A Quick Math Review

Xiang Han (SUFE)

This short note provides a quick review of some of the essential mathematical tools used in this course. Most results are stated without proofs. For a more rigorous and complete treatment, see, for example, *Mathematics for Economists* by Simon and Blume.

1. Functions

A *function* f from a set X (the *domain* of f) to a set Y, denoted as $f : X \to Y$, maps each x in X to a unique element $f(x)$ in Y. If for each $y \in Y$, there exists $x \in X$ such that $y = f(x)$, then we say f is a function *onto* Y (or, f is *surjective*). f is called *one-to-one* (or, f is *injective*) if $f(x_1) = f(x_2)$ implies $x_1 = x_2$. If f is a one-to-one function from X onto Y, then it is *bijective*.

Let $\mathbb{N} = \{1, 2, 3, ...\}$ denote the set of *natural numbers*. An (infinite) *sequence* is a function whose domain is $\mathbb{N}$, often denoted as $\{x^n\}_{n=1}^{\infty}$, or simply $\{x^n\}$. A set S is called *countable* if there exists a one-to-one function $f : S \to \mathbb{N}$.

2. Real numbers

Denote the set of *real numbers* as $\mathbb{R}$ and the set of *rational numbers* as $\mathbb{Q}$. There are much more reals than rationals.

Theorem *$\mathbb{R}$ is not countable; $\mathbb{Q}$ is countable.*

But between any two different real numbers, there exists a rational number.

Theorem *For any $x, y \in \mathbb{R}$ with $x < y$, there exists $z \in \mathbb{Q}$ such that $x < z < y$.*

Given a set of real numbers $S \subseteq \mathbb{R}$, $b \in \mathbb{R}$ is an *upper bound* of S if $x \leq b$ for each $x \in S$. $c \in \mathbb{R}$ is a *least upper bound* of S if it is an upper bound of S and $c \leq b$ for each upper bound b of S. From this definition, it is clear that a least upper bound, if it exists, is unique. While a set of real numbers with an upper bound may not include a maximum element (e.g. think of the open interval $(0, 1)$), a nice property of real numbers is that the least upper bound of such a set always exists.

Completeness Axiom *Every non-empty set S of real numbers which has an upper bound has a least upper bound.*

The least upper bound of S is denoted as "sup S". Similarly, b is a *lower bound* of S if $b \leq x$ for each $x \in S$. c is a *greatest lower bound* of S, denoted as "inf S", if it is a lower bound of S and $c \geq b$ for each lower bound b of S. From the completeness axiom, it can be seen that each non-empty set of real numbers which has a lower bound has a greatest lower bound.[1] In the simple example given above, we have $\sup(0,1) = 1$ and $\inf(0,1) = 0$.

3. Convergence and continuity

From now on, let us focus on the Euclidean space $\mathbb{R}^N$, which is of most interest in Economics. For any $x, y \in \mathbb{R}^N$, their distance is the usual Euclidean norm

$$\|x - y\| = \sqrt{\sum_{n=1}^{N}(x_n - y_n)^2}$$

Definiton A sequence $\{x^n\}$ in $\mathbb{R}^N$ *converges* to $x \in \mathbb{R}^N$, denoted as $x^n \to x$, if for every $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $\|x^n - x\| < \epsilon$ for all $n > N$.

Proposition *If $x^n \to x$, $y^n \to y$ and $x^n \geq y^n$ for all n, then $x \geq y$.*

Definition A set $S \subseteq \mathbb{R}^N$ is *open* if for every $x \in S$ there exists some $\epsilon > 0$ such that for all y with $\|y - x\| < \epsilon$, $y \in S$. $S \subseteq \mathbb{R}^N$ is *closed* if for any $\{x^n\} \subseteq S$ with $x^n \to x$, we have $x \in S$.

Proposition *A set S is closed if and only if $\mathbb{R}^N \setminus S$ is open.*

We are mostly interested in properties of *real-valued* functions.

Definition Let $S \subseteq \mathbb{R}^N$ and $f : S \to \mathbb{R}$. f is *continuous* at $x \in S$ if for any $\epsilon > 0$, there exists $\delta > 0$ such that for any $y \in S$ with $\|y - x\| < \delta$ we have $|f(y) - f(x)| < \epsilon$.

We say f is continuous on S if it is continuous at every $x \in S$. The continuity of a function can also be defined using convergent sequences.

[1] This can be shown by considering the set "$-S$", $-S = \{-x : x \in S\}$

Proposition *Let $S \subseteq \mathbb{R}^N$ and $f : S \to \mathbb{R}$. f is continuous at $x \in S$ if and only if for any $\{x^n\} \subseteq S$ with $x^n \to x \in S$, we have $f(x^n) \to f(x)$.*

4. Convex analysis

A set $S \subseteq \mathbb{R}^N$ is *convex* if for any $x, y \in S$ and $\alpha \in [0,1]$ we have $\alpha x + (1-\alpha)y \in S$.

Definition Let $S \subseteq \mathbb{R}^N$ be convex and $f : S \to \mathbb{R}$. f is *concave* if for any $x, y \in S$ and $\alpha \in [0,1]$, we have $f(\alpha x + (1-\alpha)y) \geq \alpha f(x) + (1-\alpha)f(y)$. If the inequality is strict for all $x \neq y$ and $\alpha \in (0,1)$, then f is *strictly concave.*

Convexity and *strict convexity* of a function are defined analogously by reversing the inequalities. Note that the concavity of a function f may not be preserved under a strictly increasing transformation. For example, $f(x) = \sqrt{x}$ is concave, and $g(x) = x^4, x \geq 0$, is a strictly increasing function, but $g \circ f(x) = x^2$ is not concave. It is tempting to think about what a strictly increasing transformation of a concave function looks like.

Let $f : S \to \mathbb{R}$ be concave and $g : \mathbb{R} \to \mathbb{R}$ a strictly increasing function. Consider any $x, y \in S$ and $\alpha \in [0,1]$. We have

$$\begin{aligned} g \circ f(\alpha x + (1-\alpha)y) &\geq g[\alpha f(x) + (1-\alpha)f(y)] \\ &\geq g[\alpha \cdot \min\{f(x), f(y)\} + (1-\alpha) \cdot \min\{f(x), f(y)\}] \\ &= g(\min\{f(x), f(y)\}) \\ &= \min\{g \circ f(x), g \circ f(y)\} \end{aligned}$$

In fact, $g \circ f$ is *quasiconcave*:

Definition Let $S \subseteq \mathbb{R}^N$ be convex and $f : S \to \mathbb{R}$. f is *quasiconcave* if for any $x, y \in S$ and $\alpha \in [0,1]$, we have $f(\alpha x + (1-\alpha)y) \geq \min\{f(x), f(y)\}$. If the inequality is strict for all $x \neq y$ and $\alpha \in (0,1)$, then f is *strictly quasiconcave*.

Quasiconcavity is also often characterized by convex upper contour sets.

Proposition *Let $S \subseteq \mathbb{R}^N$ be convex and $f : S \to \mathbb{R}$. f is quasiconcave if and only if for any $t \in \mathbb{R}$, the set $\{x \in S : f(x) \geq t\}$ is convex. f is strictly quasiconcave if and only if* *for any $x, y \in S$ with $x \neq y$, $\alpha \in (0,1)$ and $t \in \mathbb{R}$, we have: $f(x) \geq t$ and $f(y) \geq t$ imply $f(\alpha x + (1-\alpha)y) > t$.*

Quasiconvexity can be defined analogously. Let $S \subseteq \mathbb{R}^N$ be convex. $f : S \to \mathbb{R}$ is quasiconvex if for any $x, y \in S$ and $\alpha \in [0,1]$, we have $f(\alpha x + (1-\alpha)y) \leq \max\{f(x), f(y)\}$. If the inequality is strict for all $x \neq y$ and $\alpha \in (0,1)$, then f is strictly quasiconvex. f is quasiconvex if and only if for any $t \in \mathbb{R}$, the set $\{x \in S : f(x) \leq t\}$ is convex.

It can be easily shown that (strict) concavity implies (strict) quasiconcavity, and any strictly increasing transformation of a quasiconcave function is quasiconcave. Analogous results hold for quasiconvexity.

5. Optimization

We consider the problem of maximizing a real-valued function. We first present an existence result. A set $S \subseteq \mathbb{R}^N$ is *bounded* if there exists $r \in \mathbb{R}$ such that $\|x\| < r$ for all $x \in S$. A set $S \subseteq \mathbb{R}^N$ is *compact* if it is closed and bounded.

Theorem *Let $S \subseteq \mathbb{R}^N$ be compact. $f : S \to \mathbb{R}$ is continuous on S. Then there exists $x \in S$ such that $f(x) \geq f(y)$ for all $y \in S$.*

Next, we consider unconstrained optimization, i.e., maximizing a function $f : \mathbb{R}^N \to \mathbb{R}$.

Definition $\bar{x} \in \mathbb{R}^N$ is a *local maximizer* of f if there exists $\epsilon > 0$ such that for any $y \in \mathbb{R}^N$ with $\|y - \bar{x}\| < \epsilon$ we have $f(\bar{x}) \geq f(y)$. $\bar{x} \in \mathbb{R}^N$ is a *global maximizer* of f if $f(\bar{x}) \geq f(y)$ for every $y \in \mathbb{R}^N$.

Local and *global minimizers* are defined analogously. The next theorem gives the necessary, or first-order, conditions.

Theorem *If f is differentiable and $\bar{x}$ is a local maximizer or a local minimizer of f, then $\nabla f(\bar{x}) = 0$, i.e., $\frac{\partial f(\bar{x})}{\partial x_n} = 0$ for $n = 1, \ldots, N$.*

A common sufficient condition is concavity.

Theorem *If f is differentiable, concave and $\nabla f(\bar{x}) = 0$, then $\bar{x}$ is a global maximizer of f.*

Finally, we consider a general constrained optimization problem:

$$\text{Max}_{x\in\mathbb{R}^N}\ \ f(x)$$

$$s.t.\ \ g_1(x)=b_1,\ldots,g_M(x)=b_M$$

$$h_1(x)\leq c_1,\ldots,h_K(x)\leq c_K$$

f, g_m, $m=1,\ldots,M$, and h_k, $k=1,\ldots,K$, are all mappings from $\mathbb{R}^N$ to $\mathbb{R}$. There are M equality constraints and K inequality constraints. Assume that $N\geq M+K$.

$$C=\left\{x\in\mathbb{R}^N: g_1(x)=b_1,\ldots,g_M(x)=b_M, h_1(x)\leq c_1,\ldots,h_K(x)\leq c_K\right\}$$

is the set of points satisfying all the constraints.

Definition $\bar{x}\in C$ is a *local constrained maximizer* of f if there exists $\epsilon>0$ such that for any $y\in C$ with $\|y-\bar{x}\|<\epsilon$ we have $f(\bar{x})\geq f(y)$. $\bar{x}\in C$ is a *global constrained maximizer* of f if $f(\bar{x})\geq f(y)$ for every $y\in C$.

Given $\bar{x}\in C$, the *constraint qualification* is satisfied if the vectors in

$$\{\nabla g_m(\bar{x}): m=1,\ldots,M\}\cup\{\nabla h_k(\bar{x}): h_k(\bar{x})=c_k\}$$

are linearly independent.

The next theorem presents the first-order conditions. All of the involved functions are assumed to be differentiable.

Theorem: Kuhn-Tucker conditions *Suppose that $\bar{x}$ is a local constrained maximizer and the constraint qualification is satisfied. Then there are multipliers $\lambda_m\in\mathbb{R}$, one for each equality constraint, and $\mu_k\in\mathbb{R}_+$, one for each inequality constraint, such that after setting up the Lagrangian*

$$\mathscr{L}(x,\lambda,\mu)=f(x)+\sum_{m=1}^{M}\lambda_m(b_m-g_m(x))+\sum_{k=1}^{K}\mu_k(c_k-h_k(x))$$

we have

$$\frac{\partial \mathscr{L}(\bar{x},\lambda,\mu)}{\partial x_n}=0,\ \ n=1,...,N\ \ \textit{(first-order conditions of the Lagrangian)}$$

and,

$$\mu_k(c_k-h_k(\bar{x}))=0,\ \ k=1,...,K\ \ \textit{(complementary slackness conditions)}$$

In economics applications, we usually have non-negative constraints on the x variables. Suppose that we further require $x_n \geq 0$ for some n. Then we only need *one* modification of the above Kuhn-Tucker conditions: instead of $\frac{\partial \mathscr{L}(\bar{x},\lambda,\mu)}{\partial x_n}=0$, we now have

$$\frac{\partial \mathscr{L}(\bar{x},\lambda,\mu)}{\partial x_n}\leq 0,\ \ \text{with equality if}\ \ \bar{x}_n>0 \tag{1}$$

To see why this is true, we can explicitly add $x_n \geq 0$ as an inequality constraint to the original problem, i.e., the $(K+1)th$ inequality constraint is $h_{K+1}(x)=-x_n \leq 0$. The new Kuhn-Tucker conditions have two differences from before. First, the first order condition of the new Lagrangian, $\widetilde{\mathscr{L}}$, with respect to x_n is

$$\frac{\partial \widetilde{\mathscr{L}}(\bar{x},\lambda,\mu)}{\partial x_n}=\frac{\partial \mathscr{L}(\bar{x},\lambda,\mu)}{\partial x_n}+\mu_{K+1}=0 \tag{2}$$

Second, there is one more complementary slackness condition

$$\mu_{K+1}\bar{x}_n=0 \tag{3}$$

Recall that the multiplier corresponding to each inequality constraint must be non-negative, so $\mu_{K+1}\geq 0$. Then (2) and (3) imply

$$\frac{\partial \mathscr{L}(\bar{x},\lambda,\mu)}{\partial x_n}\leq 0\ \ \text{and}\ \ \frac{\partial \mathscr{L}(\bar{x},\lambda,\mu)}{\partial x_n}\bar{x}_n=0 \tag{4}$$

which is equivalent to (1).

Example.[2]

$$\text{Max } f(x) = x_1(x_2 + 3)$$

$$s.t.\ \ x_1 + x_2 \leq 2$$

$$x_1 \geq 0, x_2 \geq 0$$

f is continuous on $C = \{(x_1, x_2) : x_1 \geq 0, x_2 \geq 0, x_1 + x_2 \leq 2\}$ and C is compact, so a global constrained maximizer exists. Let $\bar{x} \in C$ be a global constrained maximizer.
Set up the Lagrangian:

$$\mathscr{L}(x_1, x_2, \lambda) = x_1 x_2 + 3x_1 + \lambda(2 - x_1 - x_2)$$

$\bar{x}$ must satisfy Kuhn-Tucker conditions:[3]

$$\bar{x}_2 + 3 - \lambda \leq 0, \text{ with equality if } \bar{x}_1 > 0 \tag{5}$$

$$\bar{x}_1 - \lambda \leq 0, \text{ with equality if } \bar{x}_2 > 0 \tag{6}$$

$$\lambda \geq 0 \text{ and } \lambda(2 - \bar{x}_1 - \bar{x}_2) = 0 \tag{7}$$

In addition, we also know that $\bar{x} \in C$, i.e., it satisfies all the constraints.

From (5), we have $\lambda > 0$ since $\bar{x}_2 \geq 0$. Then from (7)

$$\bar{x}_1 + \bar{x}_2 = 2 \tag{8}$$

If $\bar{x}_1 > 0$ and $\bar{x}_2 > 0$, then (5) and (6) hold with equality. Together with (8), we can solve for $\bar{x}_1$ and $\bar{x}_2$: $\bar{x}_1 = 2.5$, $\bar{x}_2 = -0.5 < 0$, contradiction.

If $\bar{x}_1 = 0$, then by (8), $\bar{x}_2 = 2$. Then (6) implies $\lambda = \bar{x}_1 = 0$, contradiction.

We are only left with the case where $\bar{x}_2 = 0$. By (8), $\bar{x}_1 = 2$. Then (5) implies $\lambda = 3$. In this case, all of the Kuhn-Tucker conditions, as well as all the constraints, are satisfied. $(\bar{x}_1 = 2, \bar{x}_2 = 0)$ is the unique global constrained maximizer.

[2] This example is taken from Qianfeng Tang's notes.

[3] It can be verified that the constraint qualification is always satisfied at a global constrained maximizer in this problem.