Math 125B, Winter 2015.

Discussion problems 9

Note. These are problems on the Chain Rule and the Inverse Function Theorem.

1. Assume $f : \mathbb{R}^2 \to \mathbb{R}^2$ is differentiable everywhere on $\mathbb{R}^2$. For each statement below determine, with proof, whether it is true or false for every such f.
(a) For every $a \in \mathbb{R}^2$, $D(f \circ f)(a) = Df(a)^2$.
(b) The function $f \circ f \circ f$ is differentiable.
(c) If f is one-to-one, then $\det D(f)(a) \neq 0$ for every $a \in \mathbb{R}^2$.

2. (a) Assume $m < n$ and let $g : \mathbb{R}^n \to \mathbb{R}^m$ and $f : \mathbb{R}^m \to \mathbb{R}^n$, with $g \in \mathcal{C}^1(\mathbb{R}^n)$ and $f \in \mathcal{C}^1(\mathbb{R}^m)$. Show that $D(f \circ g)$ exists on $\mathbb{R}^n$ and that $\det D(f \circ g) = 0$ on $\mathbb{R}^n$.
(b) If $f_1, f_2 : \mathbb{R}^3 \to \mathbb{R}$ are in $\mathcal{C}^1(\mathbb{R}^3)$, show that the function $F : \mathbb{R}^3 \to \mathbb{R}^3$, given by $F = (f_1, f_2, f_1 + f_2)$, does not have a differentiable local inverse at any point.

3. (a) Suppose $f : \mathbb{R}^3 \to \mathbb{R}$ is in $\mathcal{C}^1(\mathbb{R}^3)$. Let $g(x, y, z) = f(x - y, y - z, z - x)$. Show that $g : \mathbb{R}^3 \to \mathbb{R}$ is in $\mathcal{C}^1(\mathbb{R}^3)$ and that $\frac{\partial g}{\partial x} + \frac{\partial g}{\partial y} + \frac{\partial g}{\partial z} = 0$.
(b) Suppose $f : \mathbb{R}^2 \to \mathbb{R}$ is in $\mathcal{C}^1(\mathbb{R}^2)$. Let $g(x, y, z) = f(xyz, x + y + z)$. Show that $g : \mathbb{R}^3 \to \mathbb{R}$ is in $\mathcal{C}^1(\mathbb{R}^3)$ and that $x(z - y)\frac{\partial g}{\partial x} + y(x - z)\frac{\partial g}{\partial y} + z(y - x)\frac{\partial g}{\partial z} = 0$.
(c) Let f and g be as in (b). Assume $Df(0, 0) = \begin{bmatrix} 3 & 4 \end{bmatrix}$. Compute $Dg(0, 0, 0)$.

4. Define $f : \mathbb{R}^2 \to \mathbb{R}^2$ by (a) $f(x, y) = (x + y, 2xy^2)$; (b) $f(x, y) = (e^{2x+y}, x^3y^3)$. Describe the set of points (x, y) at which f has a differentiable local inverse. (Draw a graph representing points without such inverse.) Moreover, let g be the local inverse of f at $(1, 1)$ and $h : \mathbb{R}^2 \to \mathbb{R}^2$ given by $h(x, y) = (2x + 1, x + y)$. Compute $D(h \circ g)(f(1, 1))$.

5. Assume $f : \mathbb{R}^n \to \mathbb{R}^n$ is differentiable everywhere. Suppose $x_n \in \mathbb{R}^n$ and $y_n \in \mathbb{R}^n$ have a common limit $a = \lim x_n = \lim y_n$, and satisfy $f(x_n) = f(y_n)$. Assume $\det Df(a) \neq 0$. Show that $x_n \neq y_n$ only for finitely many n.

6. (a) Let $A \subset \mathbb{R}^n$ be an open set and $f : A \to \mathbb{R}^n$ one-to-one function such that $f \in \mathcal{C}^1(A)$ and $\det Df(a) \neq 0$ for all $a \in A$. Show that $f(G)$ is open for every open set $G \subset A$.
(b) Let $f : A \to \mathbb{R}$, where $A \subset \mathbb{R}^2$ is open and f is one-to-one. Let $g : A \to \mathbb{R}^2$ be given by $g(x, y) = (f(x, y), y)$. Show that $g(A)$ cannot be open.
(c) Let $f : \mathbb{R}^2 \to \mathbb{R}$. Show that if $f \in \mathcal{C}^1(\mathbb{R}^2)$, then f is not one-to-one.

7. (a) Let $A \subset \mathbb{R}^n$ be an open set and $a \in A$. Assume also that $f : A \to \mathbb{R}$ is differentiable at a. Assume that there exists an open set $U \subset A$ so that $a \in U$ and $f(x) \geq f(a)$ for all $x \in U$. In this case we say that f has a *local minimum* at a. Show that $Df(a) = 0$.
(b) Assume now $A \subset \mathbb{R}^n$ is an open set $f : A \to \mathbb{R}^n$ is differentiable on A. Assume that $b \in \mathbb{R}^n$ is

fixed. Let $g : A \to \mathbb{R}$ be defined by $g(x) = ||f(x) - b||^2$. Show that g is differentiable on A and its derivative is given by $Dg(x) = 2(f(x) - b)^T Df(x)$.

Remark. Here $(f(x) - b)^T$ is the *row* $1 \times n$ matrix made from the vector $f(x) - b \in \mathbb{R}^n$. The transpose T is used because by default we view vectors as columns (i.e., $n \times 1$) matrices.

(c) Within the same setting as in (b), assume that $\det Df(x) \neq 0$ for $x \in A$. Assume further that g has a local minimum at $x_0 \in A$. Show that $f(x) = b$.

Brief solutions

1. (a) No. The correct formula is $D(f \circ f)(a) = Df(f(a)) \cdot Df(a)$. For a concrete counterexample, take $f(x, y) = (x^2 + 1, y)$, and $a = (0, 0)$. Then $f(a) = (1, 0)$, and

$$Df(a) = Df(0,0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad Df(f(a)) = Df(1,0) = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

and $Df(a)^2 = Df(a) \neq Df(f(a)) = Df(f(a)) \cdot Df(a)$.

(b) Yes. Apply the chain rule twice.

(c) No. For example, $f(x, y) = (x^3, y)$ is one-to-one, but

$$Df(0,0) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

is not invertible.

2. As f and g are $\mathcal{C}^1$, Df and Dg exist everywhere. By the Chain Rule, $D(f \circ g)(x) = Df(g(x)) \cdot Dg(x)$, and therefore $\operatorname{rank} D(f \circ g)(x) \leq \operatorname{rank} Df(g(x)) \leq m < n$. Therefore $\det D(f \circ g) = 0$.

(b) This follows from (a) with $n = 3$, $m = 2$, where $g = (f_1, f_2)$ and $f(x, y) = (x, y, x + y)$; then $F = f \circ g$.

3. (a) $g \in \mathcal{C}^1(\mathbb{R}^3)$ by the partial derivatives version of the Chain Rule we proved in class. Then, also by the Chain Rule

$$\frac{\partial g}{\partial x} = f_x(x - y, y - z, z - x) - f_z(x - y, y - z, z - x)$$

and all other derivatives are similarly computed.

(b) This is a similar application to the chain rule as in (a).

(c) You can use the results of your computation in (b), and plug in $\frac{\partial f}{\partial x}(0, 0) = 3$ and $\frac{\partial f}{\partial y}(0, 0) = 4$. Another possibility is to use the Chain Rule directly in the matrix form: if $h(x, y, z) = (xyz, x+y+z)$, then $g = f \circ h$ and

$$Dh = \begin{bmatrix} yz & xz & xy \\ 1 & 1 & 1 \end{bmatrix}$$

and, as $h(0, 0, 0) = (0, 0)$ we have, by the Chain Rule,

$$Dg(0,0,0) = Df(0,0) \cdot Dh(0,0,0) = \begin{bmatrix} 3 & 4 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 4 & 4 \end{bmatrix}$$

4. (a) We have

$$Df = \begin{bmatrix} 1 & 1 \\ 2y^2 & 4xy \end{bmatrix}$$

and $\det Df = 4xy - 2x^2 = 2x(2y - x)$. Answer: f has differentiable local inverse at all points $(x, y) \in \mathbb{R}^2$ with $x \neq 0$ and $x \neq 2y$. The set of points where f does not have a differentiable local inverse is the union of two lines: $y = x/2$ and $x = 0$.

At $(1, 1)$, we have

$$Df(1,1) = \begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix}$$

and so $Dg(f(1,1)) = Df(1,1)^{-1}$. Moreover,

$$Dh = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}$$

at every point of $\mathbb{R}^2$. By the chain rule,

$$D(h \circ g)(f(1,1)) = Dh \cdot Dg(f(1,1)) = Dh \cdot Df(1,1)^{-1} = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix}^{-1} = \begin{bmatrix} 4 & -1 \\ 1 & 0 \end{bmatrix}$$

(b) Now,

$$Df = \begin{bmatrix} 2e^{2x+y} & e^{2x+y} \\ 3x^2y^3 & 3x^3y^2 \end{bmatrix}$$

and $\det Df = 3x^2y^2e^{2x+y}(2x-y)$. Answer: differentiable local inverse exists at all points $(x,y) \in \mathbb{R}^2$ except those on the union of three lines: $y = 2x$, $x = 0$, and $y = 0$. The last computation is analogous to that in (a).

5. There exists a neighborhood of a on which f is one-to-one. All but finitely many elements of either sequence are inside this neighborhood.

6. (a) By restricting f to G, we may assume $G = A$. Take a $b \in f(A)$. Then there is an $a \in A$ with $b = f(a)$. By the Inverse Function Theorem, there are open sets $U \subset A$ and $V \subset B$, so that $a \in U$ and $f : U \to V$ is one-to-one and onto. Therefore $f(U) = V$ and so $b \in V \subset f(A)$.
(b) Assume $g(A)$ is open. Take a point $(x_1, y_1) \in g(A)$. As $g(A)$ is open, there exists a point $(x_1, y_1') \in g(A)$, with $y_1' \neq y_1$. Then $g(x_0, y_0) = (x_1, y_1)$ and $g(x_0', y_0') = (x_1, y_1')$ for some points $(x_0, y_0), (x_0', y_0') \in A$. It follows that $f(x_0, y_0) = f(x_0', y_0') = x_1$ and, as f is one-to-one, $(x_0, y_0) = (x_0', y_0')$, and then $y_1 = y_1'$, a contradiction.
(c) We will show that the result is true even if f is defined on an open ball $A \subset \mathbb{R}^2$. Assume f is one-to-one on A. We cannot have both f_x and f_y constantly 0 on A, or else f is constant and hence not one-to-one on A. Without loss of generality, let's assume that $f_x(x_0, y_0) \neq 0$ for some $(x_0, y_0) \in A$. By continuity, $f_x \neq 0$ on a neighborhood of (x_0, y_0), which (by redefinition of A) we may assume to be the entire A. Define $g : A \to \mathbb{R}^2$ as in (b). Then

$$Df = \begin{bmatrix} f_x & f_y \\ 0 & 1 \end{bmatrix}$$

is invertible on A. By (a), $g(A)$ is open, $g(A)$ is not open. Contradition.

7. (a) Fix an $u \in \mathbb{R}^n$, $u \neq 0$. For small enough t, $t^{-1}(f(a+tu) - f(a)) \geq 0$. Thus, the directional derivative $D_u f(a) \geq 0$. But $D_u f(a) = Df(a)u$; this implies that, for all $u \neq 0$, $Df(a)u \geq 0$ and $-Df(a)u = Df(a)(-u) \geq 0$, and so $Df(a)u = 0$, i.e., $Df(a) = 0$.
(b) Let $h : \mathbb{R}^n \to \mathbb{R}$ be defined by $h(x) = ||x-b||^2 = \sum_{i=1}^n (x_i - b_i)^2$. Then $h \in \mathcal{C}^\infty(\mathbb{R}^n)$ and (by partial differentiation) $Dh(x) = 2(x-b)^T$. Moreover, $g = h \circ f$. By the Chain Rule, then

$$Dg(x) = Dh(f(x)) \cdot Df(x),$$

which is exactly the claim.
(c) By (a) and (b), $0 = Dg(x_0) = 2(f(x_0) - b)^T Df(x_0)$, but, as $Df(x_0)$ is invertible, this implies $f(x_0) - b = 0$.