MATHEMATICAL PRELIMINARIES

PONTUS RENDAHL

Let's start with some definitions of useful spaces. They are not difficult, but take some of the details for granted for now. Most will be clearer with time.

Real Vector Space A real vector space, X, is a set of elements (vectors) together with two operations: addition and scalar multiplication,

i) If $x, y \in X$ then $x + y \in X$
ii) If $x \in X$ then $\forall \alpha \in \mathbb{R}$, $\alpha x \in X$

In addition, $\forall x, y, z \in X$ and $\alpha, \beta \in \mathbb{R}$

a) $x + y = y + x$
b) $(x + y) + z = x + (y + z)$
c) $\alpha(x + y) = \alpha x + \alpha y$
d) $(\alpha + \beta)x = \alpha x + \beta x$
e) $(\alpha\beta)x = \alpha(\beta x)$
f) $x + 0 = x$
g) $0 \times x = 0$
h) $1 \times x = x$

Examples:

– The closed interval $[0, 1]$ is *not* a RVS
– The set X containing all infinite sequences $(x_0, x_1, x_2, \ldots)$, with $x_i \in \mathbb{R}$ *is* a RVS
– The set of all continuous functions on $[a, b]$, $C[a, b]$, is a RVS
– The set $\mathbb{N}$ is not

A metric space is a set S together with a metric (distance function) $\rho : S \times S \to \mathbb{R}$ such that $\forall x, y, z \in S$

a) $\rho(x, y) \geq 0$ and $\rho(x, y) = 0$ iff $x = y$
b) $\rho(x, y) = \rho(y, x)$
c) $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$ (triangle inequality)

Examples:

– $S = \mathbb{N}$ with $\rho(x, y) = |x - y|$ is a metric space

– $S = C[a,b]$ with $\rho(x,y) = \max_{t\in[a,b]} |x(t) - y(t)|$ is a metric space

A normed vector space is a vector space S with a *norm* $\|\cdot\| : S \to \mathbb{R}$ such that for all $x, y \in S$ and $\alpha \in \mathbb{R}$

a) $\|x\| \geq 0$, and $\|x\| = 0$ iff $x = 0$
b) $\|\alpha x\| = |\alpha| \times \|x\|$
c) $\|x + y\| \leq \|x\| + \|y\|$ (triangle inequality)

Examples:

– $S = \mathbb{R}^\ell$ with $\|x\| = (\sum_{i=1}^{\ell} x_i^2)^{1/2}$ (Euclidean space)
– $S = C[a,b]$ with $\|x\| = \sup_{t\in[a,b]} |x(t)|$ (as will become apparent, the "sup" and the "max" are equal to each other in this particular case)

In a normed vector space it is convention to define the metric $\rho(x,y) = \|x - y\|$ (can you prove that $\rho(\cdot)$ must satisfy the conditions for a metric if $\|\cdot\|$ satisfies the conditions for a norm?)

For most of the propositions below I will use the absolute value norm $|x - y|$. However, this is just to keep the analysis more transparent. All results go through using a general function $\rho(x,y)$ that satisfies the criteria for a metric listed above.

The sup and the inf If $S \subseteq \mathbb{R}$, then $\mu = \sup S$ if

i) $\mu \geq x \;\forall x \in S$
ii) $\forall \varepsilon > 0 \;\exists\; y \in S$ such that $y > \mu - \varepsilon$

$\lambda = \inf S$ if

i) $\lambda \leq x \;\forall x \in S$
ii) $\forall \varepsilon > 0 \;\exists\; y \in S$ such that $y < \lambda + \varepsilon$

If $\mu \in S$, then μ is the max of S. If $\lambda \in S$, then λ is the min of S.

Examples:

– The sup of (a,b) is b, the inf is a. The max and the min does not exist.
– The sup of $(a,b]$ is b, the inf is a. The max is b but the min does not exist.

Fact 1 For all bounded $S \subseteq \mathbb{R}$ the sub and the inf exist

Definition If the supremum μ belongs to (is an element of) S, μ is called the **maximum**. Similarly, if the infimum λ belongs to S, λ is called the **minimum**.

A sequence is a collection of numbers $(x_1, x_2, \ldots)$. Another way to see this is as a function $f : \mathbb{N} \to X$. We usually denote a sequence $\{x_n\}_{n\in\mathbb{N}}$, or $\{x_n\}_{n=1}^{\infty}$, or just $\{x_n\}$ if I'm lazy.

Subsequence A subsequence of $\{x_n\}$ is $(x_{n_1}, x_{n_2}, x_{n_3}, \ldots)$ where n_k is a sequence in $\mathbb{N}$ such that $n_1 < n_2 < n_3 < \ldots$

Alternatively $\{y_k\}$ is a subsequence of $\{x_n\}$ if $\exists\ g : \mathbb{N} \to \mathbb{N}$ such that

i) $y_k = x_{g(x)}$
ii) $k' > k \Rightarrow g(k') > g(k)$

Convergence of a sequence A sequence $\{x_n\}_{n=0}^{\infty}$ in S converges to $x \in S$ if for each $\varepsilon > 0$, $\exists N_\varepsilon$ such that

$$|x_n - x| < \varepsilon, \quad \forall\ n \geq N_\varepsilon$$

If the sequence $\{x_n\}$ converge to x, we normally write $x_n \to x$ instead of spelling out the entire definition of convergence.

Proposition 1 If $\{x_n\}$ converges to x, then x is unique

Proof. Suppose not. That is, $\exists\ y$ such that $\forall \varepsilon > 0$, $\exists\ \mathbb{N}_\varepsilon$ such that $|x_n - y| < \varepsilon\ \forall n \geq N_\varepsilon$ and $x \neq y$ (i.e. $|x - y| \neq 0$). Then $|x_n - x| + |x_n - y| < 2\varepsilon$ for some $N = \max(N_x, N_y)$. By the triangle inequality $|x - y| < 2\varepsilon$. But since $\varepsilon > 0$ was chosen arbitrarily, x must equal y, which is a contradiction. $\square$

Proposition 2 Every monotone *real* sequence that is bounded converges

Proof. (For the increasing case) $\{x_n\}$ is a bounded sequence with $x_{n+1} \geq x_n$. Define $x = \sup\{x_1, x_2, \ldots\}$. By the definition of the supremum, $x_M > x - \varepsilon$ for some $M \in \mathbb{N}$, and any $\varepsilon > 0$ (notice that M is likely to depend on the choice of ε). Thus $x \geq x_n \geq x_M > x - \varepsilon$ for all $n \geq M$. Or simply $x \geq x_n > x - \varepsilon$, so $\varepsilon > x - x_n \Rightarrow \varepsilon > |x - x_n|$. Since $\varepsilon > 0$ was arbitrarily chosen, the result follows. $\square$

Proposition 3 Every real sequence has a monotone subsequence

Proof. Let $S_m = \{x_m, x_{m+1}, \ldots\}$. If for any (*any!*) $m \in \mathbb{N}$, S_m has no maximum the result is trivial (make sure you understand this).

If S_m has a max for *each* $m \in \mathbb{N}$ then construct $x_{m_1} = \max S_1$, $x_{m_2} = \max S_{m_1+1}$, $x_{m_3} = \max S_{m_2+1}$, and the subsequence $\{x_{m_k}\}$ must be decreasing. $\square$

Proposition 4 (The Bolzano-Weierstrass Theorem) Every bounded real sequence has a convergent subsequence.

Proof. This follows as a corollary of Propositions 2 and 3. As a bounded real sequence has a monotone subsequence (Proposition 3), and every bounded monotone sequence converges, every bounded real sequence must have a convergent subsequence. □

Definition x is an **accumulation point** of the set S if for each $\varepsilon > 0$ $\exists$ $y \in S$ $(y \neq x)$ such that $\|x - y\| < \varepsilon$.

Notice that the sup, μ, of a set S satisfies $\mu - y < \varepsilon$, for any $\varepsilon > 0$ and for some $y \in S$. Thus $|\mu - y| < \varepsilon$. As a consequence, either μ belongs to S and the sup coincides with the max, **or** μ us an accumulation point. (take for instance the set $\{1, 2, 3\}$. What is the sup? Is the sup an accumulation point? Does the sup coincide with the max?)

Proposition 5 If y is an accumulation point of S, then there exist a sequence $\{y_n\}$ with $y_n \in S$ $\forall n$ that converges to y

Proof. Pick $\{\varepsilon_n\} = \{\frac{1}{n}\}_{n \in \mathbb{N}}$, and create a sequence $\{y_n\}$ such that $|y_n - y| < \varepsilon_n$. This can be done as y is an accumulation point. Now, for all $\varepsilon > 0$, $\exists$ $N \in \mathbb{N}$ such that $\varepsilon_n < \varepsilon$, $\forall n \geq N$, and thus that $|y - y_n| < \varepsilon$, $\forall n \geq N$. And $\lim_{n \to \infty} |y_n - y| = 0$ implies that $y_n \to y$ □

Definition S is a **closed set** if for each $x_n \to x$ with $x_n \in S$ $\forall n$ implies that $x \in S$.

Theorem 1 Every closed and bounded set $F \subset \mathbb{R}$ has a max and a min.

Proof. $\mu = \sup F$ exists (and is less than ∞). Either μ is the max (and the proof is done), or μ is an accumulation point. If μ is an accumulation point there is a sequence in F that converges to μ (Proposition 5). As F is closed, μ belongs to S is therefore the maximum. □

Definition A sequence $\{x_n\}$ is a **Cauchy sequence** if for any $\varepsilon > 0$ $\exists$ an N_ε such that $\rho(x_m, x_n) < \varepsilon$, $\forall$ $m, n \geq N_\varepsilon$

Proposition 6 Every Cauchy sequence is bounded

Proof. If $\{x_n\}$ is Cauchy, we know that $|x_m - x_n| < \varepsilon$, $\forall m, n \geq N_\varepsilon$. Thus, $|x_m| - |x_n| \leq |x_m - x_n| < \varepsilon$, $\forall m, n \geq N_\varepsilon$. Set $n = N_\varepsilon$. Then $|x_m| < |x_{N_\varepsilon}| + \varepsilon < \infty$, $\forall m \geq N_\varepsilon$ □

Theorem 2 The space $(\mathbb{R}, |\cdot|)$ is complete. That is every Cauchy sequence in $\mathbb{R}$ with norm $|\cdot|$ converges in $\mathbb{R}$ (to a unique point)

Proof. Let $\{x_n\}$ be a Cauchy sequence in $\mathbb{R}$. Then $\{x_n\}$ is bounded (Proposition 6) and there exist a convergent subsequence $\{x_{n_k}\}$ (Proposition 4). Thus for every $\varepsilon > 0$ there exist $N, M \geq 0$ such that

$$|x_{n_k} - x| < \varepsilon \quad \text{and} \quad |x_m - x_n| < \varepsilon \quad \text{for all } \ k \geq N_\varepsilon, \text{ and } m, n \geq M_\varepsilon$$

Pick N_ε such that $n_k \geq M_\varepsilon$ (this can be done as n_k is strictly monotone in k). Then $|x_{n_k} - x_n| < \varepsilon$. Thus

$$|x_{n_k} - x| + |x_{n_k} - x_n| < 2\varepsilon$$

And by the triangle inequality we have $|x_n - x| < 2\varepsilon$ for all $n \geq N_\varepsilon$ □

Definition A function $f : D \to I$ is **continuous** at $c \in D$ if for any $\varepsilon > 0$ $\exists$ $\delta > 0$ such that for any $x \in D$ with $|x - c| < \delta$ implies $|f(x) - f(c)| < \varepsilon$.

If this definition holds for all $c \in D$, then the function f is a continuous function.

Definition A function $f : D \to I$ is **uniformly continuous** if for any $\varepsilon > 0$ $\exists$ $\delta > 0$ such that if $s, t \in D$ and $|s - t| < \delta$ implies $|f(s) - f(t)| < \varepsilon$.

The difference between continuity and uniform continuity lies in the fact that the δ in the definition of a continuous function depends both on ε and on c. The δ in the definition of a uniformly continuous function depends only on ε. For instance, $f = 1/x$ is continuous on $(0, 1)$, but it's not uniformly continuous. What about on $[0, 1]$? It's not even continuous!

On a compact domain, every continuous function is also uniformly continuous.

Proposition 7 If $f : D \to I$ is continuous at $c \in closure(D)$ then for any $x_n \to c$ we have $f(x_n) \to f(c)$.

Proof. Pick a $\varepsilon > 0$. Then $\exists$ a $\delta(\varepsilon) > 0$ such that when $|x - c| < \delta(\varepsilon)$ then $|f(x) - f(c)| < \varepsilon$. Pick any $\{x_n\}$ in D that converges to c. Then $\exists$ an $N_{\delta(\varepsilon)}$ such that $|x_n - c| < \delta(\varepsilon)$, for all $n \geq N_{\delta(\varepsilon)}$. Thus, $|f(x_n) - f(c)| < \varepsilon$, for all $n \geq N_{\delta(\varepsilon)}$ □

Proposition 8 Suppose that for any $x_n \to c$ we have $f(x_n) \to f(c)$, then f is continuous at c.

Proof. Suppose f is *not* continuous at c. Then for some $\tilde{\varepsilon}$, $\nexists$ a δ such that $|x - c| < \delta$ implies $|f(x) - f(c)| < \tilde{\varepsilon}$ [i.e. for all $\delta > 0$, $\exists$ $\hat{x}$ such that $|\hat{x} - c| < \delta$, and $|f(\hat{x}) - f(c)| \geq \tilde{\varepsilon}$]. Define $\delta_n = \frac{1}{n}$, and create the sequence $\{x_n\}$ such that $|x_n - c| < \delta_n$ *and* $|f(x_n) - f(c)| \geq \tilde{\varepsilon}$, for all $n \in \mathbb{N}$. Clearly $x_n \to c$ but $f(x_n)$ does *not* converge to $f(c)$. □

Theorem 3 A continuous real valued function f defined on a compact (i.e. closed and bounded) set S has a maximum and a minimum.

Proof. As f is continuous and real valued and the domain is compact, the image of f must be bounded and real. Thus by Fact 1, f has a supremum; $\mu = \sup_x f(x)$. Then either μ is the max and the proof is trivial, or μ is an accumulation point of the image of f. Thus $\exists$ a sequence $y_n \to \mu$. Define $\{x_n\}$ implicitly as $y_n = f(x_n)$. The sequence $\{x_n\}$ need not to be unique, but it is bounded. As a consequence $\exists$ an $\{x_{n_k}\}$ such that $x_{n_k} \to x^*$ for some $x^* \in S$ (as S is closed and $x_{n_k} \in S$). Therefore the subsequence $y_{n_k} = f(x_{n_k})$ must converge to $y^* = f(x^*)$ as f is continuous (Proposition 7), and since $y_n \to \mu$ we have that $y_{n_k} \to \mu$.[1] Since the limit point of a sequence is unique we have that $\mu = y^*$. And since $y^* = f(x^*)$ with $x^* \in S$, y^* is the maximum. $\square$

Convergence of functions In functional analysis we are often interested in sequences of functions, $\{f_n\}$, rather than sequences of real numbers. While it is quite easy to define norms or metrics that captures the idea of convergence for real numbers, some complications arise when we are studying the convergence of functions. In general, we will consider two types of convergence; *pointwise convergence* and *uniform convergence*.

A sequence of functions $\{f_n\}$ converges pointwise to f if for *each* x, and for each $\varepsilon > 0$, there exist a N such that $|f_n(x) - f(x)| < \varepsilon$, for all $n \geq N$.

A sequence of functions $\{f_n\}$ converges uniformly to f if for *all* x, and for each $\varepsilon > 0$, there exist a N such that $|f_n(x) - f(x)| < \varepsilon$, for all $n \geq N$.

The difference between pointwise and uniform convergence is that in the former definition N depends on both ε and the particular choice of x, while in the latter N is independent of the choice of x. It is quite easy to see that if a sequence of functions converges uniformly to f, then $\{f_n\}$ is a convergent sequence under the *sup-norm*; $\|f_n - f\| = \sup_{x \in X} |f_n(x) - f(x)|$. And, conversely, if a sequence of functions converges in the sup-norm to f, then convergence is uniform.

Example: The sequence of functions $f_n = \frac{1}{x+1/n}$ on $(0,1)$ converges pointwise (to $f = \frac{1}{x}$), but it does not converge uniformly.

Definition A **complete normed vector space** is a normed vector space in which every Cauchy sequence converges to a point in the space. This space is also known as a **Banach space**.

Theorem 3 (Theorem 3.1 SLP) Let $X \subseteq \mathbb{R}^\ell$ and let $C(X)$ be the space of continuous function $f : X \to \mathbb{R}$ with the sup-norm $\|f\| = \sup_{x \in X} |f(x)|$. Then $C(X)$ is a complete normed vector space (if X is not compact $C(X)$ must be confined to continuous and bounded function)

Proof. The strategy of the proof is to pick an arbitrary Cauchy sequence of functions $\{f_n\}$, with $f_n \in C(X)$ for all $n \in \mathbb{N}$, and then show that

[1] If $\{x_n\}$ is a sequence converging to x, then every subsequence of $\{x_n\}$ is also converging to x.

(1) $f_n \to f$ uniformly
(2) And $f \in C(X)$.

Part 1. The sequence $\{f_n\}$ is Cauchy in the sup-norm. That is, $\forall \varepsilon > 0$ there exist an N such that $\sup_{x \in X} |f_m(x) - f_n(x)| < \varepsilon$, $\forall m, n \geq N$. Let's pick an arbitrary $x \in X$ and an $\varepsilon > 0$ and note that

$$|f_m(x) - f_n(x)| \leq \sup_{x \in X} |f_m(x) - f_n(x)| < \varepsilon \quad \forall m, n \geq N$$

Since $f_n(x)$ is real valued, the sequence $\{f_n(x)\}$ is a Cauchy sequence in real numbers and will therefore converge to a unique limit point (Theorem 2). Repeating this argument for all $x \in X$ gives us a candidate function $f(x)$. However, while f_n converges to f for each x (pointwise convergence), the speed of convergence (the N) will generally depend on the chosen x, and convergence in the sup-norm does not permit this. Therefore, in order to show that $f_n \to f$ uniformly, we must show that $\{f_n\}$ converges in the sup-norm, or equivalently, that for each ε, N can be chosen independently of x.

Let us use $f(x)$ as our candidate limit point. Pick an $\varepsilon > 0$ and an N_ε independently of x such that $\|f_m - f_n\| < \varepsilon$ for all $m, n \geq N_\varepsilon$ (this can be done as $\{f_n\}$ is Cauchy). Then *for all* $x \in X$, and for all $m, n \geq N_\varepsilon$ we have

$$\begin{aligned}
|f_n(x) - f(x)| &= |f_n(x) - f_m(x) + f_m(x) - f(x)| \\
&\leq |f_n(x) - f_m(x)| + |f_m(x) - f(x)| \\
&< \varepsilon + |f_m(x) - f(x)|
\end{aligned}$$

Now, for any x we can always find a m such that $|f_m(x) - f(x)| < \varepsilon$. This value of m is likely to depend on x, but that does not matter. Therefore

$$\begin{aligned}
|f_n(x) - f(x)| &< \varepsilon + |f_m(x) - f(x)| \\
&< 2\varepsilon
\end{aligned}$$

As a consequence, for any $\varepsilon > 0$ there exist an N_ε (which is independent of x) and

$$|f_n(x) - f(x)| < 2\varepsilon \quad \Rightarrow \quad \|f_n(x) - f(x)\| < 2\varepsilon$$

Part 2. Pick a k such that $\|f - f_k\| < \varepsilon$. We can do this as $\{f_n\}$ converges in the sup-norm. By assumption f_k is in $C(X)$ and is therefore continuous. Thus pick a $\delta > 0$ such that

$|x-y|<\delta$ implies $|f_k(x)-f_k(y)|<\varepsilon$. Then,

$$\begin{aligned}|f(x)-f(y)| &= |f(x)-f_k(x)+f_k(x)-f_k(y)+f_k(y)-f(y)| \\ &\leq |f(x)-f_k(x)|+|f_k(x)-f_k(y)|+|f_k(y)-f(y)| \\ &\leq 2\varepsilon+|f_k(x)-f_k(y)| \\ &\leq 3\varepsilon\end{aligned}$$

Thus for any $\varepsilon>0$ there exist a $\delta>0$ such that $|x-y|<\delta$ implies that $|f(x)-f(y)|<3\varepsilon$. As a consequence f is a continuous function and belongs to the space $C(X)$. □

Contraction Mappings Let (S,ρ) be a metric space and let $T:S\to S$ be a function mapping S to itself. Then T is a **contraction mapping** with modulus β if for some $\beta\in(0,1)$, $\rho(Tx,Ty)\leq\beta\rho(x,y)$.

Contraction Mapping Theorem If (S,ρ) is a **complete metric space** and $T:S\to S$ is a contraction mapping with modulus β then,

(1) T has exactly one fixed point, $v\in S$ (i.e. $Tv=v$).
(2) For any $v_0\in S$, $\rho(T^n v_0,v)\leq\beta^n\rho(v_0,v)$ (i.e. $T^n v_0\to v$).

Proof. The proof is in three steps

Part 1 Start with a $v_0\in S$, and define $v_{n+1}=Tv_n$. We want to show that $\{v_n\}$ is a Cauchy sequence. Notice that

$$\begin{aligned}\rho(v_{n+1},v_n) &\leq \beta\rho(v_n,v_{n-1})\leq\beta^2\rho(v_{n-1},v_{n-2}) \\ &\leq \beta^n\rho(v_1,v_0)\end{aligned}$$

and by the triangle inequality for any $m>n$

$$\rho(v_m,v_n)\leq\rho(v_m,v_{m-1})+\rho(v_{m-1},v_{m-2})+\ldots+\rho(v_{n+1},v_n)$$

Combining

$$\begin{aligned}\rho(v_m,v_n) &\leq \beta^{m-1}\rho(v_1,v_0)+\beta^{m-2}\rho(v_1,v_0)+\ldots+\beta^n\rho(v_1,v_0) \\ &= \beta^n[\beta^{m-n-1}\rho(v_1,v_0)+\beta^{m-n-2}\rho(v_1,v_0)+\ldots+\rho(v_1,v_0)] \\ &= \beta^n[\beta^{m-n-1}+\beta^{m-n-2}+\ldots+1]\rho(v_1,v_0) \\ &\leq \frac{\beta^n}{1-\beta}\rho(v_1,v_0)\end{aligned}$$

Since $\rho(v_1,v_0)\in\mathbb{R}$, for any $\varepsilon>0$ there exist an N such that $n\geq N$, $\frac{\beta^n}{1-\beta}\rho(v_1,v_0)<\varepsilon$. Thus, for any $\varepsilon>0$ there exist an N such that $\rho(v_m,v_n)<\varepsilon$, $\forall\ m,n\geq N$.

Part 2 Show that v is a fixed point $v = Tv$. We know that

$$\begin{aligned}\rho(Tv, v) &\leq \rho(Tv, v_n) + \rho(v_n, v) \\ &\leq \beta\rho(v, v_{n-1}) + \rho(v_n, v) < \beta\varepsilon + \varepsilon\end{aligned}$$

if N is sufficiently large. Thus $\rho(Tv, v) \to 0$, and $Tv = v$ follows.

Part 3 Show that v is unique. Suppose there exist a $\hat{v} \neq v$ which is another fixed point. Then

$$0 < a = \rho(\hat{v}, v) = \rho(T\hat{v}, Tv) \leq \beta\rho(\hat{v}, v) = \beta a$$

This is a contradiction. □

Corollary 1 Notice that

$$\begin{aligned}\rho(v, v_n) &\leq \rho(v, v_{n+1}) + \rho(v_{n+1}, v_n) \\ &\leq \beta\rho(v, v_n) + \rho(v_{n+1}, v_n)\end{aligned}$$

So

$$\rho(v, v_n) \leq \frac{1}{1-\beta}\rho(v_{n+1}, v_n)$$

This tells us how far off v_n is from v by computing the difference between v_n and v_{n+1}.

Corollary 2 If (S, ρ) is a complete metric space and S' is a closed subset of S, then (S', ρ) is also a complete metric space. (Proof? Homework).

Corollary 3 Let (S, ρ) be a complete metric space and let $T : S \to S$ be a contraction mapping with fixed point $v \in S$. If S' is a closed subset of S *and* $T(S') \subseteq S'$, then $v \in S'$. And if $T(S') \subseteq S'' \subseteq S'$, then $v \in S''$.

Proof. Part (i). Suppose $v_0 \in S'$. Since $T(s') \subseteq S'$, $\{T^n v_0\} \in S'$ converges to a unique v. Since S is closed $v \in S'$. *Part (ii).* $v = Tv \in S''$ □

Blackwell's sufficient conditions for a contraction mapping. Let $X \subseteq \mathbb{R}^\ell$ and let $B(X)$ denote the space of **bounded** functions with the sup-norm. Let $T : B(X) \to B(X)$ be an operator satisfying

a) **Monotonicity**. If $f, g \in B(X)$ and $f(x) \leq g(x)$ $\forall x \in X$, then $(Tf)(x) \leq (Tg)(x)$ $\forall x \in X$

b) **Discounting**. There exist a $\beta \in (0, 1)$ such that

$$[T(f + a)](x) \leq (Tf)(x) + \beta a \quad \forall f \in B(X),\ a \geq 0,\ x \in X$$

Then T is a contraction mapping with modulus β.

Proof. For *any* $f, g \in B(X)$, we have $f \leq g + \|f-g\|$ and $g \leq f + \|f-g\|$. Then by monotonicity

$$Tf \leq T(g + \|f-g\|)$$

And by discounting

$$Tf \leq T(g + \|f-g\|) \leq Tg + \beta\|f-g\|$$

Reversing the role of f and g gives

$$Tg \leq T(f + \|f-g\|) \leq Tf + \beta\|f-g\|$$

Thus

$$\|Tf - Tg\| \leq \beta\|f-g\|$$

□

Correspondences A correspondence $\Gamma : X \to 2^Y$ is a (possibly) multivalued function. A correspondence is *lower hemi-continuous* if $\forall\ y \in \Gamma(x)$ and all $x_n \to x$, $\exists y_n \in \Gamma(x_n)$ such that $y_n \to y$.

Define the graph of Γ as

$$A = \{(x,y) \in X \times Y : y \in \Gamma(x)\}$$

A correspondence is upper hemi-continuous if A is closed. That is if for all $\{x_n, y_n\}$ with $(x_n, y_n) \in A$ we have that $(x,y) \in A$.

We say that a correspondence if continuous if it is both upper- and lower hemi-continuous. For instance, suppose that $f \geq g$ and both f and g are continuous functions, then

$$\Gamma(x) = \{y \in Y : g(x) \leq y \leq f(x)\}$$

is a continuous correspondence.

Let us define the functions h and G as follows

$$h(x) = \max_{y \in \Gamma(x)} f(x,y)$$

and

$$G(x) = \{y \in \Gamma(x) : f(x,y) = h(x)\}$$

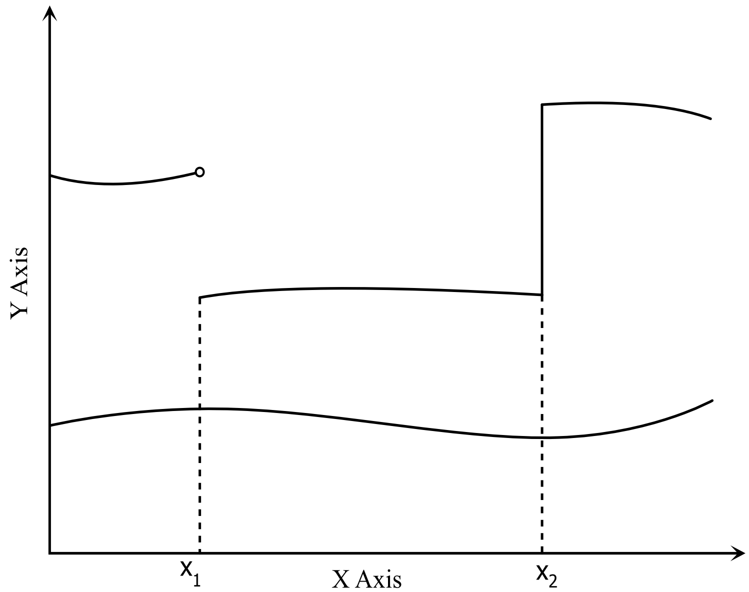

FIGURE 1. A correspondence which is not u.h.c at x_1 (but l.h.c), and not l.h.c at x_2 (but u.h.c).

A Lemma Suppose A (the graph of Γ) is a compact and convex set, and suppose $f : A \to \mathbb{R}$ is a continuous and strictly concave function (NB: strict concavity actually implies continuity). Then

$$x^* = \text{argmax}_{x \in A} f(x)$$

exists and is *unique*.

Proof. Since A is compact and f is a real valued continuous function the maximum exists. Now suppose x^* is not unique. Then there exist some $x' \neq x^*$ such that $f(x^*) = f(x')$ (with $x' \in A$). Define $x'' = \lambda x^* + (1-\lambda)x'$. By concavity

$$f(x'') > \lambda f(x^*) + (1-\lambda) f(x') = f(x^*)$$

Since A is convex $x'' \in A$. This contradicts that x^* indeed was the maximum. $\square$

The theorem of the maximum Let $f : X \times Y \to \mathbb{R}$ be strictly concave in y and continuous in x. Let $\Gamma : X \to 2^Y$ be a compact-, convex-, and continuous correspondence. Then $G(x)$ and $h(x)$ are bounded continuous functions.

Proof. Part (i) $G(x)$ is a function. By the previous lemma, y^* (defined as $h(x) = f(x, y^*)$) is unique and $G(x) = y^*$.

Part (ii) $G(x)$ is bounded. As $\Gamma(x)$ is bounded and $G(x) \in \Gamma(x)$. Hence $G(x)$ is bounded.

Part (iii) $G(x)$ is continuous. Pick a sequence $\{x_n\}$ which converges to some value x. The we will show that $y_n = G(x_n)$ converges to some value y such that $y = G(x)$.

For any $x_n \to x$, $y_n = G(x_n)$ is bounded and has a convergent subsequence with some limit y. Denote this subsequence $\{y_{n_k}\}$ with the associated $\{x_{n_k}\}$. Since Γ is u.h.c we have that $y \in \Gamma(x)$. Since Γ is l.h.c, for every $z \in \Gamma(x)$ and every $x_n \to x$, $\exists$ a sequence $\{z_n\}$ that converges to z, with $z_n \in \Gamma(x_n)$. By construction $f(x_{n_k}, y_{n_k}) \geq f(x_{n_k}, z_{n_k})$, $k = 1, 2, \ldots$ Since $x_{n_k} \to x$, $y_{n_k} \to y$, $z_{n_k} \to z$, and f is continuous we have that $f(x, y) \geq f(x, z)$. Since this holds for every $z \in \Gamma(x)$, $y = G(x)$, which is unique. Now, since every converging subsequence converges to a unique limit point, the original sequence, $\{y_n\}$, must be convergent. Thus $G(x_n) \to G(x)$ $\forall x_n \to x$.

Part (iv) $h(x)$ is bounded and continuous. $h(x)$ is bounded as Γ is compact and f is continuous. Pick any $x_n \to x$. Then $y_n = G(x_n) \to y = G(x)$. Thus $h(x_n) = f(x_n, y_n) \to h(x) = f(x, y)$ as f is a continuous function. □

The theorem of the maximum above is a slightly simplified version of Berge's theorem which is presented in Stokey, Lucas and Prescott (1989). In particular, by assuming concavity we can exploit the aforementioned lemma, which makes life easier. In the original formulation, concavity is dispensed with but $h(x)$ can still be shown to be a continuous function, while $G(x)$ is, instead, and upper hemi-continuous correspondence.

The envelope theorem Let $X \subseteq \mathbb{R}^\ell$ and let $h : X \to \mathbb{R}$ be a concave continuous function. Let $x_0 \in X$ and let D be a neighbourhood of x_0. If there is a differentiable function $g : D \to \mathbb{R}$ with $g(x_0) = h(x_0)$ and $h(x) \geq g(x)$ for all $x \in D$, then $h(x)$ is differentiable at x_0 and $h_i(x_0) = g_i(x_0)$.

Proof. Pick an $\varepsilon > 0$. Since h is concave we have that

$$\begin{aligned} h(x_0) &\geq \frac{1}{2}h(x_0+\varepsilon) + \frac{1}{2}h(x_0-\varepsilon) \\ 2h(x_0) &\geq h(x_0+\varepsilon) + h(x_0-\varepsilon) \\ h(x_0) - h(x_0-\varepsilon) &\geq h(x_0+\varepsilon) - h(x_0) \end{aligned}$$

Since $h(x_0) = g(x_0)$ and $h(x) \geq g(x)$ $\forall x \in D$ we have

$$\frac{g(x_0) - g(x_0-\varepsilon)}{\varepsilon} \geq \frac{h(x_0+\varepsilon) - h(x_0)}{\varepsilon}$$

Also since $h(x_0+\varepsilon) \geq g(x_0+\varepsilon)$ we have

$$\frac{h(x_0+\varepsilon) - h(x_0)}{\varepsilon} \geq \frac{g(x_0+\varepsilon) - g(x_0)}{\varepsilon}$$

Combining inequalities yields

$$\frac{g(x_0) - g(x_0 - \varepsilon)}{\varepsilon} \geq \frac{h(x_0 + \varepsilon) - h(x_0)}{\varepsilon} \geq \frac{g(x_0 + \varepsilon) - g(x_0)}{\varepsilon}$$

Since g is a differentiable function the limits

$$\lim_{\varepsilon \to 0} \frac{g(x_0) - g(x_0 - \varepsilon)}{\varepsilon} = \lim_{\varepsilon \to 0} \frac{g(x_0 + \varepsilon) - g(x_0)}{\varepsilon}$$

exist and equal the partial derivative $g_i(x_0)$. By the Pinching theorem, $h(x)$ must be differentiable at x_0 with derivative $h_i(x_0) = g_i(x_0)$. The proof can then be repeated with $\varepsilon < 0$ but with reversed inequalities. $\square$

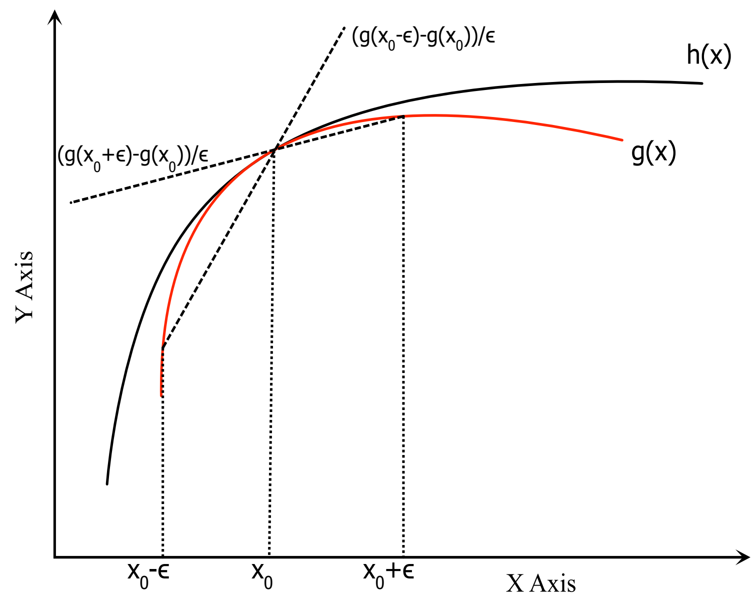

FIGURE 2. The envelope theorem.

OK, so why do we need all this mathematics? Recall that the problem we are looking to solve is given by the Bellman equation

$$v(x) = \max_{y \in \Gamma(x)} \{F(x, y) + \beta v(y)\} \tag{FE}$$

with $\beta \in (0, 1)$. Define the operator T as

$$(Tf)(x) = \max_{y \in \Gamma(x)} \{F(x, y) + \beta f(y)\}$$

Notice that if F and f are bounded, so if Tf, and T maps bounded function to bounded functions: $T : B(X) \to B(X)$. We can also easily see that T satisfy Blackwell's sufficient conditions. For instance, suppose that $f \geq g$, $\forall x \in X$. Then

$$\begin{aligned}(Tg)(x) &= \max_{y\in\Gamma(x)}\{F(x,y)+\beta g(y)\}\\ &= \{F(x,y_g)+\beta g(y_g)\}\\ &\leq \{F(x,y_g)+\beta f(y_g)\}\\ &\leq \max_{y\in\Gamma(x)}\{F(x,y)+\beta f(y)\}\\ &= (Tf)(x)\end{aligned}$$

where y_g denotes the argmax to $(Tg)(x)$. Therefore T satisfies monotonicity. In addition

$$\begin{aligned}[T(f+a)](x) &= \max_{y\in\Gamma(x)}\{F(x,y)+\beta(f(y)+a)\}\\ &= \max_{y\in\Gamma(x)}\{F(x,y)+\beta f(y)+\beta a\}\\ &= (Tf)(x)+\beta a\end{aligned}$$

so T satisfies discounting. Thus T is a contraction mapping with modulus β. It should be noted, although we haven't proved in this note, that the space of bounded function $B(X)$ with the sup-norm is indeed a complete metric space.[2] Thus, by the contraction mapping theorem T has a unique fixed point and this fixed point satisfies (FE). In addition, the sequence $\{v_n\}_{n=0}^{\infty}$ defined as

$$v_{n+1}(x) = \max_{y\in\Gamma(x)}\{F(x,y)+\beta v_n(y)\}$$

with $v_0 \in B(X)$, converges to the unique fixed point, $v(x)$.

Is there anything else to be said about $v(x)$? As we will see it is quite straightforward to show, under the appropriate assumptions, that $v(x)$ is, continuous, monotonically increasing, concave, and sometimes even differentiable.

Continuity of the value function. When is the value function continuous? Suppose that F and f are bounded continuous functions and Γ is a compact-, convex-, and continuous correspondence. Then by the theorem of the maximum, the function Tf

$$(Tf)(x) = \max_{y\in\Gamma(x)}\{F(x,y)+\beta f(y)\}$$

is also a bounded continuous function. Thus the operator T maps bounded continuous functions onto itself, $T : C(X) \rightarrow C(X)$. And as $C(X)$ is a complete metric space in the

[2]This is actually very simple to do. Try it out as an exercise!

sup-norm, we know that the fixed point of T also belongs to $C(X)$. Thus the value function is, under these conditions, a continuous function.

Montonicity of the value function. Let's temporarily introduce two new assumptions

i) Suppose that $F(x, y)$ is strictly increasing in x (in many economic applications this simple translates to utility being strictly increasing in consumption)

ii) The feasibility correspondence Γ is monotone in the sense that if $x \leq x'$ then $\Gamma(x) \subseteq \Gamma(x')$

Then if X is a convex subset of $\mathbb{R}^\ell$ and $F : A \to \mathbb{R}$ is bounded and continuous, the value function, v, is *strictly increasing*.

Proof. Let $C'(X) \subset C(X)$ be the space of bounded, continuous and *non-decreasing* functions on X, and let $C''(X) \subset C'(X)$ be the space of bounded, continuous and *increasing functions*. Since $C'(X)$ is a closed subset of $C(X)$, we only need to show that $T : C'(X) \to C''(X)$ (see corollary 2 and 3 above). Let y^* be the argmax of $(Tf)(x)$. Let $x' > x$. Then,

$$(Tf)(x') \geq F(x', y^*) + \beta v(y^*) > (Tf)(x)$$

$\square$

Concavity of the value function. Let's introduce two new assumptions

i) F is strictly concave. That is

$$F(\lambda(x, y) + (1 - \lambda)(x', y')) > \lambda F(x, y) + (1 - \lambda)F(x', y')$$

for all (x, y) and (x', y') in A with $(x, y) \neq (x', y')$ and $\lambda \in (0, 1)$

ii) Γ is convex. That is, $y \in \Gamma(x)$ and $y' \in \Gamma(x')$ implies that

$$(\lambda(y) + (1 - \lambda)y') \in \Gamma(\lambda x + (1 - \lambda)x')$$

(or alternatively that the graph of Γ, A, is convex)

If X, Γ, F and β satisfy the aforementioned assumptions, then $v(x)$ is strictly concave.

Proof. Let $C'(X) \subset C(X)$ denote the space of bounded, continuous and weakly concave functions. Notice that $C'(X)$ is a closed space. Let $C''(X) \subset C'(X)$ denote the space of bounded, continuous and strictly concave functions. Then it is sufficient to show that $T : C'(X) \to C''(X)$.

Pick $x, x' \in X$. Define $x'' = \lambda x + (1 - \lambda)x'$ for some $\lambda \in (0, 1)$. Let y denote the argmax to $(Tf)(x)$, and y' the argmax to $(Tf)(x')$. Define $y'' = \lambda y + (1 - \lambda)y'$. Then since Γ is convex we have that $y'' = \Gamma(x'')$. Suppose that $f \in C'(X)$, then

$$\begin{aligned}(Tf)(x'') &\geq F(x'', y'') + \beta f(y'') \\ &> \lambda[F(x,y) + \beta f(y)] + (1-\lambda)[F(x',y') + \beta f(y')] \\ &= \lambda(Tf)(x) + (1-\lambda)(Tf)(x')\end{aligned}$$

Thus $T : C'(X) \to C''(X)$ $\square$

Differentiability of the value function. Let X, Γ, F and β satisfy the aforementioned assumptions (including those used for concavity). In addition, let F be a differentiable function. Then we know by the envelope theorem that if there exist some differentiable function $g(x)$ such that $v(x_0) = g(x_0)$ and $v(x) \geq g(x)$ for all x in a neighbourhood of x_0, then $v(x)$ is differentiable at x_0 and it's derivative is given by $g_i(x_0)$. So in order to prove that $v(x)$ is differentiable, we must find a function $g(x)$ that satisfy the above criteria.

Let y^* denote the argmax to $v(x_0)$. That is,

$$v(x_0) = F(x_0, y^*) + \beta v(y^*)$$

And assume that $y^* \in int(\Gamma(x))$. Then consider the function

$$g(x) = F(x, y^*) + \beta v(y^*)$$

Since $y^* \in int(\Gamma(x))$ we must have

$$v(x) = \max_{y \in \Gamma(x)} \{F(x,y) + \beta v(y)\} \geq F(x, y^*) + \beta v(y^*) = g(x)$$

with equality at x_0. Thus, $v(x)$ is differentiable at x_0 with derivative

$$v_i(x_0) = g_i(x_0) = F_i(x_0, y^*)$$

where F_i refers to the derivative of F with respect to the ith argument of x.