Lecture Notes for Week 14 (First Draft)

Monotone Mappings on $\mathbb{R}$ and $\mathbb{R}^n$

Recall that a function $f : \mathbb{R} \to \mathbb{R}$ is said to be

- *increasing* if $f(x_2) \geq f(x_1)$ for all $x_1, x_2 \in \mathbb{R}$ with $x_2 \geq x_1$,
- *decreasing* if $-f$ is increasing, and
- *monotone* provided that f is either increasing or decreasing.

Monotone functions from $\mathbb{R}$ to $\mathbb{R}$ have many special properties. For example, it is a simple exercise in real analysis to prove that if $f : \mathbb{R} \to \mathbb{R}$ is monotone, continuous and satisfies

$$|f(x)| \to \infty \quad \text{as } |x| \to \infty. \tag{1}$$

then f is surjective. We shall generalize this result to mappings on infinite-dimensional spaces.

Since it is completely straightforward to "adjust" results obtained for increasing functions so that they apply to decreasing functions, we shall focus on generalizing notions associated with increasing functions. We shall refer to such mappings as "monotone" (rather than increasing).

The condition

$$f(x_2) \geq f(x_1) \quad \text{for all } x_1, x_2 \in \mathbb{R} \text{ with } x_2 \geq x_1 \tag{2}$$

is not particularly well-suited to generalization. Observe that a function $f : \mathbb{R} \to \mathbb{R}$ satisfies (2) if and only if

$$(f(x_2) - f(x_1)) \cdot (x_2 - x_1) \geq 0 \quad \text{for all } x_1, x_2 \in \mathbb{R}. \tag{3}$$

Observe also that if $f : \mathbb{R} \to \mathbb{R}$ satisfies (3) then (1) holds if and only if

$$\frac{f(x) \cdot x}{|x|} \to \infty \quad \text{as } |x| \to \infty. \tag{4}$$

A function $g : \mathbb{R}^n \to \mathbb{R}^n$ is said to be *monotone* provided that

$$(g(x) - g(y)) \cdot (x - y) \geq 0 \quad \text{for all } x, y \in \mathbb{R}^n. \tag{5}$$

It is a standard result in convex analysis (and a special case of a result that we will prove later) that a C^1-function $\phi : \mathbb{R}^n \to \mathbb{R}$ is convex if and only if $\nabla\phi : \mathbb{R}^n \to \mathbb{R}^n$ is monotone as defined above.

In extending (5) to an infinite-dimensional setting we could use inner products (or semi-inner products) but it is actually very convenient to consider mappings from a Banach space to its dual space.

Definition 14.1: Let X be a real Banach space with dual space X^*. A mapping $F : X \to X^*$ is said to be

(a) *monotone* provided that $\langle F(u) - F(v), u - v\rangle \geq 0$ for all $u, v \in X$,

(b) *strictly monotone* provided that $\langle F(u) - F(v), u - v\rangle > 0$ for all $u, v \in X$ with $u \neq v$,

(c) *bounded* provided $F[B]$ is bounded in X^* for every bounded set $B \subset X$,

(d) *coercive* provided that

$$\frac{\langle F(u), u\rangle}{\|u\|} \to \infty \quad \text{as } \|u\| \to \infty.$$

Remark 14.2: It is an immediate consequence of the definition that if $F : X \to X^*$ is strictly monotone then F is injective.

Theorem 14.3 (Browder-Minty): Let X be a real Banach space with dual space X^*. Assume that X is separable and reflexive. Let $F : X \to X^*$ be given and assume that F is monotone, bounded, continuous, and coercive. Then F is surjective.

The proof of Theorem 14.3 will be based on the Galerkin Method. (Here the idea is to replace a problem in an infinite-dimesional space with a sequence of approximating, finite-dimensional, problems, and pass to the limit.) In order to obtain solutions to the finite-dimensional problems we shall make use of a result that follows fairly easily from Brouwer's Fixed-Point Theorem.

Lemma 14.4: Let $m \in \mathbb{N}$ and $\rho > 0$ be given. Assume that $\Phi : \mathbb{R}^m \to \mathbb{R}^m$ is continuous and satisfies

$$\xi \cdot \Phi(\xi) \geq 0 \quad \text{for all } \xi \in \mathbb{R}^m \text{ with } |\xi| = \rho. \tag{6}$$

Then there exists $\eta \in \mathbb{R}^m$ with $|\eta| \leq \rho$ such that $\Phi(\eta) = 0$.

Proof: Suppose that there is no such η. Define

$$\Psi(\xi) = -\rho \frac{\Phi(\xi)}{|\Phi(\xi)|} \quad \text{for all } \xi \in \overline{B}_\rho(0), \tag{7}$$

and observe that $\Psi : \overline{B}_\rho(0) \to \overline{B}_\rho(0)$ continuously. It follows from Brouwer's Fixed-Point Theorem that there exists $\xi^* \in \overline{B}_\rho(0)$ such that $\xi^* = \Psi(\xi^*)$. Notice that

$$|\xi^*| = |\Psi(\xi^*)| = \rho. \tag{8}$$

Appealing to (6) and using (8) we see that

$$\begin{aligned} 0 \le \xi^* \cdot \Phi(\xi^*) &= -\frac{|\Phi(\xi^*)|}{\rho}\Psi(\xi^*) \cdot \xi^* \\ &= -\rho|\Phi(\xi^*)| < 0, \end{aligned}$$

and this is a contradiction. $\square$

Proof of Theorem 14.3: Let $g \in X^*$ be given. We want to find $u \in X$ such that $F(u) = g$.

Since X is separable, we may choose a linearly dependent sequence $\{x_j\}_{j=1}^\infty$ such that

$$\text{span}(x_j | j \in \mathbb{N}) \text{ is dense in } X.$$

For each $m \in \mathbb{N}$, put

$$V_m = \text{span}(x_1, x_2, \cdots, x_m).$$

Let $m \in \mathbb{N}$ be given. We seek $u_m \in V_m$ such that

$$\langle F(u_m), v\rangle = \langle g, v\rangle \quad \text{for all } v \in V_m. \tag{9}$$

To construct a solution of (9) we define $\Phi : \mathbb{R}^m \to \mathbb{R}^m$ by

$$\Phi_i(\xi) = \langle F\left(\sum_{j=1}^m \xi_j v_j\right) - g, v_i\rangle \quad \text{for all } \xi \in \mathbb{R}^m. \tag{10}$$

For ease of notation, it is convenient to put

$$v_m^\xi = \sum_{j=1}^m \xi_j x_j \quad \text{for all } \xi \in \mathbb{R}^m. \tag{11}$$

Continuity of F implies continuity of Φ. Moreover, we have

$$\begin{aligned} \xi \cdot \Phi(\xi) &= \sum_{i=1}^m \xi_i \Phi_i(\xi) \\ &= \langle F(v_m^\xi) - g, v_m^\xi\rangle \\ &= \langle F(v_m^\xi), v^\xi\rangle - \langle g, v_m^\xi\rangle \\ &\ge \left(\frac{\langle F(v_m^\xi), v_m^\xi\rangle}{\|v_m^\xi\|} - \|g\|\right). \end{aligned} \tag{12}$$

Since all norms on V_m are equivalent (and m is fixed), we know that

$$\|v_m^\xi\| \to \infty \quad \text{as } |\xi| \to \infty. \tag{13}$$

Using the continuity of Φ, (12), (13) and Lemma 14.4, we may choose $\xi^* \in \mathbb{R}^m$ such that $\Phi(\xi^*) = 0$. If we put

$$u_m = v_m^{\xi^*} \tag{14}$$

then u_m satisfies (9). In this way we generate a sequence $\{u_m\}_{m=1}^\infty$ such that u_m satisfies (9) for every $m \in \mathbb{N}$.

Since $u_m \in V_m$ for every $m \in \mathbb{N}$, it follows from (9) and a simple computation that

$$\frac{\langle F(u_m), u_m\rangle}{\|u_m\|} \le \|g\| \quad \text{for all } m \in \mathbb{N}. \tag{15}$$

Since F is coercive, it follows from (15) that the sequence $\{u_m\}_{m=1}^\infty$ is bounded.

Since X is reflexive and separable, and since $\{u_m\}_{m=1}^\infty$ and $\{F(u_m)\}_{m=1}^\infty$ are bounded we may choose a subsequence $\{u_{m_j}\}_{j=1}^\infty$ and $u \in X$, $\phi \in X^*$ such that

$$u_{m_j} \rightharpoonup u \text{ (weakly)}, \quad F(u_{m_j}) \overset{*}{\rightharpoonup} \phi \text{ (weakly}*\text{) as } j \to \infty.$$

We shall show that $\phi = g$ and $F(u) = g$. We know that

$$\langle F(u_{m_j}), v\rangle = \langle g, v\rangle \quad \text{for all } v \in V_{m_j}.$$

Since the spaces V_m are "nested", we see that

$$\langle F(u_{m_j}), v\rangle \to \langle g, v\rangle \text{ as } j \to \infty \quad \text{for all } v \in \bigcup_{m=1}^\infty V_m.$$

It follows that

$$\langle \phi, v\rangle = \langle g, v\rangle \quad \text{for all } v \in \bigcup_{m=1}^\infty V_m.$$

Since $\operatorname{span}(x_k | k \in \mathbb{N})$ is dense in X we see that

$$\phi = g. \tag{16}$$

Now let $v \in X$ be given. Since F is monotone we have

$$\begin{aligned}
0 &\le \langle F(v) - F(u_{m_j}), v - u_{m_j}\rangle \\
&\le \langle F(v), v\rangle - \langle F(v), u_{m_j}\rangle - \langle F(u_{m_j}), v\rangle + \langle F(u_{m_j}, u_{m_j}\rangle \\
&\le \langle F(v), v\rangle - \langle F(v), u_{m_j}\rangle - \langle F(u_{m_j}, v\rangle + \langle g, u_{m_j}\rangle \\
&\to \langle F(v), v\rangle - \langle F(v), u\rangle - \langle g, v\rangle + \langle g, u\rangle \quad \text{as } j \to \infty.
\end{aligned} \tag{17}$$

It follows that

$$\langle F(v) - g, v - u \rangle \geq 0 \text{ for all } v \in X. \tag{18}$$

We can use (18) to conclude that $F(u) = g$. To this end, let $w \in X$, $t > 0$ be given and put

$$v = u + tw.$$

Substituting this choice of v into (18) and using the fact that $t > 0$ we find that

$$\langle F(u + tw) - g, w \rangle \geq 0 \text{ for all } w \in X, \ t > 0. \tag{19}$$

Letting $t \downarrow 0$ in (19) we see that

$$\langle F(u) - g, w \rangle \geq 0 \text{ for all } w \in X. \tag{20}$$

Replacing w with $-w$ in (20) we see that

$$\langle F(u) - g, w \rangle = 0 \text{ for all } w \in X,$$

which implies that $F(u) = g$. $\square$

Remark 14.5: Examination of the proof of Theorem 14.3 reveals that the assumption of continuity of F can be weakened. In particular, it would be enough to know that the restriction of F to finite-dimensional subspaces is continuous and that for every $u, v, w \in X$ the mapping

$$t \to \langle F(u + tw), v \rangle$$

is continuous. This last condition is sometimes called *hemicontinuity*. See Kato for a discussion of this issue.

Fréchet and Gâteaux Differentiability

Definition 14.6: Let X and Y be Banach spaces and $F : X \to Y$ and $x_0 \in X$ be given. We say that F is *Fréchet differentiable* at x_0 provided there exists $L \in \mathcal{L}(X; Y)$ such that

$$\frac{F(x_0 + h) - F(x_0) - Lh}{\|h\|} \to 0 \text{ as } h \to 0, \ h \in X.$$

The linear operator L is called the *Fréchet derivative* of F at x_0 and we write

$$F'(x_0) = \nabla F(x_0) = DF(x_0) = L.$$

We say that F is *Fréchet differentiable on* X if F is Fréchet differentiable at each $x_0 \in X$.

Remark 14.7: If F is Fréchet differentiable at x_0 then F is continuous at x_0. [This follows easily from the definition.]

Remark 14.8: The definition of Fréchet differentiability (and most of the basic results) can easily be adapted to mappings $F : U \to Y$, where U is an open subset of X.

Remark 14.9 (Higher Differentiability):

Proposition 14.10 (Chain Rule): Let X, Y, and Z be Banach spaces, $G : X \to Y$, $F : Y \to Z$, and $x_0 \in X$ be given. Assume that G is Fréchet differentiable at x_0 and that F is Fréchet differentiable at $G(x_0)$. Then $F \circ G$ is Fréchet differentiable at x_0 and

$$(F \circ G)'(x_0) = F'(G(x_0))G'(x_0).$$

Proof: We employ the standard "little oh" notation; in particular, if we write

$$\phi(h) = o(\|h\|) \text{ as } h \to 0,$$

for a function $\phi : X \to Y$ what we really mean is that

$$\lim_{h \to 0} \frac{\phi(h)}{\|h\|} = 0.$$

For $h \in X \backslash \{0\}$ we have

$$G(x_0 + h) = G(x_0) + G'(X_0)h + o(\|h\|) \text{ as } h \to 0,$$

and consequently

$$\begin{aligned} F(G(x_0 + h)) &= F(G(x_0) + G'(x_0)h + o(\|h\|)) \text{ as } h \to 0 \\ &= F(G(x_0)) + F'(G(x_0))[G'(x_0)h + o(\|h\|)] + o(\|h\|) \text{ as } h \to 0 \\ &= F(G(x_0)) + F'(G(x_0))G'(x_0)h + o(\|h\|) \text{ as } h \to 0, \end{aligned}$$

and the result follows. □

Example 14.11: Let $\mathbb{K} = \mathbb{R}$.

(a) Let $X = L^2[0, 1]$ and define $F : X \to X$ by

$$(F(u))(x) = \sin u(x) \text{ for all } x \in [0, 1].$$

Notice that

$$\|F(u) - F(v)\|_2 \leq \|u - v\|_2 \text{ for all } u, v \in X,$$

so that F is (globally) Lipschitz continuous. You should check as an exercise for yourself that F is nowhere Fréchet differentiable. However, changing the space a bit, dramatically alters the result

(b) Let $Y = C[0,1]$ and define $G : Y \to Y$ by

$$G(u(x)) = \sin u(x) \quad \text{for all } x \in [0,1].$$

You should check for yourself as an exercise that G is everywhere Fréchet differentiable and

$$(G'(u)h)(x) = (\cos u(x))h(x) \quad \text{for all } u, h \in Y,\ x \in [0,1].$$

(In fact, $G : Y \to Y$ is actually analytic!)

Definition 14.12: Let X and Y be Banach spaces and $F : X \to Y$, $x_0, v \in X$ be given. We say that F has a Gâteaux variation at x_0 in the direction v provided that

$$\lim_{t \to 0} \frac{F(x_0 + tv) - F(x_0)}{t} \quad \text{exists;}$$

in this case we write

$$\delta F(x_0; v) = \lim_{t \to 0} \frac{F(x_0 + tv) - F(x_0)}{t}.$$

(The variable t is real in the limits above.) $\delta F(x_0; v)$ is called the *Gâteaux variation* of F at x_0 in the direction v. We say that F is *Gâteaux differentiable* at x_0 provided F has a Gâteaux variation $\delta F(x_0; u)$ at x_0 in every direction $u \in X$.

Remark 14.13: Assume that F is Gâteaux differentiable at x_0. The mapping $v \to \delta F(x_0; v)$ need not be linear (even in finite dimensions) as the example below shows. Many authors require that that the mapping $v \to \delta F(x_0; v)$ be linear and continuous as part of the definition of Gâteaux differentiability at x_0.

Example 14.14: Assume that $F : \mathbb{R}^2 \to \mathbb{R}^2$ is homogeneous of degree 1, i.e.

$$F(tx) = tF(x) \quad \text{for all } x \in \mathbb{R}^2,\ t \in \mathbb{R}.$$

Let $v \in \mathbb{R}^2$ be given. Then we have

$$\delta F(0; v) = F(v).$$

It is straightforward to construct a function $F : \mathbb{R}^2 \to \mathbb{R}^2$ that is homogeneous of degree 1, but fails to be linear.

Remark 14.15: Gâteaux differentiability at x_0 does not imply continuity at x_0 – even in finite dimensions – (and even if one assumes that the mapping $v \to \delta F(x_0; v)$ is linear and continuous) as the following example shows.

Example 14.16: Let $X = \mathbb{R}^2$, put

$$s = \{x \in \mathbb{R}^2 : x_2^2 < x_2 < 2x_2^2\},$$

and

$$F(x) = \begin{cases} 1 \text{ if } x \in S, \\ 0 \text{ if } x \notin S. \end{cases}$$

Then we have

$$\delta F(0; v) = 0 \quad \text{for all } v \in \mathbb{R}^2,$$

but F is not continuous at 0.

Proposition 14.17: Let X and Y be Banach spaces and $F : X \to Y$ and $x_0 \in X$ be given. Assume that F is Fréchet differentiable at x_0. Then F is Gâteaux differentiable at x_0 and

$$\delta F(x_0; v) = F'(x_0)v \quad \text{for all } v \in X.$$