Homework 5

Math 25b

Due March, 7 2018

Topics covered: Derivatives, inverse function theorem, implicit function theorem

Instructions:

- The homework is divided into one part for each CA. You will submit each part to the corresponding CA's mailbox on the second floor of the science center.
- If your submission to any one CA takes multiple pages, then staple them together. A stapler is available in the Cabot library in the science center.
- If you collaborate with other students, please mention this near the corresponding problems.
- Most problems from this assignment come from Spivak's *Calculus* or Spivak's *Calculus on manifolds* or Munkres' *Analysis on manifolds.* I've indicated this next to the problems (e.g. Spivak, CoM 1-2 means problem 2 of chapter 1 from Calculus on Manifolds).
- Any result that we proved in class can be freely used on the homework. If there's a result that we haven't stated in class that you want to use, then you have to prove it. If there's a result that we stated in class, but haven't proven, it's best to ask for clarification.

1 For Ellen

Problem 1 (Munkres, 7-3)**.** *Let* $f : \mathbb{R}^3 \to \mathbb{R}$ *and* $g : \mathbb{R}^2 \to \mathbb{R}$ *be differentiable. Let* $F : \mathbb{R}^2 \to \mathbb{R}$ *be defined by the equation*

$$F(x, y) = f(x, y, g(x, y)).$$

(a) Find DF *in terms of the partials of* f *and* g*.*

(b) If $F(x, y) = 0$ *for all* (x, y)*, find* $D_1 g$ *and* $D_2 g$ *in terms of the partials of* f*.*

Solution. (a). This can be done with the chain rule. $F = f \circ h$, where $h : \mathbb{R}^2 \to \mathbb{R}^3$ is defined by $h(x, y) = (x, y, g(x, y))$. We have

$$Df(x, y, z) = (D_1 f(x, y, z),\ D_2 f(x, y, z),\ D_3 f(x, y, z))$$

and

$$Dh(x, y) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ D_1 g(x, y) & D_2 g(x, y) \end{pmatrix}.$$

Then by the chain rule,

$$\begin{aligned} DF(x, y) &= Df(x, y, g(x, y)) \cdot Dh(x, y) \\ &= \big(D_1 f(x, y, g(x, y)) + D_3 f(x, y, g(x, y)) \cdot D_1 g(x, y), \\ &\quad D_2 f(x, y, g(x, y)) + D_3 f(x, y, g(x, y)) \cdot D_2 g(x, y)\big) \end{aligned}$$

(b). If $F \equiv 0$, then $DF(x, y) = 0$ for all (x, y). Then

$$D_1 g(x, y) = -\frac{D_1 f(x, y, g(x, y))}{D_3 f(x, y, g(x, y))} \quad \text{and} \quad D_2 g(x, y) = -\frac{D_2 f(x, y, g(x, y))}{D_3 f(x, y, g(x, y))}.$$

□

Problem 2 (Munkres, 8-1)**.** *Let* $f : \mathbb{R}^2 \to \mathbb{R}^2$ *be defined by the equation*

$$f(x, y) = (x^2 - y^2, 2xy).$$

(a) Show that f *is injective on* $A = \{(x, y) : x > 0\}$*. Hint. If* $f(x, y) = f(a, b)$*, then* $|f(x, y)| = |f(a, b)|$*.*

(b) What is the set $B = f(A)$*.*

(c) If g *is the inverse function, find* $Dg(0, 1)$*.*[1]

Solution. (a). You can compute that $|f(x, y)| = |(x, y)|$. Thus if $f(x, y) = f(a, b)$ we get the following system of equations:

$$x^2 - y^2 = a^2 - b^2$$
$$x^2 + y^2 = a^2 + b^2$$

[1] After you do this problem, it's worth thinking about how much simpler life was last semester when we were dealing with *linear* functions!

$$xy = ab.$$

If we add the first two these equations, we find $x^2 = a^2$. Since $s(x) = x^2$ is injective on $(0, \infty)$, $x^2 = a^2$ implies $x = a$. Now we use the last equation to conclude that $y = b$.

(b). We'll show the image of f is $\mathbb{R}^2 \setminus \{(a, 0) : a \leq 0\}$, i.e. $\mathbb{R}^2$ without the negative x-axis. Fix $(a, b) \in \mathbb{R}^2$. We want to solve the system of equations $x^2 - y^2 = a$ and $2xy = b$ when restricted to A. This last restriction implies $x > 0$, so we can solve $y = b/(2x)$. Then the first equation gives

$$x^2 - \frac{b^2}{4x^2} = a$$

Writing $z = x^2$, we can write this as a quadratic equation

$$z^2 - az - \frac{b^2}{4} = 0.$$

Now use the quadratic formula to find

$$x^2 = z_\pm = \frac{a \pm \sqrt{a^2 + b^2}}{2}$$

Since $x^2 > 0$, in order to have a solution we need the right-hand side to be positive. Note that $|a| \leq \sqrt{a^2 + b^2}$ with equality if and only if $b = 0$.

If $b \neq 0$, then either $z_+ = a + \sqrt{a^2 + b^2} > 0$ or $z_- = a - \sqrt{a^2 + b^2} > 0$, so we can take x to be the unique positive square root of z_+ or z_- (whichever is positive).

If $b = 0$, then $z_\pm = a + |a| = \begin{cases} 2a & a > 0 \\ 0 & a < 0 \end{cases}$. Then if $b = 0$ and $a > 0$, then $x^2 = z_+$ has a unique positive solution.

In total we've shown that $f(A) = \mathbb{R}^2 \setminus \{(a, 0) : a \leq 0\}$.

(c). Using (b), we compute $g(0, 1) = f^{-1}(0, 1) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$. Then by the chain rule

$$Dg(0, 1) = Df(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})^{-1}.$$

Now $Df(x, y) = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix}$, so $Df(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}) = \sqrt{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$. Then

$$Dg(0, 1) = \frac{1}{2\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

$\square$

2 For Charlie

Problem 3 (Munkres, 8-3)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be given by the equation $f(x) = |x|^2 \cdot x$. Show that f has derivative of all orders and that f carries the unit ball $B_1 = \{x \in \mathbb{R}^n : |x| \leq 1\}$ to itself in an injective fashion. Show however that the inverse function is not differentiable at 0.*

Solution. First note that $|f(x)| = |x|^3$ so if $|x| \leq 1$ then $|f(x)| \leq 1$, so f sends B_1 to itself. Furthermore f is injective, as we now show. First note $f(0) = 0$ if and only if $x = 0$. Furthermore, note that if $|u| = 1$ and $t \in (0, 1]$, then $f(tu) = t^3 u$. Thus if $|u| = 1 = |v|$ and $s, t \in (0, 1]$, then $f(tu) = f(sv)$ implies $t^3 u = s^3 v$. In particular $v = (t/s)^3 u$ are linearly dependent. Since $1 = |v| = |t/s|^3 |u| = |t/s|^3$ it follows that $t/s = 1$, i.e. $s = t$ and so $v = u$. This shows f is injective.

To compute Df, write f explicitly

$$f(x_1, \ldots, x_n) = (\sum_k x_k^2)(x_1, \ldots, x_n).$$

Then f has coordinates $f_i(x) = x_i \cdot \sum x_k^2$ and partial derivatives

$$D_j f_i(x) = \begin{cases} 2x_i x_j & i \neq j \\ \sum x_k^2 + 2x_i^2 & i = j \end{cases}$$

Note that these functions are polynomials, so they have derivatives of all orders. We know that for $y = su$ where $|u| = 1$ then $f^{-1}(su) = s^{1/3} u$ is the inverse of f. But $Df^{-1}(0)$ does not exist: if it did then $Df^{-1}(0) = Df(0)^{-1}$. But we know that $Df(0)$ is the zero matrix, which is not invertible. $\square$

Problem 4 (Munkres, 8-4)**.** *This exercise is a reality check of your concrete understanding of inverse function theorem (and chain rule). Define $g : \mathbb{R}^2 \to \mathbb{R}^2$ and $f : \mathbb{R}^2 \to \mathbb{R}^3$ by*

$$g(x, y) = (2ye^{2x}, xe^y) \quad \textit{and} \quad f(x, y) = (3x - y^2, 2x + y, xy + y^3).$$

(a) Show that there is a neighborhood of $(0, 1)$ that g carries in a bijective fashion to a neighborhood of $(2, 0)$.

(b) Find $D(f \circ g^{-1})$ at $(2, 0)$.

Solution. (a). We'll use the inverse function theorem. Compute

$$Dg(x, y) = \begin{pmatrix} 4ye^{2x} & 2e^{2x} \\ e^y & xe^y \end{pmatrix}$$

so $Dg(0, 1) = \begin{pmatrix} 4 & 2 \\ e & 0 \end{pmatrix}$, which is invertible, since the determinant is $-2e \neq 0$. Then by IFT, there is a neighborhood U around $(0, 1)$ so that f is invertible on U.

(b). By the chain rule,

$$D(fg^{-1})(2, 0) = Df(g^{-1}(2, 0)) \cdot Dg^{-1}(2, 0) = Df(0, 1) \cdot Dg^{-1}(2, 0).$$

We compute

$$Df(x,y) = \begin{pmatrix} 3 & -2y \\ 2 & 1 \\ y & x+3y^2 \end{pmatrix} \quad \text{and} \quad Df(0,1) = \begin{pmatrix} 3 & -2 \\ 2 & 1 \\ 1 & 3 \end{pmatrix}, \quad \text{and} \quad Dg^{-1}(2,0) = \begin{pmatrix} 0 & -2 \\ -e & 4 \end{pmatrix}.$$

Putting this all together, we get

$$D(fg^{-1})(2,0) = \begin{pmatrix} 2e & -14 \\ -e & 0 \\ -3e & 10 \end{pmatrix}$$

□

3 For Michele

Problem 5 (Munkres, 8-5). *Let A be open in $\mathbb{R}^n$, and let $f : A \to \mathbb{R}^n$ be C^1. Assume $Df(x)$ is invertible for $x \in A$. Show that $B = f(A)$ is open, even if f is not injective. Hint: You have to find the place in the proof from class where we used injectivity in a nontrivial way!*

Solution. Fix $b = f(a) \in B$. We have to find the place in the proof where we used injectivity. It occurs in the claim that there is an open rectangle Q and $d > 0$ so that every point of $f(\mathrm{bd}Q)$ is distance at least d from b. If f is not injective, we have to worry that maybe some point $x \in \mathrm{bd}(Q)$ maps to b, in which case we couldn't find such a $d > 0$.

To fix this, we'll find $\delta > 0$ so that if $x \in B_\delta(a)$ and $f(x) = b$, then $x = a$.

Since $Df(a)$ is invertible, there exists $c > 0$ so that $|Df(a) \cdot h| \geq c|h|$ for every $h \in \mathbb{R}^n$. This follows from a previous homework problem: setting $E = Df(a)$, we know there exists $M > 0$ so that $|E^{-1}u| \leq M|u|$ for every u. Setting $u = Eh$, we get $|h| \leq M|Eh|$ or $|Eh| \geq \frac{1}{M}|h|$, as desired.

Since f is differentiable, given $\epsilon = c/2$ we can find $\delta > 0$ so that $0 < |x - a| < \delta$ implies

$$\frac{|f(x) - f(a) - Df(a)(x - a)|}{|x - a|} < \epsilon.$$

Now suppose for a contradiction that $x \in B_\delta(a)$ and $f(x) = f(a)$. Then the inequality above becomes

$$\frac{|Df(a)(x - a)|}{|x - a|} < \epsilon = c/2.$$

But we also know that $|Df(a)(x - a)| \geq c|x - a|$, so in total we arrive at

$$c < \frac{|Df(a)(x - a)|}{|x - a|} < c/2,$$

a contradiction.

Once we have Q and $d > 0$, the rest of the proof goes through without any problem. □

Problem 6 (Spivak, CoM 2-37). *(a) Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a continuously differentiable function. Show that f is not 1-1. Hint: If, for example, $D_1 f(x, y) \neq 0$ for all (x, y) in some open set A, consider $g : A \to \mathbb{R}^2$ defined by $g(x, y) = (f(x, y), y)$.*

(b) Generalize this result to the case of a continuously differentiable function $f : \mathbb{R}^n \to \mathbb{R}^m$ with $m < n$. [2]

Solution. (a) Intuitively you should believe this because a linear map $\mathbb{R}^2 \to \mathbb{R}$ cannot be injective. At $z \in \mathbb{R}^2$ a differentiable function $\mathbb{R}^2 \to \mathbb{R}$ is approximated by the derivative $Df(z)$, which is linear, so we expect f to have a "1-dimensional kernel" at z. This can be made rigorous as follows.

Fix $z = 0$, and set $c = f(z)$. If $Df(w) = 0$ for all w near z, then f is constant near 0, which is as non-injective as it gets. Now assume $Df(w)$ is not identically zero near z. In particular, we can

[2] This problem is related to the Implicit Function Theorem. It can be proved using this theorem, but I'd like you to *not* quote it. It might be helpful to read the proof of the implicit function theorem in Munkres. Additional information/intuition available at office hours.

replace z with a point with $Df(z) \neq 0$. Then either $D_1 f(0) \neq 0$ or $D_2 f(0) \neq 0$ (or both). Without loss of generality, assume $D_1 f(0) \neq 0$.

Consider the function $F : \mathbb{R}^2 \to \mathbb{R}^2$ given by $F(x, y) = (f(x, y), y)$. Note that $F(0) = (c, 0)$ and

$$DF(x,y) = \begin{pmatrix} D_1 f(x,y) & D_2 f(x,y) \\ 0 & 1 \end{pmatrix},$$

so $DF(0)$ is invertible. By the inverse function theorem, there exists a neighborhood U of 0 so that $F : U \to F(U)$ is invertible with C^1 inverse $G : F(U) \to U$.

Since $F(x, y) = (f(x, y), y)$ is constant in the second variable, so is G, so G has the form $G(x, y) = (g_1(x, y), y)$. In particular,

$$G(c, y) = (g_1(c, y), y),$$

and since $F \circ G = 1$, it follows that

$$(c, y) = FG(c, y) = F(g_1(c, y), y) = (f(g_1(c, y), y), y)$$

In particular,

$$f(g_1(c, y), y) = c.$$

This formula holds for y sufficiently small, so in particular, if y is small, then $f(g_1(c, y), y) = f(0)$, which shows that f is not injective.

(b) We can generalize the argument above to show that if $f : \mathbb{R}^n \to \mathbb{R}$ is C^1, then f is not injective. If $Df(w)$ is zero in a neighborhood then f is locally constant, and otherwise, we can find z so that $Df(z) \neq 0$. Without loss of generality then $D_1 f(z) \neq 0$, and by the inverse-function-theorem argument we can write $\{f(w) = f(z)\}$ as the graph of a function $g : \mathbb{R}^{n-1} \to \mathbb{R}$ near z. In particular, this means that the set $\{w : f(w) = f(z)\}$ is nontrivial, which implies that f is not injective.

Now we can proceed inductively. Given $f = (f_1, \ldots, f_m) : \mathbb{R}^n \to \mathbb{R}^m$ with $m < n$, we can apply the argument above to $f_1 : \mathbb{R}^n \to \mathbb{R}$ to write $\{w : f_1(w) = f_1(z)\}$ as the graph of $g : \mathbb{R}^{n-1} \to \mathbb{R}$ near z. Now we're reduced to studying $(f_2, \ldots, f_m)$ on the graph of g, which is a function $\mathbb{R}^{n-1} \to \mathbb{R}^{m-1}$. This allows us to prove the result by induction. □

Problem 7 (Spivak, CoM 2-38)**.** *Define* $f : \mathbb{R}^2 \to \mathbb{R}^2$ *by* $f(x, y) = (e^x \cos y, e^x \sin y)$. *Show that* $\det f'(x, y) \neq 0$ *for all* (x, y) *but* f *is not 1-1.*

Solution. Compute

$$Df(x,y) = \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix},$$

and $\det Df(x, y) = e^x$, which is always positive. Thus f is locally injective. However, f fails to be (globally) injective: let k be any integer and consider $x = 1$ and $y_k = 2k\pi$. Then $f(x, y_k) = (e, 0)$. This shows f is not injective. □

4 For Natalia

Problem 8 (2-39)**.** *Use the function $f : \mathbb{R} \to \mathbb{R}$ defined by*

$$f(x) = \begin{cases} \frac{x}{2} + x^2 \sin \frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

to show that continuity of the derivative cannot be eliminated from the hypothesis of Theorem 2-11 (the inverse function theorem). Hint. Show that f is differentiable, but f' is not continuous, and that f is not injective in any neighborhood of 0. The last step is the tricky part – it might help to remember what f', f'' tell us about f.

Solution. For $x \neq 0$ we compute

$$f'(x) = \frac{1}{2} + 2x \sin(1/x) - \cos(1/x).$$

Using the definition of the derivative we compute $f'(0) = 1/2$. Observe that $f'(x)$ is not continuous at 0 since $\lim_{x\to 0} f'(x)$ does not exist.

Fix an integer $k > 0$. Note that if $a = (2k\pi + \frac{\pi}{2})^{-1}$ and $b = (2k\pi)^{-1}$, then $f'(a) = \frac{1}{2} + \frac{4}{(4k+1)\pi}$ and $f'(b) = -\frac{1}{2}$.

Since $f'(x)$ is continuous on $(0, \infty)$, and $f(a) > 0 > f(b)$, there exists a point $c \in (a, b)$ with $f'(c) = 0$. We claim that c is a local maximum. To show this we use the second derivative test. Compute

$$f''(x) = 2 \sin(1/x) - \frac{2}{x} \cos(1/x) - \frac{1}{x^2} \sin(1/x).$$

If $y \in [2k\pi, 2k\pi + \frac{\pi}{2}]$ then $\cos(y), \sin(y) > 0$, so it follows that for $x = \frac{1}{y}$ we have $f''(x) < 0$. In particular, this means that for k large enough we can ensure that $f'' < 0$ on (a, b) and so c is a local maximum.

Finally, since c is a local it follows that f is not injective on a neighborhood of c. Since as k goes to infinity, c will go to 0, this shows that f is not injective on any neighborhood of 0. ☐

Problem 9 (Spivak, CoM 2-40)**.** *Use the implicit function theorem to re-do Problem 2-15(c). (It should be less painful this time!)* [3]

Solution. Suppose given functions $A : \mathbb{R} \to M_n(\mathbb{R})$ and $b : \mathbb{R} \to \mathbb{R}^n$. We show that there exist a function $s : \mathbb{R} \to \mathbb{R}^n$ so that $A(t)s(t) = b(t)$ for all t and that s is differentiable.

Consider the function $F : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ given by $F(t, v) = A(t)v - b(t)$. Note that $F(t, A(t)^{-1}b(t)) = 0$. Give v coordinates $v_1, \ldots, v_n$. Write $F = (F_1, \ldots, F_n)$, where $F_i(t, v) = \sum_{r=1}^{n} A(t)_{ir} v_r - b_i(t)$. Then we can compute the $n \times n$ matrix of partial derivatives $(\partial F/\partial v)$:

$$D_j F_i(t, v) = A(t)_{ij}.$$

By assumption this is invertible for every t, so for each fixed $t_0 \in \mathbb{R}$ there exists $\epsilon > 0$ and $s : (t_0 - \epsilon, t_0 + \epsilon) \to \mathbb{R}^n$ so that $F(t, s(t)) = 0$. In fact we know that s is defined uniquely by $s(t) = A(t)^{-1}b(t)$. We conclude by the inverse function theorem that s is C^1. Furthermore, we can

[3] If you set up the implicit function theorem in the right way, this will be extremely painless.

compute the derivative by the chain rule applied to $\mathbb{R} \to \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ defined by $t \mapsto (t, s(t)) \mapsto A(t)s(t) - b(t)$. We find that

$$0 = A'(t)s(t) + A(t)s'(t) - b'(t) + A(t)s'(t),$$

which we can solve for $s'(t)$ to get

$$s'(t) = A(t)^{-1}\big[b'(t) - A'(t)s(t)\big].$$

$\square$