# MA2321—Analysis in Several Variables School of Mathematics, Trinity College Michaelmas Term 2017 Lecture 24 (November 30, 2017)

David R. Wilkins

# 11. The Inverse and Implicit Function Theorems

## 11.1. Local Invertibility of Differentiable Functions

**Definition**

Let $\varphi\colon X \to \mathbb{R}^n$ be a continuous function defined over an open set $X$ in $\mathbb{R}^n$ and mapping that open set into $\mathbb{R}^n$, and let $\mathbf{p}$ be a point of $X$. A *local inverse* of the map $\varphi\colon X \to \mathbb{R}^n$ around the point $\mathbf{p}$ is a continuous function $\mu\colon W \to X$ defined over an open set $W$ in $\mathbb{R}^n$ that satisfies the following conditions:

**(i)** $\mu(W)$ is an open set in $\mathbb{R}^n$ contained in $X$, and $\mathbf{p} \in \mu(W)$;

**(ii)** $\varphi(\mu(\mathbf{y})) = \mathbf{y}$ for all $\mathbf{y} \in W$.

If there exists a function $\mu\colon W \to X$ satisfying these conditions, then the function $\varphi$ is said to be *locally invertible* around the point $\mathbf{p}$.

### Lemma 11.1

*Let $\varphi \colon X \to \mathbb{R}^n$ be a continuous function defined over an open set $X$ in $\mathbb{R}^n$ and mapping that open set into $\mathbb{R}^n$, let $\mathbf{p}$ be a point of $X$. and let $\mu \colon W \to X$ be a local inverse for the map $\phi$ around the point $\mathbf{p}$. Then $\varphi(\mathbf{x}) \in W$ and $\mu(\varphi(\mathbf{x})) = \mathbf{x}$ for all $\mathbf{x} \in \mu(W)$.*

**Proof**

The definition of local inverses ensures that $\mu(W)$ is an open subset of $X$, $\mathbf{p} \in \mu(W)$ and $\varphi(\mu(\mathbf{y})) = \mathbf{y}$ for all $\mathbf{y} \in W$. Let $\mathbf{x} \in \mu(W)$. Then $\mathbf{x} = \mu(\mathbf{y})$ for some $\mathbf{y} \in W$. But then $\varphi(\mathbf{x}) = \varphi(\mu(\mathbf{y})) = \mathbf{y}$, and therefore $\varphi(\mathbf{x}) \in W$. Moreover $\mu(\varphi(\mathbf{x})) = \mu(\mathbf{y}) = \mathbf{x}$, as required. ∎

Let $\varphi\colon X \to \mathbb{R}^n$ be a continuous function defined over an open set $X$ in $\mathbb{R}^n$ and mapping that open set into $\mathbb{R}^n$, let $\mathbf{p}$ be a point of $X$. and let $\mu\colon W \to X$ be a local inverse for the map $\phi$ around the point $\mathbf{p}$. Then the function from the open set $\mu(W)$ to the open set $W$ that sends each point $\mathbf{x}$ of $\mu(W)$ to $\varphi(x)$ is invertible, and its inverse is the continuous function from $W$ to $\varphi(W)$ that sends each point $\mathbf{y}$ of $W$ to $\mu(\mathbf{y})$. A function between sets is *bijective* if it has a well-defined inverse. A continuous bijective function whose inverse is also continuous is said to be a *homeomorphism*. We see therefore that the restriction of the map $\varphi$ to the image $\mu(W)$ of the local inverse $\mu\colon W \to X$ determines a homeomorphism from the open set $\mu(W)$ to the open set $W$.

**Example**

The function $\varphi\colon \mathbb{R}^2 \to \mathbb{R}^2 \setminus \{(0,0)\}$ defined such that

$$\varphi(u,v) = (e^u \cos v, e^u \sin v)$$

for all $u, v \in \mathbb{R}^2$ is locally invertible, though it is not bijective. Indeed, given $(u_0, v_0) \in \mathbb{R}$, let

$$\begin{aligned} W &= \{(r\cos(v_0+\theta), r\sin(v_0+\theta)) : \\ &\qquad r,\theta \in \mathbb{R},\ \ r > 0 \text{ and } -\pi < \theta < \pi\}, \end{aligned}$$

and let

$$\mu(r\cos(v_0+\theta), r\sin(v_0+\theta)) = (\log r, v_0+\theta)$$

whenever $r > 0$ and $-\pi < \theta < 1$. Then $W$ is an open set in $\mathbb{R}^2$, the function $\mu\colon W \to \mathbb{R}^2$ is continuous,

$$\mu(W) = \{(u,v) \in \mathbb{R}^2 : v_0 - \pi < v < v_0 + \pi\},$$

and $\mu(\varphi(u,v)) = (u,v)$ for all $(u,v) \in \mu(W)$.

A continuously differentiable function may have a continuous inverse, but that inverse is not guaranteed to be differentiable, as the following example demonstrates.

**Example**

Let $f\colon \mathbb{R} \to \mathbb{R}$ be defined so that $f(x) = x^3$ for all real numbers $x$. The function $f$ is continuously differentiable and has a continuous inverse $f^{-1}\colon \mathbb{R} \to \mathbb{R}$, where $f^{-1}(x) = \sqrt[3]{x}$ when $x \geq 0$ and $f^{-1}(x) = -\sqrt[3]{-x}$ when $x < 0$. This inverse function is not differentiable at zero.

**Lemma 11.2**

*Let $\varphi\colon X \to \mathbb{R}^n$ be a continuously differentiable function defined over an open set $X$ in $\mathbb{R}^n$. Suppose that $\varphi$ is locally invertible around some point $\mathbf{p}$ of $X$. Suppose also that a local inverse to $\varphi$ around $\mathbf{p}$ is differentiable at the point $\varphi(\mathbf{p})$. Then the derivative $(D\varphi)_{\mathbf{p}}\colon \mathbb{R}^n \to \mathbb{R}^n$ of $\varphi$ at the point $\mathbf{p}$ is an invertible linear operator on $\mathbb{R}^n$. Thus if*

$$\varphi(x_1, x_2, \ldots, x_n) = (y_1, y_2, \ldots, y_n),$$

*for all $(x_1, x_2, \ldots, x_n) \in X$, where $y_1, y_2, \ldots, y_n$ are differentiable functions of $x_1, x_2, \ldots, x_n$, and if $\varphi$ has a differentiable local inverse around the point $\mathbf{p}$, then the Jacobian matrix*

$$\begin{pmatrix} \dfrac{\partial y_1}{\partial x_1} & \dfrac{\partial y_1}{\partial x_2} & \cdots & \dfrac{\partial y_1}{\partial x_n} \\ \dfrac{\partial y_2}{\partial x_1} & \dfrac{\partial y_2}{\partial x_2} & \cdots & \dfrac{\partial y_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial y_n}{\partial x_1} & \dfrac{\partial y_n}{\partial x_2} & \cdots & \dfrac{\partial y_n}{\partial x_n} \end{pmatrix}$$

is invertible at the point $\mathbf{p}$.

**Proof**

Let $\mu\colon W \to X$ be a local inverse of $\varphi$ around $\mathbf{p}$, where $W$ is an open set in $\mathbb{R}^n$, $\mathbf{p} \in \mu(W)$, $\mu(W) \subset X$ and $\mu(\varphi(\mathbf{x})) = \mathbf{x}$ for all $\mathbf{x} \in \mu(W)$. Suppose that $\mu\colon W \to X$ is differentiable at $\varphi(\mathbf{p})$. The identity $\mu(\varphi(\mathbf{x})) = \mathbf{x}$ holds throughout the open neighbourhood $\mu(W)$ of point $\mathbf{p}$. Applying the Chain Rule (Proposition 9.8), we find that $(D\mu)_{\varphi(\mathbf{p})}(D\varphi)_{\mathbf{p}}$ is the identity operator on $\mathbb{R}^n$. It follows that the linear operators $(D\mu)_{\varphi(\mathbf{p})}$ and $(D\varphi)_{\mathbf{p}}$ on $\mathbb{R}^n$ are inverses of one another, and therefore $(D\varphi)_{\mathbf{p}}$ is an invertible linear operator on $\mathbb{R}^n$. The result follows. ∎

**Definition**

A function $\mu\colon W \to X$ between subsets $W$ and $X$ of Euclidean spaces is said to be *Lipschitz continuous* if there exists a positive constant $C$ such that

$$|\mu(\mathbf{u}) - \mu(\mathbf{v})| \leq C|\mathbf{u} - \mathbf{v}|$$

for all $\mathbf{u}, \mathbf{v} \in W$.

It follows from Corollary 9.11 that a continuously differentiable function is Lipschitz continuous throughout some sufficiently small open neighbourhood of any given point in its domain.

### Lemma 11.3

*Let $\varphi\colon X \to \mathbb{R}^n$ be a continuously differentiable function defined over an open set $X$ in $\mathbb{R}^n$ that is locally invertible around some point of $X$ and let $\mu\colon W \to X$ be a local inverse for $\varphi$. Suppose that $\varphi\colon X \to \mathbb{R}^n$ is continuously differentiable and that the local inverse $\mu\colon W \to X$ is Lipschitz continuous throughout $W$. Then $\mu\colon W \to X$ is continuously differentiable throughout $W$.*

**Proof**

The function $\mu\colon W \to X$ is Lipschitz continuous, and therefore there exists a positive constant $C$ such that

$$|\mu(\mathbf{y}) - \mu(\mathbf{w})| \leq C\,|\mathbf{y} - \mathbf{w}|$$

for all $\mathbf{y}, \mathbf{w} \in W$. Let $\mathbf{q} \in W$, let $\mathbf{p} = \mu(\mathbf{q})$, and let $S$ be the derivative of $\varphi$ at $\mathbf{p}$. Then

$$S\mathbf{v} = \lim_{t \to 0} \frac{1}{t}(\varphi(\mathbf{p} + t\mathbf{v}) - \varphi(\mathbf{p}))$$

for all $\mathbf{v} \in \mathbb{R}^n$ (see Lemma 9.5). If $|t|$ is sufficiently small then $\mathbf{p} + t\mathbf{v} \in \mu(W)$.

It then follows from Lemma 11.1 that

$$t\mathbf{v} = \mu(\varphi(\mathbf{p} + t\mathbf{v})) - \mu(\varphi(\mathbf{p})),$$

and therefore

$$|t||\mathbf{v}| \leq C\,|\varphi(\mathbf{p} + t\mathbf{v}) - \varphi(\mathbf{p})|.$$

It follows that

$$|S\mathbf{v}| = \lim_{t\to 0} \frac{1}{|t|}|\varphi(\mathbf{p} + t\mathbf{v}) - \varphi(\mathbf{p})| \geq \frac{1}{C}|\mathbf{v}|$$

for all $\mathbf{v} \in \mathbb{R}^n$, and therefore $S\mathbf{v} \neq \mathbf{0}$ for all non-zero vectors $\mathbf{v}$. It then follows from basic linear algebra that the linear operator $S$ on $\mathbb{R}^n$ is invertible. Moreover $|S^{-1}\mathbf{v}| \leq C|\mathbf{v}|$ for all $\mathbf{v} \in \mathbb{R}^n$.

Now

$$\lim_{\mathbf{x}\to\mathbf{p}} \frac{1}{|\mathbf{x}-\mathbf{p}|} |\varphi(\mathbf{x}) - \varphi(\mathbf{p}) - S(\mathbf{x}-\mathbf{p})| = 0,$$

because the function $\varphi$ is differentiable at $\mathbf{p}$. Also $\mu(\mathbf{y}) \neq \mathbf{p}$ when $\mathbf{y} \neq \mathbf{q}$, because $\mathbf{q} = \varphi(\mathbf{p})$ and $\mathbf{y} = \varphi(\mu(\mathbf{y}))$. The continuity of $\mu$ ensures that $\mu(\mathbf{y})$ tends to $\mathbf{p}$ as $\mathbf{y}$ tends to $\mathbf{q}$. It follows that

$$\lim_{\mathbf{y}\to\mathbf{q}} \frac{1}{|\mu(\mathbf{y})-\mathbf{p}|} |\mathbf{y} - \mathbf{q} - S(\mu(\mathbf{y})-\mathbf{p})| = 0$$

(see Proposition 4.16). Now

$$|S^{-1}(\mathbf{y}-\mathbf{q}) - (\mu(\mathbf{y})-\mathbf{p})| \leq C|\mathbf{y} - \mathbf{q} - S(\mu(\mathbf{y})-\mathbf{p})|$$

for all $\mathbf{y} \in W$. Also

$$\frac{1}{|\mathbf{y}-\mathbf{q}|} \leq \frac{C}{|\mathbf{p}-\mu(\mathbf{y})|}$$

for all $\mathbf{y} \in W$ satisfying $\mathbf{y} \neq \mathbf{q}$.

It follows that

$$\frac{1}{|\mathbf{y}-\mathbf{q}|}|\mu(\mathbf{y})-\mathbf{p}-S^{-1}(\mathbf{y}-\mathbf{q})| \leq \frac{C^2}{|\mu(\mathbf{y})-\mathbf{p}|}\,|\mathbf{y}-\mathbf{q}-S(\mu(\mathbf{y})-\mathbf{p})|.$$

It follows that

$$\lim_{\mathbf{y}\to\mathbf{q}} \frac{1}{|\mathbf{y}-\mathbf{q}|}|\mu(\mathbf{y})-\mathbf{p}-S^{-1}(\mathbf{y}-\mathbf{q})| = 0$$

(see Proposition 4.9), and therefore the function $\mu$ is differentiable at $\mathbf{q}$ with derivative $S^{-1}$. Thus $(D\mu)_{\mathbf{q}} = (D\varphi)_{\mathbf{p}}^{-1}$ for all $\mathbf{q} \in W$. It follows from this that $(D\mu)_{\mathbf{q}}$ depends continuously on $\mathbf{q}$, and thus the function $\mu$ is continuously differentiable on $W$, as required. ∎

## 11.2. Convergence of Contractive Sequences

### Proposition 11.4

*Let $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ be an infinite sequence of points in n-dimensional Euclidean space $\mathbb{R}^n$, and let $\lambda$ be a real number satisfying $0 < \lambda < 1$. Suppose that*

$$|\mathbf{x}_{j+1} - \mathbf{x}_j| \leq \lambda |\mathbf{x}_j - \mathbf{x}_{j-1}|$$

*for all integers j satisfying $j > 1$. Then the infinite sequence $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ is convergent.*

**Proof**

We show that an infinite sequence of points in Euclidean space satisfying the stated criterion is a Cauchy sequence and is therefore convergent. Now the infinite sequence satisfies

$$|\mathbf{x}_{j+1} - \mathbf{x}_j| \leq C\lambda^j$$

for all positive integers $j$, where $C = |\mathbf{x}_2 - \mathbf{x}_1|/\lambda$. Let $j$ and $k$ be positive integers satisfying $j < k$. Then

$$\begin{aligned} |\mathbf{x}_k - \mathbf{x}_j| &= \left|\sum_{s=j}^{k-1} (\mathbf{x}_{s+1} - \mathbf{x}_s)\right| \leq \sum_{s=j}^{k-1} |\mathbf{x}_{s+1} - \mathbf{x}_s| \\ &\leq C\sum_{s=j}^{k-1} \lambda^s = C\lambda^j \frac{1-\lambda^{k-j}}{1-\lambda} < \frac{C\lambda^j}{1-\lambda}. \end{aligned}$$

We now show that the infinite sequence $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ is a Cauchy sequence. Let some positive real number $\varepsilon$ be given. Then a positive integer $N$ can be chosen large enough to ensure that $C\lambda^N < (1-\lambda)\varepsilon$. Then $|\mathbf{x}_k - \mathbf{x}_j| < \varepsilon$ whenever $j \geq N$ and $k \geq N$. Therefore the given infinite sequence is a Cauchy sequence. Now all Cauchy sequences in $\mathbb{R}^n$ are convergent (see Theorem 2.8). Therefore the given infinite sequence is convergent, as required. ∎