point of $f(\cdot)$ such that $|D^2 f(x^*)| \neq 0$ and $D^2 f(x^*)$ is neither positive definite nor negative definite, then x^* is a saddle point.

Proof: We skip the proof. ■

Theorem (Necessary Conditions for Local Extreme Points): Suppose that $f(x) = f(x_1, \ldots, x_n)$ is defined on a set $S \subseteq \mathbb{R}^n$, and x^* is an interior stationary point in S. Assume that f is C^2 in a ball around x^*. Then,

1. x^* is a local minimum point $\Rightarrow D^2 f(x^*)$ is positive semidefinite.

2. x^* is a local maximum point $\Rightarrow D^2 f(x^*)$ is negative semidefinite.

Proof: We skip the proof. ■

Constrained Optimization: Equality Constraints (Ch. 3.3, 3.4)

Equality Constraints as a Tangent Hyperplane

If the constraints do bite at an optimum x, we need to have some knowledge of what the constraint set looks like in a neighborhood of x in order to characterize the behavior of the objective function $f(\cdot)$ around x.

A set of equality constraints in $\mathbb{R}^n$, $g(x) = \mathbf{0}$, i.e.,

$$\begin{aligned} g^1(x) &= 0 \\ g^2(x) &= 0 \\ &\vdots \\ g^m(x) &= 0 \end{aligned}$$

defines a subset of $\mathbb{R}^n$ which is best viewed as a hypersurface.

We write

$$\underbrace{Dg(x)}_{m \times n \text{ matrix}} = \begin{pmatrix} \nabla g^1(x) \\ \nabla g^2(x) \\ \vdots \\ \nabla g^m(x) \end{pmatrix} = \begin{pmatrix} \frac{\partial g^1(x)}{\partial x_1} & \frac{\partial g^1(x)}{\partial x_2} & \cdots & \frac{\partial g^1(x)}{\partial x_n} \\ \frac{\partial g^2(x)}{\partial x_1} & \frac{\partial g^2(x)}{\partial x_2} & \cdots & \frac{\partial g^2(x)}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g^m(x)}{\partial x_1} & \frac{\partial g^m(x)}{\partial x_2} & \cdots & \frac{\partial g^m(x)}{\partial x_n} \end{pmatrix}.$$

Equality Constraints: The Lagrange Problem (Ch. 3.3)

Let $S \subseteq \mathbb{R}^n$. A general maximization problem with equality constraints is of the form

$$\max_{x=(x_1,\ldots,x_n)\in S} f(x_1,\ldots,x_n) \text{ s.t. } g^j(x) = 0\ \forall j = 1,\ldots,m\ (m < n)\ (*)$$

We say $x \in S$ is **feasible** if $g^j(x) = 0$ for each $j \in \{1,\ldots,m\}$.

Define the **Lagrangian**,

$$\mathcal{L}(x,\lambda) = f(x) - \lambda_1 g^1(x) - \cdots - \lambda_m g^m(x)$$

where $\lambda_1, \ldots, \lambda_m$ are called **Lagrange multipliers**.

The necessary first-order conditions for optimality are then:

$$\nabla \mathcal{L}(x,\lambda) = \nabla f(x) - \sum_{j=1}^{m} \lambda_j \nabla g^j(x) = \mathbf{0}$$

$$\iff \frac{\partial \mathcal{L}(x)}{\partial x_i} = \frac{\partial f(x)}{\partial x_i} - \sum_{j=1}^{m} \lambda_j \frac{\partial g^j(x)}{\partial x_i} = 0, \ \forall i = 1, \ldots, n \ (**)$$

Theorem (Conditions for Extreme Points with Equality Constraints):

(**Necessity**) Suppose that the functions $f(\cdot)$ and $g^1(\cdot), \ldots, g^m(\cdot)$ are defined on a set S in $\mathbb{R}^n$ and $x^* = (x_1^*, \ldots, x_n^*)$ is an interior point of S that solves the maximization problem $(*)$. Assume further that $f(\cdot)$ and $g^1(\cdot), \ldots, g^m(\cdot)$ are C^1 in an open ball around x^*, and that rank$(Dg(x^*)) = m$. Then, there exist unique numbers $\lambda_1, \ldots, \lambda_m$ such that the first-order conditions $(**)$ are valid.

(**Sufficiency**) Suppose that there exist numbers $\lambda_1, \ldots, \lambda_m$ and a feasible x^* which together satisfy the first-order conditions $(**)$. Then, if the Lagrangian $\mathcal{L}(x, \lambda_1, \ldots, \lambda_m)$ is concave in x, then x^* solves the maximization problem $(*)$.

Proof: (**Necessity**) We content ourselves with a heuristic argument based on the simplest formulation.

Consider

$$\max_{(x,y)\in\mathbb{R}^2} f(x,y) \text{ subject to } g(x,y)=c.$$

Let (x^*,y^*) be a local maximum point of f to the above constrained optimization problem. So, we must have $g(x^*,y^*)=c$.

We consider a pair of "small" numbers $(\Delta x,\Delta y)\in\mathbb{R}^2$ such that $g(x^*+\Delta x,y^*+\Delta y)=g(x^*,y^*)$. Then, we have

$$\begin{aligned}\Delta g \quad &= \quad g(x^*+\Delta x,y^*+\Delta y)-g(x^*,y^*)\\ &\underbrace{\approx}_{\text{linear approx}} \quad g_1'(x^*,y^*)\Delta x+g_2'(x^*,y^*)\Delta y=0.\end{aligned}$$

Assuming that $g_1'(x^*, y^*) \neq 0$ (i.e., rank$(Dg(x^*, y^*)) = m$), we derive

$$\Delta x = -\frac{g_2'(x^*, y^*)}{g_1'(x^*, y^*)}\Delta y. \quad (*)$$

Since (x^*, y^*) is a local maximum point to the constrained optimization problem,

$$\begin{aligned}
0 &\geq f(x^* + \Delta x, y^* + \Delta y) - f(x^*, y^*) \\
&\underbrace{\approx}_{\text{linear approx}} f_1'(x^*, y^*)\Delta x + f_2'(x^*, y^*)\Delta y \\
&= \left(-\frac{f_1'(x^*, y^*)}{g_1'(x^*, y^*)}g_2'(x^*, y^*) + f_2'(x^*, y^*)\right)\Delta y. \quad (\because \ (*))
\end{aligned}$$

Since Δy could be positive or negative, we must have

$$-\frac{f_1'(x^*, y^*)}{g_1'(x^*, y^*)} g_2'(x^*, y^*) + f_2'(x^*, y^*) = 0. \qquad (**)$$

Define

$$\lambda^* \equiv \frac{f_1'(x^*, y^*)}{g_1'(x^*, y^*)}.$$

Then, $(**)$ can be translated into:

$$\begin{aligned} f_1'(x^*, y^*) &= \lambda^* g_1'(x^*, y^*), \\ f_2'(x^*, y^*) &= \lambda^* g_2'(x^*, y^*). \end{aligned}$$

(**Sufficiency**) Suppose that the Lagrangian $\mathcal{L}(x, \lambda)$ is concave in x. The first-order necessary conditions imply that x^* is a stationary point of the Lagrangian. Then, by the sufficiency result for unconstrained maximization with $\mathcal{L}(x, \lambda)$ being the objective function,

$$\mathcal{L}(x^*, \lambda) = f(x^*) - \sum_{j=1}^{m} \lambda_j g^j(x^*) \geq f(x) - \sum_{j=1}^{m} \lambda_j g^j(x) = \mathcal{L}(x, \lambda) \ \forall x \in S$$

But for all feasible x, we have $g^j(x) = 0$ and of course, $g^j(x^*) = 0$ for all $j = 1, \ldots, m$. This implies that $f(x^*) \geq f(x)$ for all feasible x. Thus, x^* solves the maximization problem $(*)$. ■

Example:

$$\max_{x,y,z} f(x,y,z) = x+2z \text{ subject to } \begin{cases} g^1(x,y,z) = x+y+z = 1 \\ g^2(x,y,z) = x^2+y^2+z = 7/4 \end{cases}$$

We setup the Lagrangian:

$$\mathcal{L}(x,y,z) = x + 2z - \lambda_1(x+y+z-1) - \lambda_2(x^2+y^2+z-7/4).$$

The first-order conditions are:

$$\begin{aligned} \mathcal{L}_1^{'} &= 1 - \lambda_1 - 2\lambda_2 x = 0 \quad \text{(i)} \\ \mathcal{L}_2^{'} &= -\lambda_1 - 2\lambda_2 y = 0 \quad \text{(ii)} \\ \mathcal{L}_3^{'} &= 2 - \lambda_1 - \lambda_2 = 0 \quad \text{(iii)} \end{aligned}$$

(iii) $\Rightarrow \lambda_2 = 2 - \lambda_1$

Plugging $\lambda_2 = 2 - \lambda_1$ into (ii), we obtain

$$-\lambda_1 - 4y + 2\lambda_1 y = 0 \Leftrightarrow \lambda_1(2y-1) = 4y \Rightarrow y \neq 1/2.$$

So,

$$\lambda_1 = \frac{4y}{2y-1}.$$

Plugging $\lambda_1 = 4y/(2y-1)$ and $\lambda_2 = 2 - \lambda_1$ into (i), we obtain

$$y = 2x - 1/2.$$

Plugging $y = 2x - 1/2$ into the constraints, we translate the constraint equalities into:

$$3x + z = 3/2 \text{ and } 5x^2 - 2x + z = 3/2.$$

$3x + z = 3/2 \Rightarrow z = -3x + 3/2.$

Plugging $z = -3x + 3/2$ into $5x^2 - 2x + z = 3/2$, we obtain

$$5x(x - 1) = 0 \Rightarrow x = 0 \text{ or } 1.$$

Case 1: $x = 0$

We obtain $y = -1/2$ and $z = 3/2$. In this case,

$$f(0, -1/2, 3/2) = 3.$$

Case 2: $x = 1$

We obtain $y = 3/2$ and $z = -3/2$. In this case,

$$f(1, 3/2, -3/2) = -2.$$

Hence, $(x, y, z) = (0, -1/2, 3/2)$ is the only possible candidate for the solution. And the associated Lagrange multipliers are $\lambda_1 = \lambda_2 = 1$.

When $\lambda_1 = \lambda_2 = 1$,

$$\mathcal{L}(x, y, z) = -x^2 - y^2 - y + 11/4.$$

The associated Hessian matrix is

$$D^2\mathcal{L} = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

We check its principal minors: $\Delta^2_{(1)}\mathcal{L} = -2, -2, 0$;

$$\Delta^2_{(2)}\mathcal{L} = \begin{vmatrix} -2 & 0 \\ 0 & 0 \end{vmatrix} = 0; \begin{vmatrix} -2 & 0 \\ 0 & 0 \end{vmatrix} = 0; \begin{vmatrix} -2 & 0 \\ 0 & -2 \end{vmatrix} = 4.$$

and $\Delta^2_{(3)}\mathcal{L} = 0$. Therefore, $\mathcal{L}$ is a concave function given $\lambda_1 = \lambda_2 = 1$. By the sufficiency result of the Lagrangian approach, we confirm $(0, -1/2, 3/2)$ is the solution.

Sufficiency for Local Extreme Points (Ch. 3.4)

$$\max_{x \in S} f(x) \quad \text{subject to} \quad g^j(x) = 0, j = 1, \ldots, m \ (m < n).$$

In general, we define the bordered Hessian determinants, for $r = m + 1, \ldots, n$:

$$B_r(x^*) = \begin{vmatrix} 0 & \cdots & 0 & \frac{\partial g^1(x^*)}{\partial x_1} & \cdots & \frac{\partial g^1(x^*)}{\partial x_r} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \frac{\partial g^m(x^*)}{\partial x_1} & \cdots & \frac{\partial g^m(x^*)}{\partial x_r} \\ \frac{\partial g^1(x^*)}{\partial x_1} & \cdots & \frac{\partial g^m(x^*)}{\partial x_1} & \mathcal{L}''_{11}(x^*) & \cdots & \mathcal{L}''_{1r}(x^*) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g^1(x^*)}{\partial x_r} & \cdots & \frac{\partial g^m(x^*)}{\partial x_r} & \mathcal{L}''_{r1}(x^*) & \cdots & \mathcal{L}''_{rr}(x^*) \end{vmatrix}$$

The determinant B_r is the $(m+r)$th leading principal minor of $(m+n)\times(m+n)$ bordered matrix

$$\begin{pmatrix} \mathbf{0}_{m\times m} & \underbrace{Dg(x^*)}_{m\times n} \\ \underbrace{(Dg(x^*))^T}_{n\times m} & \underbrace{D^2\mathcal{L}(x^*)}_{n\times n} \end{pmatrix} = \begin{pmatrix} 0 & \cdots & 0 & \frac{\partial g^1(x^*)}{\partial x_1} & \cdots & \frac{\partial g^1(x^*)}{\partial x_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \frac{\partial g^m(x^*)}{\partial x_1} & \cdots & \frac{\partial g^m(x^*)}{\partial x_n} \\ \frac{\partial g^1(x^*)}{\partial x_1} & \cdots & \frac{\partial g^m(x^*)}{\partial x_1} & \mathcal{L}''_{11}(x^*) & \cdots & \mathcal{L}''_{1n}(x^*) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g^1(x^*)}{\partial x_n} & \cdots & \frac{\partial g^m(x^*)}{\partial x_n} & \mathcal{L}''_{n1}(x^*) & \cdots & \mathcal{L}''_{nn}(x^*) \end{pmatrix}.$$

Theorem (Sufficiency for Local Maximum): Suppose there is a point $x^* \in S \subseteq \mathbb{R}^n$ satisfying $g(x^*) = \mathbf{0}$, and a $\lambda = (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m$ such that

$$\nabla f(x^*) - \sum_{j=1}^{m} \lambda_j \nabla g^j(x^*) = \underbrace{\mathbf{0}}_{n\times 1} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

$(-1)^r B_r(x^*) > 0$ for $r = m+1, \ldots, n \Rightarrow x^*$ is the local maximum to $(*)$.

$(-1)^m B_r(x^*) > 0$ for $r = m+1, \ldots, n \Rightarrow x^*$ is the local minimum to $(*)$.

Proof: We skip the proof. ■

Example:

$$\max_{x,y,z} f(x,y,z) = x^2+y^2+z^2 \text{ subject to } \begin{cases} g^1(x,y,z) = x + 2y + z = 30 \\ g^2(x,y,z) = 2x - y - 3z = 10 \end{cases}$$

We setup the Lagrangian:

$$\mathcal{L}(x,y,z) = x^2+y^2+z^2-\lambda_1(x+2y+z-30)-\lambda_2(2x-y-3z-10).$$

The first-order conditions are:

$$\begin{aligned} \mathcal{L}_1' &= 2x - \lambda_1 - 2\lambda_2 = 0; \\ \mathcal{L}_2' &= 2y - 2\lambda_1 + \lambda_2 = 0; \\ \mathcal{L}_3' &= 2z - \lambda_1 + 3\lambda_2 = 0. \end{aligned}$$

Considering two equality constraints, the unique solution to the first-order conditions is $(x,y,z) = (10,10,0)$. The associated Lagrange multipliers are $\lambda_1 = 12$ and $\lambda_2 = 4$.

We compute the Bordered Hessian determinant B_3:

$$B_3(x,y,z) = \begin{vmatrix} 0 & 0 & \partial g^1/\partial x & \partial g^1/\partial y & \partial g^1/\partial z \\ 0 & 0 & \partial g^2/\partial x & \partial g^2/\partial y & \partial g^2/\partial z \\ \partial g^1/\partial x & \partial g^2/\partial x & \mathcal{L}''_{xx} & \mathcal{L}''_{xy} & \mathcal{L}''_{xz} \\ \partial g^1/\partial y & \partial g^2/\partial y & \mathcal{L}''_{yx} & \mathcal{L}''_{yy} & \mathcal{L}''_{yz} \\ \partial g^1/\partial z & \partial g^2/\partial z & \mathcal{L}''_{zx} & \mathcal{L}''_{zy} & \mathcal{L}''_{zz} \end{vmatrix}$$

$$B_3(10,10,0) = \begin{vmatrix} 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 2 & -1 & -3 \\ 1 & 2 & 2 & 0 & 0 \\ 2 & -1 & 0 & 2 & 0 \\ 1 & -3 & 0 & 0 & 2 \end{vmatrix} = 150 > 0$$

This implies $(-1)^2 B_3(10,10,0) > 0$. Hence, $(10,10,0)$ is the local minimum point.

Constrained Optimization: Inequality Constraints (Ch. 3.5)

Consider the following optimization problem with inequality constraints.

$$\max_{x \in S} f(x) \text{ subject to } \begin{cases} g^1(x_1, \ldots, x_n) \leq 0 \\ g^2(x_1, \ldots, x_n) \leq 0 \\ \qquad \vdots \\ g^m(x_1, \ldots, x_n) \leq 0 \end{cases}$$

A vector $x = (x_1, \ldots, x_n)$ that satisfies all the constraints is called **feasible** (or admissible).

Throughout we assume that $f(\cdot)$ and all the $g^j(\cdot)$ are C^1 functions.

An inequality constraint $g^j(x) \leq 0$ is said to be **binding (active)** at x if $g^j(x) = 0$ and **non-binding (inactive)** at x if $g^j(x) < 0$.

Minimizing $f(x)$ is equivalent to maximizing $-f(x)$. Moreover, $g^j(x) \geq 0$ can be rewritten as $-g^j(x) \leq 0$. In this way, most constrained optimization problems can be expressed as the above form.

We define the Lagrangian exactly the same as before.

$$\mathcal{L}(x) = f(x) - \lambda \cdot g(x) = f(x) - \sum_{j=1}^{m} \lambda_j g^j(x),$$

where $\lambda = (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m$ are the Lagrangian multipliers.

Again, the first-order partial derivatives of the Lagrangian are equated to 0:

$$\nabla \mathcal{L}(x) = \nabla f(x) - \sum_{j=1}^{m} \lambda_j \nabla g^j(x) = \mathbf{0};$$

$$\Leftrightarrow \quad \frac{\partial \mathcal{L}(x)}{\partial x_i} = \frac{\partial f(x)}{\partial x_i} - \sum_{j=1}^{m} \lambda_j \frac{\partial g^j(x)}{\partial x_i} = 0, \ \forall i = 1, \ldots, n \ (KT-1)$$

In addition, we introduce the **complementary slackness conditions**. For all $j = 1, \ldots, m$,

$$\lambda_j \geq 0 \text{ and } \lambda_j = 0 \text{ if } g^j(x) < 0 \text{ (KT-2)}$$

An alternative formulation of KT-2 is that for any $j = 1, \ldots, m$,

$$\lambda_j \geq 0 \text{ and } \lambda_j g^j(x) = 0.$$

In particular, if $\lambda_j > 0$, one must have $g^j(x) = 0$. However, it is perfectly possible to have both $\lambda_j = 0$ and $g^j(x) = 0$.

Conditions (KT-1) and (KT-2) together are often called the **Kuhn-Tucker conditions**.

Suppose one can find a point x^* at which $f(\cdot)$ is stationary and $g^j(x^*) < 0$ for all $j = 1, \ldots, m$. Then, the Kuhn-Tucker conditions will automatically be satisfied by x^* together with all the Lagrangian multipliers $\lambda_j = 0$ for all $j = 1, \ldots, m$.

Theorem (Sufficiency for the Kuhn-Tucker Conditions I): Consider the maximization problem and suppose that x^* is feasible and satisfies conditions (KT-1) and (KT-2). If the Lagrangian $\mathcal{L}(x) = f(x) - \lambda \cdot g(x)$ (with the λ values obtained from solving KT-1 and KT-2) is concave, then x^* is a solution to the maximization problem.

Proof: Since $\mathcal{L}(x, \lambda)$ is concave by assumption and $\nabla\mathcal{L}(x^*) = \mathbf{0}$ from $(KT - 1)$, by the sufficiency result on unconstrained optimization, we have that for all $x \in S$,

$$f(x^*) - \sum_{j=1}^{m} \lambda_j g^j(x^*) \geq f(x) - \sum_{j=1}^{m} \lambda_j g_j(x)$$

Rearranging gives the equivalent inequality:

$$f(x^*) - f(x) \geq \sum_{j=1}^{m} \lambda_j \left(g^j(x^*) - g^j(x) \right).$$

Thus, it suffices to show

$$\sum_{j=1}^{m} \lambda_j \left(g^j(x^*) - g^j(x) \right) \geq 0$$

for every feasible x, because this will imply that x^* solves the maximization problem.

Suppose that $g^j(x^*) < 0$. Then (KT-2) shows that $\lambda_j = 0$.

Suppose that $g^j(x^*) = 0$, we have $\lambda_j(g^j(x^*) - g^j(x)) = -\lambda_j g^j(x) \geq 0$ because x is feasible, i.e., $g^j(x) \leq 0$ and $\lambda_j \geq 0$. Then, we have $\sum_{j=1}^{m} \lambda_j \left(g^j(x^*) - g^j(x) \right) \geq 0$ as desired. ■

Theorem (Sufficiency for the Kuhn-Tucker Conditions II): Consider the maximization problem and suppose that x^* is feasible and satisfies conditions $(KT-1)$ and $(KT-2)$. If $f(\cdot)$ is concave and each $\lambda_j g^j(x)$ (with the λ values obtained from the recipe) is quasiconvex, then x^* is the solution to the maximization problem.

Proof: We want to show that $f(x) - f(x^*) \leq 0$ for all feasible x. Since $f(\cdot)$ is concave, then, according to the first-order derivative characterization of concavity of $f(\cdot)$,

$$f(x) - f(x^*) \leq \nabla f(x^*) \cdot (x - x^*) \underbrace{=}_{(KT-1)} \sum_{j=1}^{m} \lambda_j \nabla g^j(x^*) \cdot (x - x^*)$$

where we use the first order condition $(KT-1)$.

It therefore suffices to show that for all $j = 1, \dots, m$, and all feasible x,

$$\lambda_j \nabla g^j(x^*) \cdot (x - x^*) \leq 0.$$

The above inequality is satisfied for those j such that $g^j(x^*) < 0$, because then $\lambda_j = 0$ from the complementary slackness condition $(KT-2)$.

For those j such that $g^j(x^*) = 0$, we have $g^j(x) \leq g^j(x^*)$ (because x is feasible), and hence $-\lambda_j g^j(x) \geq -\lambda_j g^j(x^*)$ because $\lambda_j \geq 0$.

Since the function $-\lambda_j g^j(x)$ is quasiconcave (because $\lambda_j g^j(x)$ is quasiconvex), it follows from the first-order derivative characterization of quasiconcavity that $\nabla(-\lambda_j g^j(x^*)) \cdot (x - x^*) \geq 0$, and thus, $\lambda_j \nabla g^j(x^*) \cdot (x - x^*) \leq 0$. ■

Necessity of the Kuhn and Tucker Condition and Constraint Qualification (Ch. 3.5)

Consider the following maximization problem with inequality constraints.

$$\max_{x \in S} f(x) \text{ subject to } g^j(x) \leq 0, \; j = 1, \ldots, m$$

The following condition plays an important role when one uses the Kuhn-Tucker condition.

Definition: A solution x^* to the constrained maximization problem satisfies the **constraint qualification** if the gradient vectors $\nabla g^j(x^*)$ $(1 \leq j \leq m)$ corresponding those constraints that are active (binding) at x^*, are linearly independent.

Theorem (Necessity for Kuhn–Tucker Conditions): Suppose that $x^* = (x_1^*, \ldots, x_n^*)$ solves the constrained maximization problem where $f(\cdot)$ and $g^1(\cdot), \ldots, g^m(\cdot)$ are C^1 functions. Suppose furthermore that x^* satisfies the constraint qualification. Then, there exist unique numbers $\lambda_1, \ldots, \lambda_m$ such that $(KT-1)$ and $(KT-2)$ hold at $x = x^*$.

Proof: We skip the proof. ■

Example:

$$\max_{x,y} f(x,y) = xy + x^2 \text{ subject to } \begin{cases} g^1(x,y) = x^2 + y \leq 2 \\ g^2(x,y) = -y \leq -1 \end{cases}$$

We set up the Lagrangian:

$$\mathcal{L}(x,y) = xy + x^2 - \lambda_1(x^2 + y - 2) - \lambda_2(-y + 1).$$

The K-T conditions are:

$$\begin{array}{rl} \text{(i)} & \mathcal{L}_x^{'} = y + 2x - 2\lambda_1 x = 0 \\ \text{(ii)} & \mathcal{L}_y^{'} = x - \lambda_1 + \lambda_2 = 0 \\ \text{(iii)} & \lambda_1 \geq 0,\ x^2 + y \leq 2, \text{ and } \lambda_1(x^2 + y - 2) = 0 \\ \text{(iv)} & \lambda_2 \geq 0,\ y \geq 1, \text{ and } \lambda_2(-y + 1) = 0. \end{array}$$

Case 1: Both constraints are binding

Then, $x^2 + y = 2$ and $y = 1 \Rightarrow x = \pm 1$ and $y = 1$.

When $x = y = 1$, (i) and (ii) yield $\lambda_1 = 3/2$ and $\lambda_2 = 1/2$.

Thus, $(x, y) = (1, 1)$ with $\lambda_1 = 3/2$ and $\lambda_2 = 1/2$ is a solution candidate.

When $x = -1$ and $y = 1$, (i) and (ii) yield $\lambda_1 = 1/2$ and $\lambda_2 = 3/2$.

Since $f(1, 1) = 2$ and $f(-1, 1) = 0$, $(x, y) = (1, 1)$ with $\lambda_1 = 3/2$ and $\lambda_2 = 1/2$ is a solution candidate.

Case 2: g^1 is binding but g^2 is not.

Then, $x^2 + y = 2$ and $y > 1$. From (iv), $\lambda_2 = 0$. From (ii), $x = \lambda_1$.

Plugging $\lambda_1 = x$ into (i), we obtain

$$y + 2x - 2x^2 = 0 \underbrace{\Rightarrow}_{y=2-x^2} 3x^2 - 2x - 2 = 0.$$

The solutions are $x = (1 \pm \sqrt{7})/3$. But $x = \lambda_1 \geq 0$, only $x = (1 + \sqrt{7})/3$ is admissible.

However, plugging this into $y = 2 - x^2$, we obtain

$$y = \frac{2}{9}(5 - \sqrt{7}) < \frac{2}{9}(5 - 2) = \frac{2}{3} < 1,$$

which contradicts $y > 1$.

So, there is no solution candidate in this case.

Case 3: g^1 is not binding but g^2 is binding.

Then, $x^2 + y < 2$ and $y = 1$. From (iii), $\lambda_1 = 0$.

Then (i) gives $x = -1/2$ and (ii) gives $\lambda_2 = 1/2$.

Thus, $(x, y) = (-1/2, 1)$ with $\lambda_1 = 0$ and $\lambda_2 = 1/2$ is a solution candidate.

Case 4: Both constraints are not binding.

Then, $x^2 + y < 2$ and $y > 1$.

(iii) and (iv) gives $\lambda_1 = \lambda_2 = 0$.

However, plugging $\lambda_1 = \lambda_2 = 0$ into (i) and (ii), we have $y = 0$, which contradicts $y \geq 1$.

So, there is no solution candidate in this case.

The two solution candidates are $f(1, 1) = 2$ and $f(-1/2, 1) = -1/4$ and the objective function is highest at $(1, 1)$.

Weierstraas' theorem applies here.

Define

$$D = \left\{(x, y) \in \mathbb{R}^2 |\ x^2 + y \leq 2\right\} \cap \left\{(x, y) \in \mathbb{R}^2 |\ y \geq 1\right\}$$

as the set of all feasible points in this question.

We claim that D is a closed set. Basically, we need to show the following two facts: (1) any set involving a weak inequality is closed and (2) the intersection of two closed sets is closed.

Define $B = \{(x, y) \in \mathbb{R}^2 |\ x^2 + y^2 \leq 5\}$ as a closed ball around $(0, 0)$ with radius $\sqrt{5}$.

We claim $D \subseteq B$, which implies that D is a bounded set.

Since $x^2 \leq 2 - y$ and $y \geq 1$, we have $x^2 \leq 1$.

Fix $(x, y) \in D$. Then,

$$\begin{aligned} x^2 + y^2 & \underbrace{\leq}_{x^2 \leq 1} 1 + y^2 \underbrace{\leq}_{y \leq 2 - x^2} 1 + (2 - x^2)^2 \\ & = x^4 - 4x^2 + 5 = (x^2 - 2)^2 + 1 \underbrace{\leq}_{0 \leq x^2 \leq 1} 5. \end{aligned}$$

This implies $(x, y) \in B$. Thus, D is a bounded set.

Since the objective function f is continuous and the set of all feasible points is compact, by Weierstrass theorem, this constrained maximization problem has a solution.

Checking the constraint qualification

The gradients of the two constraints are $\nabla g^1(x, y) = (2x, 1)$ and $\nabla g^2(x, y) = (0, -1)$.

In Case 1, suppose, by way of contradiction, that there exist $\alpha \neq 0$ and a feasible point (x, y) that falls into Case 1 such that $\nabla g^1(x, y) = \alpha \nabla g^2(x, y)$. This implies that $(2x, 1) = (0, -\alpha)$. This equality holds only when $x = 0$ and $\alpha = -1$.

However, in Case 1, we have $x^2 + y = 2$ and $y = 1$, which further imply that $x = \pm 1$. This contradicts the previous conclusion that $x = 0$. Therefore, these two vectors are linearly independent.

In Case 2, only g^1 is binding so that we only look at $\nabla g^1(x, y) = (2x, 1)$, which is linearly independent because it is not the zero vector.

In Case 3, only g^2 is binding so that we only look at $\nabla g^2(x, y) = (0, -1)$, which is linearly independent because it is not the zero vector.

In Case 4, the constraint qualification trivially holds.

So, the constraint qualification holds.

We conclude that $(1, 1)$ is the solution to the problem.