Problem Set 9 Solution Set

Anthony Varilly

Math 112, Spring 2002

1. *End of Chapter 6, Exercise 5.* Calculate the Jacobians of the following functions:

   (a) $f(x,y) = \sin(x^2 + y^3)$. We have

   $$\frac{\partial f}{\partial x} = \cos(x^2 + y^3) \cdot 2x \text{ and } \frac{\partial f}{\partial y} = \cos(x^2 + y^3) \cdot 3y^2$$

   Thus by Theorem 6.2.2, the Jacobian matrix is

   $$\begin{pmatrix} 2x\cos(x^2 + y^3) & 3y^2\cos(x^2 + y^3) \end{pmatrix}.$$

   □

   (f) $f(x,y) = x^{y+z}$. We have

   $$\frac{\partial f}{\partial x} = (y+z)\, x^{y+z-1} \text{ and } \frac{\partial f}{\partial y} = (\log x)\, x^{y+z} \text{ and } \frac{\partial f}{\partial z} = (\log x)\, x^{y+z}$$

   Thus by Theorem 6.2.2, the Jacobian matrix is

   $$\begin{pmatrix} (y+z)\, x^{y+z-1} & (\log x)\, x^{y+z} & (\log x)\, x^{y+z} \end{pmatrix}.$$

   □

2. *End of Chapter 6, Exercise 7.* Find the critical points of the following functions and determine whether they are local minima, maxima, or saddle points.

   (a) $f(x,y) = x^3 + 6x^2 + 3y^2 - 12xy + 9x$.

   Solution. The critical points are those points (x,y) for which

   $$\frac{\partial f}{\partial x} = 3x^2 + 12x - 12y + 9 = 0 \text{ and } \frac{\partial f}{\partial y} = 6y - 12x = 0.$$

   Therefore $y = 2x$, which implies $3(x^2 - 4x + 3) = 0$, and so $x = 1$ or $x = 3$. Therefore the critical points of f are $(3,6)$ and $(1,2)$. The matrix of the Hessian at (x,y) is

   $$\begin{pmatrix} 6x + 12 & -12 \\ -12 & 6 \end{pmatrix}.$$

At $(3,6)$ the matrix of the Hessian evaluates to

$$\begin{pmatrix} 30 & -12 \\ -12 & 6 \end{pmatrix}$$

and $\Delta_1 = 30$, $\Delta_2 = 36$. Hence, f has a local minimum at $(3,6)$.
At $(1,2)$ the matrix of the Hessian evaluates to

$$\begin{pmatrix} 18 & -12 \\ -12 & 6 \end{pmatrix}$$

and $\Delta_1 = 18$, $\Delta_2 = -36$. Hence, f has a saddle point at $(1,2)$.

□

(c) $f(x,y) = \cos 2x \cdot \sin y + z^2$.

Solution. The critical points are those points (x,y) for which

$$\frac{\partial f}{\partial x} = -2\sin 2x \cdot \sin y = 0, \quad \frac{\partial f}{\partial y} = \cos 2x \cos y = 0, \text{ and } \frac{\partial f}{\partial z} = 2z = 0$$

Clearly, it is always true that $z = 0$. From $\partial f/\partial x = 0$, either $\sin 2x = 0$, which implies $x = n\pi/2$, or $\sin y = 0$, which implies $y = m\pi$. From $\partial f/\partial y = 0$, when $x = n\pi/2$ we have $\cos y = 0$, which implies $y = (2k+1)\pi/2$, and when $y = m\pi$ we have $\cos 2x = 0$, which implies $x = (2j+1)\pi/4$. Therefore the critical points of f are $\left(\frac{n\pi}{2}, \frac{(2k+1)\pi}{2}, 0\right)$ and $\left(\frac{(2j+1)\pi}{4}, m\pi, 0\right)$, where $k, j, m, n \in \mathbb{Z}$. The matrix of the Hessian at (x,y) is

$$\begin{pmatrix} -4\cos 2x \sin y & -2\sin 2x \cos y & 0 \\ -2\sin 2x \cos y & -\cos 2x \sin y & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

At $\left(\frac{n\pi}{2}, \frac{(2k+1)\pi}{2}, 0\right)$, the matrix of the Hessian is

$$\begin{pmatrix} -4\cos n\pi \sin((2k+1)\pi/2) & 0 & 0 \\ 0 & -\cos n\pi \sin((2k+1)\pi/2) & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

It follows that $\Delta_1 = \begin{cases} -4 & \text{if } n+k \text{ even} \\ 4 & \text{if } n+k \text{ odd} \end{cases}$, $\Delta_2 = 4$, and $\Delta_3 = 8$. Hence f has local minima at $\left(\frac{n\pi}{2}, \frac{(2k+1)\pi}{2}, 0\right)$ if $n+k$ odd. If $n+k$ even, then $\Delta_3 > 0$ implies f cannot have a minimum value at $\left(\frac{n\pi}{2}, \frac{(2k+1)\pi}{2}, 0\right)$, and $\Delta_1 < 0$ implies f cannot have a maximum value at $\left(\frac{n\pi}{2}, \frac{(2k+1)\pi}{2}, 0\right)$. But since $\left(\frac{n\pi}{2}, \frac{(2k+1)\pi}{2}, 0\right)$ are critical points, they are saddle points.

At $\left(\frac{(2j+1)\pi}{4}, m\pi, 0\right)$, the matrix of the Hessian is

$$\begin{pmatrix} 0 & -2\sin\left((2j+1)\pi/2\right)\cos m\pi & 0 \\ -2\sin\left((2j+1)\pi/2\right)\cos m\pi & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

It follows that $\Delta_1 = 0$, $\Delta_2 = -4$, and $\Delta_3 = -8$. Hence f has saddle points at $\left(\frac{(2j+1)\pi}{4}, m\pi, 0\right)$ since $\Delta_2 < 0$. ☐

3. *End of Chapter 6, Exercise 8.* Show that if $f : A \in \mathbb{R}^2 \to \mathbb{R}$ has a critical point $x_0 \in A$ and we let

$$\Delta = \frac{\partial^2 f}{\partial x_1 \partial x_1} \cdot \frac{\partial^2 f}{\partial x_2 \partial x_2} - \left(\frac{\partial^2 f}{\partial x_1 \partial x_2}\right)^2$$

be evaluated at x_0, then

(a) $\Delta > 0$ and $\partial^2 f/\partial x_1 \partial x_1 > 0$ imply f has a local minimum at x_0.

Solution. From Theorem 6.9.4, f has a local minimum at x_0 if x_0 is a critical point of f such that $H_{x_0}(f)$ is positive definite. Thus we need only show that $H_{x_0}(f)$ is positive definite if $\Delta > 0$ and $\partial^2 f/\partial x_1 \partial x_1 > 0$. To be sure, we prove the following lemma:

Lemma: Let $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ be a symmetric 2×2 matrix, where $a, b, c \in \mathbb{R}$. Then A is positive definite if $a > 0$ and $ac - b^2 > 0$. A is negative definite if $a < 0$ and $ac - b^2 > 0$.

Proof. Let $v = (x, y)^\top$ be an arbitrary non-zero vector. Then

$$\begin{aligned} v^\top A v &= \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = ax^2 + 2bxy + cy^2 \\ &= ax^2 + 2bxy + \frac{b^2y^2}{a} - \frac{b^2y^2}{a} + cy^2 \\ &= a\left(x + \frac{by}{a}\right)^2 + \left(c - \frac{b^2}{a}\right)y^2 \end{aligned}$$

If A is to be positive definite, then $v^\top Av > 0$ for all v. In particular, for $y = 0$, we must have $ax^2 > 0$, which implies $a > 0$. Also, when $x = -(b/a)y$, $(c - b^2/a)y^2 > 0$ implies $ac - b^2 > 0$. If A is to be negative definite, then $v^\top Av < 0$ for all v. In particular, for $y = 0$, we must have $ax^2 < 0$, which implies $a < 0$. Also, when $x = -(b/a)y$, $(c - b^2/a)y^2 < 0$ implies $a(c - b^2/a) > 0$, which yields $ac - b^2 > 0$. ☐

Since the matrix of the Hessian clearly satisfies the conditions of this lemma, it follows that $H_{x_0}(f)$ is positive definite if $\Delta = ac - b^2 > 0$ and $\partial^2 f/\partial x_1 \partial x_1 = a > 0$. ☐

(b) $\Delta > 0$ and $\partial^2 f/\partial x_1 \partial x_1 < 0$ imply f has a local maximum at x_0.

Solution. Since the matrix of the Hessian clearly satisfies the conditions of the lemma, it follows that $H_{x_0}(f)$ is negative definite if $\Delta = ac - b^2 > 0$ and $\partial^2 f/\partial x_1 \partial x_1 = a < 0$. ☐

(c) $\Delta < 0$ implies x_0 is a saddle point of f.

Solution. If x_0 is not a local max or min, then it must be a saddle point. Similarly, if it is not the case that $\Delta > 0$ and $\partial^2 f/\partial x_1 \partial x_1 > 0$ or $\Delta > 0$ and $\partial^2 f/\partial x_1 \partial x_1 < 0$, then it must be true that $\Delta < 0$. Therefore, $\Delta < 0$ implies x_0 is a saddle point of f. □

4. *End of Chapter 6, Exercise 12.* A function $f : \mathbb{R}^n \to \mathbb{R}$ is called homogenous of degree m if $f(tx) = t^m f(x)$ for all $x \in \mathbb{R}^n$ and $t \in \mathbb{R}$. If f is differentiable, show that for $x \in \mathbb{R}^n$,

$$\mathbf{D}f(x)\,x = mf(x)\,, \text{ that is, } \sum_{i=1}^{n} x_i \frac{\partial f}{\partial x_i} = mf(x)\,.$$

Show that maps multilinear in k variables give rise to homogenous functions of degree k. Give other examples.

Solution. By definition of the directional derivative,

$$\mathbf{D}f(x)\,x = \lim_{h\to 0} \frac{f(x+hx) - f(x)}{h} = \lim_{h\to 0} \frac{f((1+h)\,x) - f(x)}{h}$$

Using the fact that f is homogeneous of degree m, we get

$$\begin{aligned} \mathbf{D}f(x)\,x &= \lim_{h\to 0} \frac{(1+h)^m f(x) - f(x)}{h} = \lim_{h\to 0} \left(\frac{(1+h)^m - 1}{h} \right) f(x) \\ &= \lim_{h\to 0} \left(\frac{1^m + \binom{m}{1}h + \binom{m}{2}h^2 + \cdots + \binom{m}{m}h^m - 1}{h} \right) f(x) \\ &= \lim_{h\to 0} \left(m + \binom{m}{2}h + \cdots + \binom{m}{m}h^{m-1} \right) f(x) = mf(x). \end{aligned}$$

as desired.

k-linear maps are characterized by the property

$$\begin{aligned} & L(x_1, \ldots, x_{i-1}, \alpha u + \beta w, x_{i+1}, \ldots, x_n) \\ = \ & \alpha L(x_1, \ldots, x_{i-1}, u, x_{i+1}, \ldots, x_n) + \beta(x_1, \ldots, x_{i-1}, w, x_{i+1}, \ldots, x_n) \end{aligned}$$

If we define $g(x) = L(\underbrace{x, \ldots, x}_{k \text{ times}})$, then it follows that

$$g(tx) = L(tx, \ldots, tx) = t^k L(x, \ldots, x) = t^k g(x).$$

Therefore, maps multilinear in k variables give rise to homogenous functions of degree k.

An example of a non-linear homogenous function is $f(x, y) = x^2 + y^2$. This is homogeneous of degree 2 since $f(kx, ky) = k^2 \left(x^2 + y^2\right) = k^2 f(x, y)$. □

5. *End of Chapter 6, Exercise 13.* Use the chain rule to find derivatives of the following, where $f(x, y, z) = x^2 + yz$, $g(x, y) = y^3 + xy$, and $h(x) = \sin x$:

(a) $F(x, y, z) = f(h(x), g(x, y), z)$.

Solution. Here $f(h, g, z) = h^2 + gz$. The chain rule gives

$$\begin{aligned}
\frac{\partial F}{\partial x} &= \frac{\partial f}{\partial h}\frac{\partial h}{\partial x} + \frac{\partial f}{\partial g}\frac{\partial g}{\partial x} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial x} = 2\sin x\cos x + zy + 0 \\
\frac{\partial F}{\partial y} &= \frac{\partial f}{\partial h}\frac{\partial h}{\partial y} + \frac{\partial f}{\partial g}\frac{\partial g}{\partial y} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial y} = 0 + z(3y^2 + x) + 0 \\
\frac{\partial F}{\partial z} &= \frac{\partial f}{\partial h}\frac{\partial h}{\partial z} + \frac{\partial f}{\partial g}\frac{\partial g}{\partial z} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial z} = 0 + 0 + (y^3 + xy)
\end{aligned}$$

Therefore $\mathbf{D}F(x, y, z) = \begin{pmatrix} \sin 2x + yz & xz + 3y^2 z & y^3 + xy \end{pmatrix}$.
Alternatively, we can use Jacobean matrices: $\mathbf{D}F(x, y, z) = \mathbf{D}F(f(x)) \circ \mathbf{D}f(\mathbf{x})$. In this case

$$\mathbf{D}F(x, y, z) = \begin{pmatrix} 2h & z & g \end{pmatrix} \cdot \begin{pmatrix} \cos x & 0 & 0 \\ y & 3y^2 + x & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and we get the same answer as before. □

6. *End of Chapter 6, Exercise 15.* Let $f : \mathbb{R} \to \mathbb{R}$ be differentiable. Assume there is no $x \in \mathbb{R}$ such that f and f' both vanish at x. Show that $S = \{x \mid 0 \leq x \leq 1,\ f(x) = 0\}$ is finite.

   Solution. Assume that S is infinite. Then by the Bolzano-Weierstrass theorem, S has an accumulation point $c \in [0, 1]$. So there are x_n in S with $x_n \to c$. Since f is differentiable, it is continuous, which implies that $\lim_{n\to\infty} f(x_n) = f(c) = 0$. Then using the definition of the derivative,
   $$f'(c) = \lim_{n\to\infty} \frac{f(x_n) - f(c)}{x_n - c} = \lim_{n\to\infty} \frac{0 - 0}{x_n - c} = 0$$
   This last equality follows because $x_n \in S$. But this contradicts the assumption that there is no $x \in \mathbb{R}$ such that f and f' both vanish at x. Thus S must be finite. □

7. *End of Chapter 6, Exercise 18.* Prove that the equation $x^3 + bx + c = 0$ where $b > 0$ has exactly one solution $x \in \mathbb{R}$.

   Solution. Let $f(x) = x^3 + bx + c$, where $b > 0$. From the Intermediate Value Theorem, we know that f must have at least one root since it is an odd polynomial. Since $f'(x) = 3x^2 + b > 0$, f is an increasing function for all $x \in \mathbb{R}$. We now claim that $f(x) = 0$ only once. For if there are x_1, x_2 such that $f(x_1) = f(x_2) = 0$, then by Rolle's Theorem, there exists $\xi \in (x_1, x_2)$ such that $f'(\xi) = 0$, which contradicts the fact that $f'(x) > 0$ for all $x \in \mathbb{R}$. □

8. *End of Chapter 6, Exercise 20.* Let $L : \mathbb{R}^n \to \mathbb{R}^n$ be a linear map. Define $\|L\| = \inf\{M \mid \|Lx\| \leq M\|x\|$ for
   Show that $\|\cdot\|$ is a norm on the space of linear maps of $\mathbb{R}^n$ to $\mathbb{R}^m$.

Solution. Let $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ denote the space of linear maps from $\mathbb{R}^n$ to $\mathbb{R}^m$. We first show that $\|L_1\| \geq 0$ for $L_1 \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. We have

$$\|L_1\| = \inf\{M \mid \|L_1 x\| \leq M\|x\|\} = \inf\{M \mid \|y\| \leq M\|x\|\}$$

if we let $L_1 x = y \in \mathbb{R}^m$. From properties of the vector norm, $\|y\| \geq 0$ and $\|x\| \geq 0$, so it follows that $M \geq 0$ and thus $\inf M \geq 0$.

Secondly, we show that $\|L_1\| = 0$ if and only if $L_1 = 0$. We have

$$\begin{aligned} \|L_1\| &= 0 \iff \inf\{M \mid \|L_1 x\| \leq M\|x\|\} = 0 \\ &\iff \|L_1 x\| \leq 0 \text{ for all } x \in \mathbb{R}^n \\ &\iff \|L_1 x\| = 0 \text{ for all } x \in \mathbb{R}^n \iff L_1 = 0. \end{aligned}$$

Thirdly, we show that for $\alpha \in \mathbb{R}$, we have $\|\alpha L_1\| = |\alpha|\|L_1\|$. We have

$$\begin{aligned} \|\alpha L_1\| &= \inf\{M \mid \|\alpha L_1 x\| \leq M\|x\|\} = \inf\{M \mid \|\alpha y\| \leq M\|x\|\} \\ &= \inf\left\{M \mid \|y\| \leq \frac{M}{|\alpha|}\|x\|\right\} = \inf\{|\alpha| N \mid \|\alpha y\| \leq N\|x\|\} \\ &= |\alpha| \inf\{N \mid \|\alpha y\| \leq N\|x\|\} = |\alpha|\|L_1\|. \end{aligned}$$

Finally for $L_1, L_2 \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, we show that $\|L_1 + L_2\| \leq \|L_1\| + \|L_2\|$. Let $\|L_1\| = \inf\{M_1 \mid \|L_1 x\| \leq M_1\|x\|\}$, $\|L_2\| = \inf\{M_2 \mid \|L_2 x\| \leq M_2\|x\|\}$, and $\|L_1 + L_2\| = \inf\{M \mid \|(L_1 + L_2)x\| \leq$ Note that

$$\begin{aligned} \|(L_1 + L_2)x\| &= \|L_1 x + L_2 x\| \\ &\leq \|L_1 x\| + \|L_2 x\| \leq M_1\|x\| + M_2\|x\| \end{aligned}$$

For all M_1 and M_2 for which the above inequalities $\|L_1 x\| \leq M_1\|x\|$ and $\|L_2 x\| \leq M_2\|x\|$ hold, they must be valid for $\inf M_1$ and $\inf M_2$. So we get

$$\begin{aligned} \|(L_1 + L_2)x\| &\leq \inf\{M_1 \mid \|L_1 x\| \leq M_1\|x\|\} + \inf\{M_2 \mid \|L_2 x\| \leq M_2\|x\|\} \\ &= \|L_1\| + \|L_2\| \end{aligned}$$

It is now convenient to redefine $\|L\| = \sup\{M \mid \|Lx\| \geq M\|x\| \text{ for all } x \in \mathbb{R}^n\}$. Then $\|L_1 + L_2\| = \sup\{M \mid \|(L_1 + L_2)x\| \geq M\|x\|\}$. So we now have

$$M\|x\| \leq \|(L_1 + L_2)x\| \leq \|L_1\| + \|L_2\|$$

For all M for which the above inequality holds, it must hold for $\sup M$. Thus we have

$$\sup\{M \mid \|(L_1 + L_2)x\| \geq M\|x\|\} \leq \|L_1\| + \|L_2\|$$

which implies

$$\|L_1 + L_2\| \leq \|L_1\| + \|L_2\|$$

as desired. □

9. *End of Chapter 6, Exercise 29.* Let $f_n(x) = xe^{-nx}$, $x \in [0, \infty)$, $n = 0, 1, 2, \ldots$.

(a) Show that $f(x) = \sum_{n=0}^{\infty} f_n(x)$ exists. Compute f explicitly.

Solution. $f(x) = \sum_{n=0}^{\infty} xe^{-nx} = x\sum_{n=0}^{\infty}(e^{-x})^n$. For $x > 0$, $|e^{-x}| < 1$, so we have a convergent geometric series which sums to $1/(1-e^{-x}) = e^x/(e^x - 1)$. Therefore $f(x) = xe^x/(e^x - 1)$ for $x > 0$. At $x = 0$, $f(0) = \sum_{n=0}^{\infty} f_n(0) = 0$. □

(b) Is f continuous?

Solution. No, f is not continuous because

$$\lim_{x\to 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x\to 0} \frac{e^x}{e^x - 1} = 1$$

using l'Hôpital's Rule. Therefore $f(x) \nrightarrow f(0)$ as $x \to 0$ and so f is not continuous. □

(c) Find a suitable set on which the convergence is uniform.

Solution. Consider the interval $[a, b]$, where $0 < a < b$. Clearly $x \le b$ and $e^{-x} \le e^{-a}$. Therefore $f_n(x) = xe^{-nx} \le be^{-an}$. If we let $M_n = be^{-an}$, then $\sum_{n=0}^{\infty} M_n = b\sum_{n=0}^{\infty}(e^{-a})^n$ is a convergent geometric series since $a > 0$ implies $e^{-a} < 1$. By the Weierstrass M test, we have uniform convergence on the interval $[a, b]$. But since a and b are arbitrary positive numbers, we have shown uniform convergence on $[a, \infty)$, $a > 0$. □

(d) May we differentiate term by term?

Solution. The f_n are clearly differentiable for $x \in (0, \infty)$, with $f_n'(x) = (1 - nx)e^{-nx}$. From this, we see that the f_n' are continuous for $x \in (0, \infty)$. We have already showed that $\sum f_n$ converges uniformly (hence pointwise) on $(0, \infty)$.
We will now show that $\sum f_n'$ converges uniformly on $(0, \infty)$. Consider the interval $[a, b]$, where $0 < a < b$. We have $\sum_{n=0}^{\infty} f_n'(x) = \sum_{n=0}^{\infty} e^{-nx} - \sum_{n=0}^{\infty} nxe^{-nx}$. For $x \in [a, b]$, $e^{-nx} < 1$, and so $\sum_{n=0}^{\infty} e^{-nx}$ is a convergent geometric series. In Example 5.1.9, this was shown to imply uniform convergence on $[a, b]$. For the series $\sum_{n=0}^{\infty} nxe^{-nx}$, take $M_n = nbe^{-an}$, similar to part (c). Then $\sum_{n=0}^{\infty} M_n = b\sum_{n=0}^{\infty} ne^{-an}$ is seen to be convergent in $[a, b]$ by an easy application of the ratio test. So by the Weierstrass M test, $\sum_{n=0}^{\infty} nxe^{-nx}$ converges uniformly on $[a, b]$. Thus, since $\sum_{n=0}^{\infty} f_n'(x)$ is the difference between two uniformly convergent series, it is also uniformly convergent on $[a, b]$. But since a and b are arbitrary positive numbers, we have shown uniform convergence on $(0, \infty)$. □

Therefore, we conclude that term-by-term differentation is valid.

10. *End of Chapter 6, Exercise 35.* Let $f : (a, b) \to \mathbb{R}$ be twice differentiable. Suppose f vanishes at three distinct points. Prove that there is a $c \in (a, b)$ such that $f''(c) = 0$.

Solution. Suppose $f(x_1) = f(x_2) = f(x_3) = 0$. WLOG, let $x_1 < x_2 < x_3$. Then by Rolle's Theorem, there exists $c_1 \in (x_1, x_2)$ and $c_2 \in (x_2, x_3)$ such that $f'(c_1) = f'(c_2) = 0$. Using Rolle's Theorem again, there exists $c \in (c_1, c_2)$ such that $f''(c) = 0$. □