MATH 6701, AUTUMN 2024

The Inverse Mapping Theorem

A *metric space* is a pair (X, d) consisting of a set X and a *distance function* $\mathrm{d} : X \times X \to [0, \infty)$ such that $\mathrm{d}(x, x') = \mathrm{d}(x', x)$, $\mathrm{d}(x, x'') \le \mathrm{d}(x, x') + \mathrm{d}(x', x'')$ for any $x, x', x'' \in X$, and $\mathrm{d}(x, x') > 0$ unless $x = x'$. A sequence $x_i \in X$, $i = 1, 2, \ldots$ of points in X then is said to *converge to a limit* $x \in X$ if $\mathrm{d}(x_i, x) \to 0$ as $i \to \infty$, and it is called a *Cauchy sequence* if $\mathrm{d}(x_i, x_j) \to 0$ as i, j simultaneously tend to ∞. The metric space (X, d) is called *complete* if every Cauchy sequence in (X, d) converges.

Any subset $K \subseteq X$ of a metric space (X, d) forms a metric space (K, d) with d restricted to $K \times K$.

The *open ball* $B_r(z) \subseteq X$ (with the center $z \in X$ and the radius $r > 0$) in the metric space (X, d) is defined by $B_r(z) = \{x \in X : \mathrm{d}(x, z) < r\}$, and the *closed ball* $\overline{B}_z(r)$ by $\overline{B}_z(r) = \{x \in X : \mathrm{d}(x, z) \le r\}$. A set $U \subseteq X$ is called *open* if it is the union of some (possibly empty, or infinite) collection of open balls. A *neighborhood* of a point $x \in X$ is any open set containing x; as for manifolds, a sequence x_i, $i = 1, 2, \ldots$ of points in X converges to a limit $x \in X$ if and only if each neighborhood of x contains the x_i for all but finitely many i.

A *norm* in a real or complex vector space V is a function $V \to [0, \infty]$, usually written as $v \mapsto \|v\|$ (or $v \mapsto |v|$, when $\dim V < \infty$), such that $\|v\| > 0$ if $v \ne 0$, and $\|v + w\| \le \|v\| + \|w\|$, $\|\lambda v\| = |\lambda| \cdot \|v\|$ for $v, w \in V$ and all scalars λ. With a fixed norm, V is called a *normed vector space*, and it naturally becomes a metric space (V, d) with $\mathrm{d}(v, w) = \|v - w\|$. A normed vector space is called a *Banach space* if it is complete as a metric space.

Banach's Fixed-Point Theorem. *Let $K \subseteq X$ be a subset of a metric space (X, d) such that (K, d) is complete and let a mapping $h : K \to X$ satisfy the condition*

$$\mathrm{d}(h(x), h(x')) \le C\, \mathrm{d}(x, x')$$

for all $x, x' \in K$ and some C with $0 \le C < 1$. If, moreover,

(a) *there is $z \in K$ with $h^i(z) \in K$ for all integers $i \ge 0$,*

or

(b) *$B_r(z) \subseteq K$ for some $z \in K$ and $r = (1 - C)^{-1}\, \mathrm{d}(z, h(z))$,*

$B_0(z)$ being the empty set, then there exists a unique $x \in K$ with $h(x) = x$.

Proof. Uniqueness of x is clear as $C < 1$. To establish its existence, set $z_i = h^i(z)$ as long as it makes sense for a given $z \in K$ and integers $i \ge 0$. Then $\mathrm{d}(z_i, z_{i+1}) \le C^i\, \mathrm{d}(z, h(z))$ (induction on $i \ge 0$), and so, for integers $j \ge 0$ such that z_{i+j} exists,

$$\mathrm{d}(z_i, z_{i+j}) \le \sum_{s=i}^{i+j-1} \mathrm{d}(z_s, z_{s+1}) \le \left[\sum_{s=i}^{\infty} C^s\right] \mathrm{d}(z, h(z)) = \frac{C^i}{1 - C}\, \mathrm{d}(z, h(z))\,. \tag{1}$$

We may assume that $h(z) \neq z$. Then, for r as in (b), (1) yields $\mathrm{d}(z_i, z_{i+j}) < r$ and, setting $i = 0$, we see that (b) implies (a). On the other hand, for z as in (a), (1) shows that the z_i form a Cauchy sequence, and we can take $x = \lim_{i\to\infty} z_i$.

Corollary. *Given a complete metric space* (X, d) *and a mapping* $h : X \to X$ *such that* $\mathrm{d}(h(x), h(x')) \leq C\,\mathrm{d}(x, x')$ *for all* $x, x' \in X$ *and some* C *with* $0 \leq C < 1$*, there exists a unique* $x \in X$ *with* $h(x) = x$*.*

Let V, W now be finite-dimensional real or complex vector spaces carrying fixed norms (both denoted by $|\ |$). Any linear mapping $T : V \to W$ is *bounded* in the sense that $|Tv| \leq C|v|$ for some constant $C \geq 0$ and all $v \in V$ [**DG**, p.198, Problem 11]. The smallest constant C with this property is called the *operator norm* of T and denoted by $|T|$. If, moreover, $U \subseteq V$ is an open set, $h : U \to V$ is a C^1 mapping and x, z are points in U such that U contains the whole segment $\overline{xz} = \{z + t(x - z) : 0 \leq t \leq 1\}$ connecting z to x, then we have the estimate

$$(2) \qquad |h(x) - h(z)| \leq |x - z| \cdot \sup_{u \in \overline{xz}} |dh_u|$$

– see [**DG**, p.199, Problem 19] – involving the operator norm of $dh_u : V \to W$, the supremum (which, in fact, is a maximum) being finite since $\overline{xz}$ is compact and the function $u \mapsto |dh_u|$ is continuous.

Lemma *Let* U, U' *be open sets in finite-dimensional vector spaces* V, W*, respectively, and let a* C^k *mapping* $F : U \to U'$ *with* $1 \leq k \leq \infty$ *be one-to-one and onto, and such that, for each* $x \in U$*, the differential* $dF_x : V \to W$ *is a linear isomorphism. Then the inverse mapping* $F^{-1} : U' \to U$ *is also* C^k *differentiable.*

Proof. Fix norms in V, W (both denoted $|\ |$). For any fixed $z \in V$, differentiability of F at z means that

$$(3) \qquad F(x) - F(z) = dF_z(x - z) + \alpha(x, z)\,, \qquad \frac{\alpha(x, z)}{|x - z|} \to 0 \quad \text{as } x \to z\,.$$

Since $|dF_z(x - z)| \geq 2C|x - z|$ for some constant $C > 0$ [**DG**, p.198, Problem 13], choosing $\varepsilon > 0$ with $|\alpha(x, z)| \leq |x - z|$ for all $x \in U$ with $|x - z| < \varepsilon$, we obtain, for such x, $|F(x) - F(z)| \geq |dF_z(x - z)| - |\alpha(x, z)|$ in view of (3) and [**DG**, p.22, Problem 10], that is, $|F(x) - F(z)| \geq C|x - z|$ for all x sufficiently close to any fixed $z \in U$ and a suitable $C > 0$, depending on z. (Thus, F^{-1} is continuous.) Applying $(dF_z)^{-1}$ to both sides of (3) and writing $\zeta = F(z)$, $\xi = F(x)$, we obtain

$$F^{-1}(\xi) - F^{-1}(\zeta) = (dF_z)^{-1}(\xi - \zeta) + \beta(\xi, \zeta)$$

with $\beta(\xi, \zeta) = -(dF_z)^{-1}\alpha(x, z)$. Thus, F^{-1} is differentiable at ζ and $d(F^{-1})_\zeta = (dF_z)^{-1}$ since. due to the estimate $|\xi - \zeta| \geq C|x - z|$.

$$|\xi - \zeta|^{-1}|\beta(\xi, \zeta)| \leq C^{-1}|x - z|^{-1}|\beta(\xi, \zeta)| = C^{-1}|(dF_z)^{-1}(|x - z|^{-1}\alpha(x, z))| \to 0$$

as $\xi \to \zeta$. Induction on s now shows that $\zeta \mapsto d(F^{-1})_\zeta = (dF_{F^{-1}(\zeta)})^{-1}$ is a C^{s-1} differentiable mapping for each $s = 1, \ldots, k$. This completes the proof.

We have the following fundamental result.

The Inverse Mapping Theorem. *Let $F : M \to N$ be a C^k mapping between C^k manifolds, $1 \le k \le \infty$. If the differential $A = dF_z : T_zM \to T_wN$ at a point $z \in M$ is a linear isomorphism, where $w = F(z)$, then, for a suitable neighborhood U of z in M, the image $F(U)$ is an open subset of N and $F : U \to F(U)$ is a C^k diffeomorphism.*

Proof. Using local coordinates, we may assume that M, N are open subsets of finite-dimensional real vector spaces V, W endowed with some fixed norms, both denoted $|\;|$. For any fixed $y \in W$, define the C^k mapping $h : M \to V$ by $h(x) = x + A^{-1}(y - F(x))$. As $dh_x = \mathrm{Id} - A^{-1}dF_x$, we have $dh_z = 0$ and hence there is a closed ball $\overline{B}_\varepsilon(z)$ centered at z, of some radius $\varepsilon > 0$, with the operator-norm inequality $|dh_x| \le C$ for all $x \in \overline{B}_z(\varepsilon)$ and any chosen constant $C \in (0,1)$. (Rather than fixing ε, we should be ready to replace it by $\varepsilon' \in (0,\varepsilon)$.) On the other hand, $\mathrm{d}(z, h(z)) = |z - h(z)| \le |A^{-1}|\cdot|y - w|$ and so, whenever $y \in U' = B_w(\delta)$ with $w = F(z)$ and $\delta = (1-C)\varepsilon/|A^{-1}|$, the assumptions of Banach's fixed-point theorem will be satisfied, according to (2), by $X = V$, $K = \overline{B}_z(\varepsilon)$, this C, our h (depending on y), and $r = \varepsilon$ in assumption (b). The resulting unique $x \in K = \overline{B}_z(\varepsilon)$ with $h(x) = x$, that is, $y = F(x)$, must actually lie in $B_z(\varepsilon)$. In fact, we may replace it with $\varepsilon' \in (0,\varepsilon)$ chosen so that our given $y \in U' = B_w(\delta)$ lies in $B_w(\delta')$, where $\delta' = (1-C)\varepsilon'/|A^{-1}|$, and invoke the existence and uniqueness of x for these ε' and δ'. Consequently, $F : U \to U'$ is one-to-one and onto, where $U = B_z(\varepsilon) \cap F^{-1}(U')$, and our assertion follows from the above lemma. Note that the lemma is applicable since – due to the freedom of making ε smaller – we may select our U to be a subset of any prescribed neighborhood of z.