Problem Set #08- Answer key

Second Calculus Problem Set

(1) Consider the function $f(x, y, z) = xyz$, with $y = x^2$ and $z = x^{1/3}$.

(a) Rewrite f as a function $g : \mathbb{R} \to \mathbb{R}$ alone and compute $g'(\cdot)$. Using g', approximate the change in f when x increases by 0.1 units, starting from $(8, 64, 2)$.

Ans:

$$\begin{aligned} g(x) &= f(x, x^2, x^{1/3}) = x \times x^2 \times x^{1/3} \\ &= x^{10/3} \quad \text{so that} \quad g'(x) = 10/3x^{7/3} \\ g'(8) &= 10/3 \times 8^{7/3} = 10/3 * 2^7 = 426.6667 \\ dg &= df = 0.1 \times 10/3 \times 128 = 42.6. \end{aligned}$$

(b) Compute the total derivative of f with respect to x. Using the total derivative, approximate the change in f when x increases by 0.1 units, starting from $(8, 64, 2)$.

Ans:

$$\begin{aligned} \frac{df}{dx} &= f_x + f_y \frac{dy}{dx} + f_z \frac{dz}{dx} = yz + xz \times 2x + yx \times 1/3x^{-2/3} \\ &= x^{7/3} + x^{4/3} \times 2x + x^3 \times 1/3x^{-2/3} \\ &= x^{7/3}(1 + 2 + 1/3) = 10/3x^{7/3} \\ \text{so that} \quad df &= \frac{df(8)}{dx} = 0.1 \times 10/3 \times 8^{7/3} = 42.6 \end{aligned}$$

(c) Write down the differential of f at $(8, 64, 2)$. Using the differential, approximate the change in f when x increases by 0.1 units, starting from $(8, 64, 2)$.

Ans:

$$\bigtriangledown f(x,y,z) = \begin{bmatrix} yz & xz & yx \end{bmatrix} \text{ so that } \bigtriangledown f(8,64,2) = \begin{bmatrix} 128 & 16 & 512 \end{bmatrix}$$

$$df = \bigtriangledown f(x,y,z) \begin{bmatrix} dx \\ y'(x)dx \\ z'(x)dx \end{bmatrix}' = \begin{bmatrix} 128 & 16 & 512 \end{bmatrix} \begin{bmatrix} 0.1 \\ 16 \times 0.1 \\ 0.0833 \times 0.1 \end{bmatrix}$$

$$= 12.8 + 16 \times 0.16 + 512 \times 0.025 = 42.7$$

(d) Identify the direction h^* that (x,y,z) moves in, starting from $(8,64,2)$, when x increases. Write down the directional derivative of f in the direction $\mathbf{h}^*$, i.e., $f_{\mathbf{h}^*}(\cdot,\cdot,\cdot)$, and evaluate this derivative at $(8,64,2)$. Using $f_{\mathbf{h}^*}(8,64,2)$, approximate the change in f when x increases by 0.1 units, starting from $(8,64,2)$.

Ans: When x increases by one, the vector (x,y,z) increases in the direction $(dx, y'(x)dx, z'(x)dx)$. When $x = 8$ and $dx = 1$, therefore, (x,y,z) increases in the direction $h^* = (1, 16, 0.0833)$. The *unit length* vector pointing in this direction is $h^*/||h^*|| = (0.0624, 0.9980, 0.0052)$. Therefore, using the differential to compute the directional derivative, $f_{\mathbf{h}^*}(8,64,2) = \begin{bmatrix} 128 & 16 & 512 \end{bmatrix} \begin{bmatrix} 0.0624 \\ 0.9980 \\ 0.0052 \end{bmatrix} = 26.613$. Now a dx of 0.1 induces a shift in $\mathbb{R}^3$ of $0.1h^*$, which has length $0.1||h^*||$. $df = f_{\mathbf{h}^*}(8,64,2) \times 0.1||h^*|| = 42.67$

(e) Check to see that all four of these distinct methods give you the same answer!

Ans: Amazingly, they do!

(2) Recall that a function $f : \mathbb{R}^n \to \mathbb{R}^m$ is nothing more than m functions, $f^1 ... f^m$, each mapping $\mathbb{R}^n \to \mathbb{R}$, and stacked on top of each other.

(a) Using this fact, write down a formal definition of the directional derivative of f at x_0 in the direction $\mathbf{h} \in \mathbb{R}^n$, for a function $f : \mathbb{R}^n \to \mathbb{R}^m$. Your definition should be of the form

$$\text{blah, blah} \quad = \quad \lim_{k\to\infty} \frac{\text{blah} \qquad\qquad \text{blah}}{\text{blah, blah}}$$

Ans: Definition: Given $f : \mathbb{R}^n \to \mathbb{R}^m$ and $\mathbf{h} \in \mathbb{R}^n$, the directional derivative of f at $\mathbf{x}_0$ in the direction $\mathbf{h}$ is given by, for $i = 1, ...m$,

$$f^i_{\mathbf{h}} \quad = \quad \lim_{|\mathbf{k}|\to\infty} \frac{\left(f^i(\mathbf{x}_0 + \mathbf{h}/k) - f^i(\mathbf{x}_0)\right)}{||\mathbf{h}||/k}$$

(b) Consider the function $f : \mathbb{R}^2 \to \mathbb{R}^2$ defined by, for $i = 1, 2$, $f^i(x, y) = x^{i/3}y^{1-i/3}$. Using the formal definition in (a) above, compute $f_{\mathbf{h}^*}(27, 8)$, where $\mathbf{h}^* = (54, 16)$. (Hint: $(27, 8) + (54, 16)/k = (27, 8)(1 + 2/k)$).

Ans:

$$f^i_{\mathbf{h}^*}(27,8) \quad = \quad \lim_{|\mathbf{k}|\to\infty} \frac{\left(f^i\left((27,8) + (54,16)/k\right) - f^i(27,8)\right.}{||\mathbf{h}||/k}$$

$$= \quad \lim_{|\mathbf{k}|\to\infty} \frac{\left(f^i\left((27,8)(1+2/k)\right) - f^i(27,8)\right.}{||\mathbf{h}||/k}. \qquad \text{Therefore}$$

$$f^1_{\mathbf{h}^*}(27,8) \quad = \quad \lim_{|\mathbf{k}|\to\infty} \frac{\left((3\times 4)(1+2/k)\right) - (3\times 4)}{||\mathbf{h}||/k} \quad = \quad \frac{24}{||\mathbf{h}||}. \qquad \text{Similarly}$$

$$f^2_{\mathbf{h}^*}(27,8) \quad = \quad \lim_{|\mathbf{k}|\to\infty} \frac{\left((9\times 2)(1+2/k)\right) - (9\times 2)}{||\mathbf{h}||/k} \quad = \quad \frac{36}{||\mathbf{h}||} \qquad \text{Therefore}$$

$$f_{\mathbf{h}^*}(27,8) \quad = \quad \begin{bmatrix} 0.4261 \\ 0.6392 \end{bmatrix}$$

(c) Now compute $f_{\mathbf{h}^*}(27, 8)$ using the differential of f at $(27, 8)$.

Ans:

$$\begin{aligned}
\bigtriangledown f^i(x,y) &= \begin{bmatrix} \frac{if^i(x,y)}{3x} & \frac{(3-i)f^i(x,y)}{3y} \end{bmatrix} \\
f^i_{\mathbf{h}^*}(27,8) &= \begin{bmatrix} \frac{if^i(x,y)}{3x} & \frac{(3-i)f^i(x,y)}{3y} \end{bmatrix} \begin{bmatrix} 54/||\mathbf{h}^*|| \\ 16/||\mathbf{h}^*|| \end{bmatrix} \qquad \text{Therefore} \\
\begin{bmatrix} f^1_{\mathbf{h}^*}(27,8) \\ f^2_{\mathbf{h}^*}(27,8) \end{bmatrix} &= \begin{bmatrix} \begin{bmatrix} \frac{12}{3\times 27} & \frac{2\times 12}{3\times 8} \end{bmatrix} \begin{bmatrix} 54/||\mathbf{h}^*|| \\ 16/||\mathbf{h}^*|| \end{bmatrix} \\ \begin{bmatrix} \frac{2\times 18}{3\times 27} & \frac{18}{3\times 8} \end{bmatrix} \begin{bmatrix} 54/||\mathbf{h}^*|| \\ 16/||\mathbf{h}^*|| \end{bmatrix} \end{bmatrix} \\
&= \frac{1}{||\mathbf{h}^*||} \begin{bmatrix} 8+16 \\ 24+12 \end{bmatrix} \quad = \quad \begin{bmatrix} 0.4261 \\ 0.6392 \end{bmatrix}
\end{aligned}$$

(d) Check to see that all of these three distinct methods give you the same answer!

Ans: Amazingly, they do!

(3) Consider the function $f(\mathbf{x}) = x_1^\rho + x_2^\rho$, where $\rho \in (-\infty, 1]$. *The whole point here is to use the differential of* $\triangledown f$ *to answer the following questions, ie., to answer all parts of the question, approximate* $\triangledown f(\mathbf{x}+\mathbf{h}) - \triangledown f(\mathbf{x})$ *using the differential of* $\triangledown f$ *at* $\mathbf{x}$*, evaluated at* $\mathbf{h}$*. There are lots of other ways to answer these questions, but the purpose of this question is to give you practice in using the differential of a vector-valued function.*

(a) Check that, *up to a first order approximation*,[1] f is homothetic (cf the notes for lecture CALCULUS3,[2] specifically the second example in the subsection entitled Four Graphical Examples.[3])

Ans: The following argument is not entirely rigorous, but can be made so. We need to show that if $\mathbf{x}$ and $\mathbf{dx}$ are colinear, (i.e., if $\mathbf{dx} = \lambda\mathbf{x}$), then $\triangledown f(\mathbf{x})$ and $\triangledown f(\mathbf{x}+\mathbf{dx})$ are colinear, i.e., $\triangledown f(\mathbf{x}+\mathbf{dx}) = \delta \triangledown f(\mathbf{x})$, for some $\delta > 0$.
Now $\triangledown f(\mathbf{x}) = \begin{bmatrix} \rho x_1^{\rho-1} & \rho x_2^{\rho-1} \end{bmatrix}$ so that $Hf(\mathbf{x}) = \begin{bmatrix} \rho(\rho-1)x_1^{\rho-2} & 0 \\ 0 & \rho(\rho-1)x_2^{\rho-2} \end{bmatrix}$ Therefore, for $\mathbf{dx} = \lambda\mathbf{x}$),

$$\begin{aligned}
\triangledown f(\mathbf{x}+\mathbf{dx}) &\approx \triangledown f(\mathbf{x}) + \begin{bmatrix} \rho(\rho-1)x_1^{\rho-2} & 0 \\ 0 & \rho(\rho-1)x_2^{\rho-2} \end{bmatrix} \begin{bmatrix} dx_1 \\ dx_2 \end{bmatrix} \\
&= \begin{bmatrix} \rho x_1^{\rho-1} \\ \rho x_2^{\rho-1} \end{bmatrix} + \begin{bmatrix} \rho(\rho-1)x_1^{\rho-2} & 0 \\ 0 & \rho(\rho-1)x_2^{\rho-2} \end{bmatrix} \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \end{bmatrix} \\
&= \begin{bmatrix} \rho x_1^{\rho-1} \\ \rho x_2^{\rho-1} \end{bmatrix} + \begin{bmatrix} \lambda\rho(\rho-1)x_1^{\rho-1} \\ \lambda\rho(\rho-1)x_2^{\rho-1} \end{bmatrix} = \rho(1+\lambda(\rho-1)) \begin{bmatrix} x_1^{\rho-1} \\ x_2^{\rho-1} \end{bmatrix} \\
&= (1+\lambda(\rho-1)) \triangledown f(\mathbf{x})
\end{aligned}$$

The argument is not entirely rigorous because of the approximation relationship above. To make it fully rigorous, we could use the 2-dimensional analog of the Taylor Lagrange theorem.

[1] The qualifier "up to a first order approximation" means: you should *pretend* that the answer you get using the differential is exactly correct, even though in fact it is only approximately correct, and then only for small $\mathbf{h}$'s, because there are non-zero higher order terms in the Taylor expansion of $\triangledown f$.

[2] In the example in the notes, you don't need the caveat about up to a first order approximation, because the higher order terms in the Taylor approximation are all zero. In this example they are not.

[3] The lecture notes tend to change, and sometimes the problem sets don't keep up. If this reference is no longer current, please notify Leo.

(b) When $\rho > 0$, does f exhibit increasing, constant or decreasing returns to scale? Is your answer true for all $\rho \in (0, 1]$. (Again, your answer should be in terms of what happens to the *gradient vector* as you move out along a ray.)

Ans: We'll show that f exhibits decreasing returns to scale whenever $0 < \rho < 1$. When $\rho = 0$, the function is flat; when $\rho < 0$, the function decreases so that the idea of "returns to scale" doesn't mean much. For $\rho > 0$, it is sufficient to show that as you increase scale, i.e., move from $\mathbf{x}$ to $(1 + \lambda)\mathbf{x}$, for any $\lambda > 0$, the length of the gradient vector shrinks. From the answer to the preceding part, note that for $\mathbf{dx} = \lambda\mathbf{x}$, and $\lambda > 0$, $\triangledown f(\mathbf{x} + \mathbf{dx}) \approx (1 + \lambda(\rho - 1)) \triangledown f(\mathbf{x})$. For $\rho < 1$, $(1 + \lambda(\rho - 1)) < 1$, so the gradient vector indeed shrinks. When $\rho = 1$, $(1 + \lambda(\rho - 1)) = 1$, so the gradient vector remains constant, which is equivalent to constant returns to scale.

(c) Fix $\mathbf{x} = (\alpha, \alpha)$, and consider $\mathbf{h} = (-0.1, 0.1)$. Approximate $\triangledown f(\mathbf{x} + \mathbf{h})$, for (i) $\rho = 1/2$; (ii) $\rho = -1/2$; (iii) $\rho = -10$.

Ans:

$$\begin{aligned}
\triangledown f(\mathbf{x} + \mathbf{h}) &\approx \rho\alpha^{\rho-1}\begin{bmatrix}1\\1\end{bmatrix} + \begin{bmatrix}\rho(\rho-1)\alpha^{\rho-2} & 0\\ 0 & \rho(\rho-1)\alpha^{\rho-2}\end{bmatrix}\begin{bmatrix}-0.1\\0.1\end{bmatrix}\\
&= \begin{cases}
\frac{1}{2}\alpha^{-1/2}\left\{\begin{bmatrix}1\\1\end{bmatrix} + 0.1\begin{bmatrix}\frac{1}{2}\alpha^{-1}\\-\frac{1}{2}\alpha^{-1}\end{bmatrix}\right\} & \text{if } \rho = 1/2\\
-\frac{1}{2}\alpha^{-3/2}\left\{\begin{bmatrix}1\\1\end{bmatrix} + 0.1\begin{bmatrix}-\frac{3}{2}\alpha^{-1}\\\frac{3}{2}\alpha^{-1}\end{bmatrix}\right\} & \text{if } \rho = -1/2\\
-10\alpha^{-11}\left\{\begin{bmatrix}1\\1\end{bmatrix} + 0.1\begin{bmatrix}-11\alpha^{-1}\\11\alpha^{-1}\end{bmatrix}\right\} & \text{if } \rho = -10
\end{cases}
\end{aligned}$$

(d) How does the curvature of the level sets of this function change as you move out along a ray through the origin. In particular, discuss the effect of the magnitude of α on the *rate of change* in the direction of $\triangledown f$ as you add $\mathbf{h} = (-\beta, \beta)$ to $\mathbf{x} = (\alpha, \alpha)$.

Ans: Consider the answer to part (c): in each case, the magnitude of the rotation is determined by α^{-1}. When α is very large, the rotation is miniscule, when α is miniscule, the rotation is huge. Thus as you move out along a ray through the origin, the change in curvature generated by a *given* shift $\mathbf{h} = (-\beta, \beta)$ becomes less and less; that is, level sets become flatter and flatter as you move out along any ray.

(4) Suppose that $f : \mathbb{R} \to \mathbb{R}$ is (n+1) times continously differentiable.

(a) Show that a sufficient condition for f to a attain a strict (local) maximum at x_0 is that for some even number n. the derivatives $f^{(k)}(x_0)$ are zero for $k = 1...n-1$, and $f^{(n)}(x_0)$ is negative

Ans:

(1) We are given that n is even. Hence $(x - x_0)^n > 0$ for $x \neq x_0$
(2) Since $f^{(n)}(x_0) < 0$ and the first $n-1$ terms in the Taylor expansion are all zero, we have that $T_n(f, x, dx) < 0$, for all dx.
(3) From the Taylor Young theorem, we know that if the n'th order Taylor expansion is nonzero for dx sufficiently small, then for all dx in some neighborhood U of 0, the absolute value of the expansion dominates the absolute value of the remainder term.
(4) Conclude that for all $dx \in U$, $dx \neq 0$, $f(x + dx) < f(x)$.

(b) If $f^{(k)}(x_0)$ is zero for $k = 1...n-1$ and $f^{(n)}(x_0)$ is non-zero, show that there exists an ϵ-neighborhood around x_0 where the absolute value of the n^{th}-order Taylor expansion is bigger than the absolute value of the remainder term $R_n(x)$.

Ans: The n'th term in the expansion is $\frac{f^{(n)}(x_0)(x-x_0)^n}{n!}$. For some $\eta \in [x_0, x]$, the remainder term can be written as $\frac{f^{(n+1)}(\eta)(x-x_0)^{n+1}}{(n+1)!} = \frac{f^{(n+1)}(\eta)(x-x_0)^n(x-x_0)}{(n+1)n!}$. We need to show that for x sufficiently close to x_0. $\left|\frac{f^{(n)}(x_0)(x-x_0)^n}{n!}\right| > \left|\frac{f^{(n+1)}(\eta)(x-x_0)^n(x-x_0)}{(n+1)n!}\right|$. If $f^{(n+1)}(\eta) = 0$, the inequality holds trivially. Assume therefore that $f^{(n+1)}(\eta) \neq 0$. Extracting the common term, the required inequality will hold if $\left|f^{(n)}(x_0)\right| > \left|\frac{f^{(n+1)}(\eta)(x-x_0)}{(n+1)}\right|$. Since $f_{(n+1)}(\cdot)$ is continuous, and $[x, x_0]$ is a compact set, $|f_{(n+1)}(\cdot)|$ attains a maximum on this set. (This is a consequence of Weierstrass theorem.) Let $\bar{f}_{n+1}$ denote this maximum and pick $\epsilon < \left|\frac{(n+1)f^{(n)}(x_0)}{\bar{f}_{n+1}}\right|$. For x such that $x - x_0 < \epsilon$,

$$\left|\frac{f^{(n+1)}(\eta)(x-x_0)}{(n+1)}\right| < \frac{\epsilon \bar{f}_{n+1}}{(n+1)} < \left|f^{(n)}(x_0)\right| \qquad \text{as required}$$

(c) Give a counter example to show that the result in part (a) would be false if the words "for some even n" were replaced with "for some $n > 0$".

Ans: Take $f(x) = -x^3$ and note that $f_{(1)}(0) = f_{(2)} = 0$ while $f_{(3)}(0) = -6$; This function exhibits the property that for so for the (odd) number $n = 3$, the derivatives $f^{(k)}(x_0)$ are zero for $k = 1...n-1$, and $f^{(n)}(x_0)$ is negative In this case, all of the conditions of the theorem are satisfied except that n is odd. And x^3 isn't maximized at zero.

(d) Explain carefully, but in as few a words as possible, why the argument in (a) works for even n but not for odd n.

Ans: If n is odd, then we cannot conclude, as we did in step (1) of our answer to (a) above, that $(x - x_0)^{(n)} > 0$ for $x \neq x_0$. Therefore we cannot conclude that the sign of the n'th term in the series is determined by the sign of $f_{(n)}$.

(e) Show that the n^{th}-order Taylor expansion around any point x_0 of a polynomial of degree n (i.e. a function of the form $f(x) = \sum_{k=0}^{n} \alpha_k x^k$) is perfectly accurate, regardless of the magnitude of dx.

Ans: Recall that that for $g(t) = \alpha t^k$

If $n \leq k$: $g_{(n)}(t) = t * (t-1) * ... * (t-n) * \alpha t^{k-n}$

If $n > k$: $g_{(n)}(t) = 0$

Hence the $(n+1)^{th}$ derivative of a function $f(x) = \sum_{k=0}^{n} \alpha_k x^k$ is equal to zero and the remainder term $R_n(x) = f_{(n+1)}(\eta)\frac{(x-x_0)^{n+1}}{(n+1)!}$ is zero. And $R_n(x) = 0$ implies that the expansion is perfectly accurate.

(f) Show that if f is an arbitrary polynomial of degree 2, i.e., $f(x) = ax^2 + bx + c$, then for any point x_0, if you add to $f(x_0)$ the 2^{nd}-order Taylor expansion around f, the expression you get is precisely the original function f. More precisely, show by writing out the Taylor expansion explicitly, that for arbitrary dx, $f(x_0 + dx) = f(x_0) + T_2(f, x_0, dx)$.

Ans: Let $f(x) = ax^2 + bx + c$, so that the second order Taylor expansion of f around x_0 is

$$\begin{aligned} & f'(x_0)\,\mathsf{dx} + 0.5 f''(x_0)\,\mathsf{dx}^2 \\ = \quad & (2ax_0 + b)\,\mathsf{dx} + 0.5 \times 2a\,\mathsf{dx}^2 \\ = \quad & (2ax_0 + b)\,\mathsf{dx} + a\,\mathsf{dx}^2 \end{aligned}$$

Therefore,

$$\begin{aligned} f(x_0 + \mathsf{dx}) \quad &= \quad ax_0^2 + bx_0 + c + (2ax_0 + b)\,\mathsf{dx} + a\,\mathsf{dx}^2 \\ &= \quad a(x_0^2 + 2x\,\mathsf{dx} + \mathsf{dx}) + b(x_0 + \mathsf{dx}) + c \\ &= \quad a(x_0 + \mathsf{dx})^2 + b(x_0 + \mathsf{dx}) + c \end{aligned}$$

which is precisely the original function!