Math 600 Lecture 37

Definition: Let (X,d) be a metric space and let $\varphi: X \to X$ be given. We say that φ is a contraction (or a contractive mapping) iff there exists $\lambda \in (0,1)$ such that

$$\forall u,v \in X,\ d(f(u),f(v)) \le \lambda d(u,v).$$

Obviously, a contractive mapping is uniformly continuous on X.

Theorem (the contractive mapping theorem): Let (X,d) be a complete metric space and let $\varphi: X \to X$ be a contraction. Then there exists a unique $x \in X$ such that $\varphi(x) = x$. (Such an x is called a fixed point for φ.)

Proof: Let x_0 be any point of X and define $\{x_n\} \subset X$ by

$$x_{n+1} = \varphi(x_n),\ n = 0,1,2,\dots.$$

Then

$$d(x_2,x_1) = d(\varphi(x_1),\varphi(x_0)) \le \lambda d(x_1,x_0),$$

$$d(x_3,x_2) = d(\varphi(x_2),\varphi(x_1)) \le \lambda d(x_2,x_1) \le \lambda^2 d(x_1,x_0),$$

$$d(x_4,x_3) = d(\varphi(x_3),\varphi(x_2)) \le \lambda d(x_3,x_2) \le \lambda^3 d(x_1,x_0),$$

$$\vdots$$

$$d(x_{n+1},x_n) \le \lambda^n d(x_1,x_0).$$

Therefore, for any $m > n$,

$$d(x_m, x_n) \leq d(x_m, x_{m-1}) + d(x_{m-1}, x_{m-2}) + \cdots + d(x_{n+1}, x_n)$$

$$\leq \lambda^{m-1} d(x_1, x_0) + \lambda^{m-2} d(x_1, x_0) + \cdots + \lambda^n d(x_1, x_0)$$

$$= \left(\sum_{j=n}^{m-1} \lambda^j\right) d(x_1, x_0)$$

$$= \left(\lambda^n \sum_{j=0}^{m-1-n} \lambda^j\right) d(x_1, x_0)$$

$$< \frac{\lambda^n}{1-\lambda} d(x_1, x_0) \to 0 \text{ as } n \to \infty.$$

But this implies that $\{x_n\}$ is Cauchy, and hence, since X is complete, there exists $x \in X$ such that $x_n \to x$. Since φ is continuous on X, we have

$$\varphi(x) = \lim_{n\to\infty} \varphi(x_n) = \lim_{n\to\infty} x_{n+1} = x,$$

and thus x is a fixed point for φ.

Now suppose $x' \in X$ also satisfies $\varphi(x') = x'$. Then

$$d(x', x) = d(\varphi(x'), \varphi(x)) \leq \lambda\, d(x', x)$$

$$\Rightarrow d(x', x) = 0 \text{ (since } \lambda \in (0,1))$$

$$\Rightarrow x' = x.$$

Thus x is the unique fixed point of φ. //

The implicit function theorem

Suppose $f:\mathbb{R}^m\times\mathbb{R}^n\to\mathbb{R}^n$. Then, for each $x\in\mathbb{R}^m$,

$$f(x,y)=0$$

represents a system of n equations in n unknowns:

$$f_1(x,y_1,\dots,y_n)=0,$$

$$f_2(x,y_1,\dots,y_n)=0,$$

$$\vdots$$

$$f_n(x,y_1,\dots,y_n)=0.$$

It seems reasonable to expect that we can solve for y in terms of x, that is, to determine a function $\psi:\mathbb{R}^n\to\mathbb{R}^n$ (or with a domain that is a subset of $\mathbb{R}^n$) such that

$$f(x,\psi(x))=0 \quad \forall x.$$

The implicit function theorem gives conditions under which this is valid locally.

Theorem: Let E be an open subset of $\mathbb{R}^m\times\mathbb{R}^n$, let $f:E\to\mathbb{R}^n$ be differentiable on E, and assume that Df is continuous on E. Suppose that $(x_0,y_0)\in E$ satisfies

$$f(x_0,y_0)=0,$$

$$D_y f(x_0,y_0) \text{ is nonsingular (invertible).}$$

Then exist open sets $U\subset\mathbb{R}^m$, $V\in\mathbb{R}^n$ such that

$$x_0\in U,\ y_0\in V,\ U\times V\subset E$$

and $\psi: U\to V$ such that

$$f(x, \psi(x)) = 0 \quad \forall x \in U.$$

Moreover, for all $x \in U$, $y = \psi(x)$ is the only point in V satisfying

$$f(x,y) = 0.$$

Finally, ψ is continuously differentiable.

Before we can prove this theorem, we must understand the derivative $Df(x,y)$ and partial derivatives $D_x f(x,y)$ and $D_y f(x,y)$.

Recall that, if f is differentiable at (x,y), then

$$f(x+u, y+v) = f(x,y) + Df(x,y)(u,v) + o(\|(u,v)\|),$$

where

$$\|(u,v)\| = \sqrt{\|u\|^2 + \|v\|^2}.$$

Write $L = Df(x,y) \in \mathcal{L}(\mathbb{R}^m \times \mathbb{R}^n, \mathbb{R}^n)$. We can define

$$L_y \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n), \; L_y v = L(0,v) \quad \forall v \in \mathbb{R}^n$$

and

$$L_x \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n), \; L_x u = L(u,0) \quad \forall u \in \mathbb{R}^m.$$

Since $(u,v) = (u,0) + (0,v)$ and L is linear, we have

$$\begin{aligned} L(u,v) &= L((u,0)+(0,v)) = L(u,0) + L(0,v) \\ &= L_x u + L_y v \qquad \forall (u,v) \in \mathbb{R}^m \times \mathbb{R}^n. \end{aligned}$$

We define the partial derivatives $D_x f(x,y)$ and $D_y f(x,y)$ to be the unique elements of $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ and $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$, respectively, satisfying

$$Df(x,y)(u,v) = D_x f(x,y)u + D_y f(x,y)v \quad \forall (u,v) \in \mathbb{R}^m \times \mathbb{R}^n.$$

Note that

$$f(x+u, y+v) = f(x,y) + Df(x,y)(u,v) + o\,(\|(u,v)\|)$$

$$\Rightarrow f(x+u, y) = f(x,y) + Df(x,y)(u,0) + o\,(\|(u,0)\|)$$

$$\Rightarrow f(x+u,y) = f(x,y) + D_x f(x,y)\,u + o\,(\|u\|).$$

Thus, regarding y as constant, $f(\cdot,y)$ is differentiable at x and its derivative at x is $D_x f(x,y)$.

Similarly, regarding x as constant, $f(x,\cdot)$ is differentiable at y and its derivative at y is $D_y f(x,y)$.

$$\boxed{f \text{ differentiable at } (x,y) \Rightarrow (f(\cdot,y) \text{ is differentiable at } x \text{ and } f(x,\cdot) \text{ is differentiable at } y)}$$

The converse is not true: Suppose

$$f(x+u, y) = f(x,y) + D_x f(x,y)u + o\,(\|u\|),$$

$$f(x, y+v) = f(x,y) + D_y f(x,y)v + o\,(\|v\|)$$

$$\Rightarrow f(x+u, y+v) - f(x,y) = \underbrace{f(x+u, y+v) - f(x, y+v)}_{?} + f(x, y+v) - f(x,y)$$

$$= \quad ? \quad + D_y f(x,y)v + o\,(\|v\|)$$

(We cannot go any further because we don't know that $f(\cdot, y+v)$ is differentiable at x for all v. Even if we were given this hypothesis, we would need to know, at least, that $D_x f(x, y+v)$ is continuous in v.)