MATH 20700: Honors Analysis in Rn I

Problem Set 6

Hung Le Tran

05 Nov 2023

Textbook: Pugh's Real Mathematical Analysis *Collaborators: Lucio, Hung Pham, Duc*

Problem 6.1 (5.7 done)

Two norms $|\cdot|_1$ and $|\cdot|_2$ on a vector space are **comparable** if there are positive constants c, C such that for all nonzero vectors in V we have

$$c \leq \frac{|v|_1}{|v|_2} \leq C.$$

(a) Prove that comparability is an equivalence relation on norms

(b) Prove that any two norms on a finite-dimensional vector space are comparable. [Hint: Use Theorem 3]

(c) Consider the norms

$$|f|_{L^1} = \int_0^1 |f(t)|dt, |f|_{C^0} = \max\{|f(t)| : t \in [0,1]\}$$

defined on the infinite-dimensional vector space C^0 of continuous functions $f : [0,1] \to \mathbb{R}$. Show that the norms are not comparable by finding functions $f \in C^0$ whose integral norm is small but whose C^0 norm is 1.

Solution

(a) To prove that it is an equivalence relation, we have to prove that it is reflexive, symmetric and transitive.

1. Reflexive: $|\cdot|_1$ is trivially comparable to itself, with $c = C = 1$.
2. Symmetric: If $|\cdot|_1$ is comparable to $|\cdot|_2$, in the sense that

$$c \leq \frac{|v|_1}{|v|_2} \leq C.$$

then

$$1/C \leq \frac{|v|_2}{|v|_1} \leq 1/c.$$

so $|\cdot|_2$ is comparable to $|\cdot|_1$.

3. Transitive: Suppose $|\cdot|_1$ is comparable to $|\cdot|_2$, $|\cdot|_2$ is comparable to $|\cdot|_3$, i.e.

$$\begin{aligned} c_1 \leq & \frac{|v|_1}{|v|_2} \leq C_1 \\ c_2 \leq & \frac{|v|_2}{|v|_3} \leq C_2 \\ \Rightarrow c_1 c_2 \leq & \frac{|v|_1}{|v|_3} \leq C_1 C_2 \end{aligned}$$

so $|\cdot|_1$ is comparable to $|\cdot|_3$.

From 3 points above, comparability is an equivalence relation. $\square$

(b) Let $|\cdot|_1$ and $|\cdot|_2$ be norms on V of finite dimension n.

Then there exists an isomorphism $T_1 : \mathbb{R}^n \to (V, |\cdot|_1)$, $T_2 : \mathbb{R}^n \to (V, |\cdot|_2)$. By Theorem 3, $\|T_1\|$, $\|T_1^{-1}\|$, $\|T_2\|$ and $\|T_2^{-1}\|$ are finite and nonzero.

Therefore,

$$\frac{|v|_1}{|v|_2} = \frac{|v|_1}{|v|_{\mathbb{R}^n}} \frac{|v|_{\mathbb{R}^n}}{|v|_2} \leq \|T_1\| \|T_2^{-1}\|; \frac{|v|_1}{|v|_{\mathbb{R}^n}} \frac{|v|_{\mathbb{R}^n}}{|v|_2} \geq \frac{1}{\|T_1^{-1}\| \|T_2\|}$$

so $|\cdot|_1$ and $|\cdot|_2$ are comparable.

(c) Given small $0 < \varepsilon < 1/2$

$$f(x) = \begin{cases} 0 & \text{for } x \in [0, 1/2 - \varepsilon] \\ 1 - 1/2\varepsilon + x/\varepsilon & \text{for } x \in [1/2 - \varepsilon, 1/2] \\ 1 + 1/2\varepsilon - x/\varepsilon & \text{for } x \in [1/2, 1/2 + \varepsilon] \\ 0 & \text{for } x \in [1/2 + \varepsilon, 1] \end{cases}$$

In short, f is 0 from 0 to $1/2 - \varepsilon$, then linearly interpolates 0 to 1 for $x = 1/2 - \varepsilon$ to $x = 1/2$, then linearly interpolates 1 to 0 for $x = 1/2$ to $x = 1/2 + \varepsilon$, then is 0 from $1/2 + \varepsilon$ onwards.

Then $|f|_{L^1} = 1(2\varepsilon)/2 = \varepsilon$, while $|f|_{C^0} = 1$. $\square$

Problem 6.2 (5.8* done)

Let $|\cdot| = |\cdot|_{C^0}$ be the sup norm on C^0 as in the previous exercise. Define an integral transformation $T : C^0 \to C^0$ by

$$T : f \mapsto \int_0^x f(t) dt$$

(a) Show that T is linear, continuous and find its norm.

(b) Let $f_n(t) = \cos(nt), n = 1, 2, \ldots$. What is $T(f_n)$?

(c) Is the set of functions $K = \{f_n : n \in \mathbb{N}\}$ closed? Bounded? Compact?

(d) Is $T(K)$ compact? How about its closure?

Solution

Reiterate that the function space is $C^0 = C^0([0,1], \mathbb{R})$.

(a)

1. WTS T is linear:

Let $f_1, f_2 \in C_0; c \in \mathbb{R}$. Then

$$T(f_1 + cf_2) = \int_0^x ((f_1 + cf_2)(t))dt = \int_0^x (f_1(t) + cf_2(t))dt = T(f_1) + T(cf_2)$$

so T is linear.

2. WTS T is continuous:

Let $f_1, f_2 \in C_0$. Then

$$\begin{aligned}|T(f_1) - T(f_2)| &= \left|\int_0^x (f_1 - f_2)(t)dt\right| \\ &\leq |x| d_{sup}(f_1, f_2) \\ &\leq 1 d_{sup}(f_1, f_2)\end{aligned}$$

so T is 1-Lipschitz. It is therefore continuous.

3. Find $\|T\|$: For $f \in C^0$ such that $|f| = 1$, then

$$\begin{aligned}|Tf| &= \max_{x\in[0,1]} \left\{\left|\int_0^x f(t)dt\right|\right\} \\ &\leq \max_{x\in[0,1]} \{|x||f|\} \\ &= |f| = 1\end{aligned}$$

And the maximum is achieved with $f \equiv 1, f \in C^0$. It follows that

$$\|T\| = \sup_{|f|=1} |Tf| = 1$$

(b)

$$\begin{aligned}T(f_n) &= \int_0^x \cos(nt)dt \\ &= \left[\frac{\sin(nt)}{n}\right]_0^x = \sin(nx)/n\end{aligned}$$

(c)

1. WTS $K = \{f_n : n \in \mathbb{N}\}$ is closed.

Suppose there exists $(f_j = \cos(n_j t))_{j\in\mathbb{N}} \subseteq K$ such that $f_j \xrightarrow{j\to\infty} f \in C^0$.

Case 1: There exists $N \in \mathbb{N}$ such that $\forall\, j \geq N, n_j = n_N$.

Then $f_j = \cos(n_j t) \xrightarrow{j\to\infty} \cos(n_N t)$ trivially, and $\cos(n_N t) \in K \Rightarrow K$ is closed.

Case 2: There doesn't exist $N \in \mathbb{N}$ such that $\forall\, j \geq N, n_j = n_N$. This means that one can construct a subsequence j_k such that

$$n_{j_1} < n_{j_2} < \cdots < n_{j_k} < n_{j_{k+1}} < \cdots$$

in the following way:

Pick $j_1 = 1$. Pick j_2 as the next $j > j_1$ such that $n_j > n_{j_1}$. Given j_k, pick j_{k+1} as the next $j > j_k$ such that $n_j > n_{j_k}$.

If this process can't be continued at some L, that means for all $j > j_L$, $n_j \leq n_{j_L}$.

Then let $D = \min_{i_1,i_2\leq L; i_1\neq i_2} d_{sup}(f_{n_{j_{i_1}}}, f_{n_{j_{i_2}}})$ then $D > 0$ (D is the minimum of the pairwise sup distance of all $f_{n_{j_i}}$ of $j_i \leq j_L$). d_{sup} of $f_{n_{j_{i_1}}}$ and $f_{n_{j_{i_2}}}$ is clearly positive, since $n_{j_{i_1}} \neq n_{j_{i_2}}$.

But this means that $\{f_j\}$ is not Cauchy! Since for all $J \in \mathbb{N}$, there would always exist $j, j' \geq \max\{J, j_L\}$ such that $n_j \neq n_{j'}; n_j, n_{j'} \leq n_{j_L}$ (if there doesn't exist such j and j' then the first condition of the case (Case 2) would be false).

Then,

$$d_{sup}(f_m, f_n) \geq D \not< \varepsilon.$$

Therefore, we can indeed construct such a subsequence j_k satisfying

$$n_{j_1} < n_{j_2} < \cdots < n_{j_k} < n_{j_{k+1}} < \cdots$$

Denote $g_k = f_{n_{j_k}}$, then $(g_k)_{k\in\mathbb{N}}$ is a subsequence of (f_j), so it converges to the same limit. $g_k \xrightarrow{k\to\infty} f$ (in C^0).

We now seek to point out a contradiction for f.

Firstly, $g_k(0) = \cos(0) = 1 \Rightarrow f(0) = 1$.

Secondly, define $x_k = \frac{\pi}{2n_{j_k}}$ then

$$g_k(x_k) = \cos(n_{j_k}(x_k)) = \cos(\pi/2) = 0$$

Also note that $x_k \xrightarrow{k\to\infty} 0$, since $n_{j_k} \xrightarrow{k\to\infty} \infty$.

Now fix $\varepsilon > 0$.

Since $f \in C^0$, there exists $\delta > 0$ such that $d(0, x') \Rightarrow |f(0) - f(x')| < \varepsilon/2$. Since $x_k \xrightarrow{k\to\infty} 0$, there exists L_1 such that $k \geq L_1 \Rightarrow d(0, x_k) < \delta$.

Since $g_k \xrightarrow{k\to\infty} f$ in C^0, there exists L_2 such that $k \geq L_2 \Rightarrow d_{sup}(f, g_k) < \varepsilon/2$.

Take $L = \max\{L_1, L_2\}$ then

$$|f(0) - g_L(x_L)| \le |f(0) - f(x_L)| + |f(x_L) - g_L(x_L)| < 2\varepsilon$$

but $g_L(x_L) = 0 \Rightarrow f(0) = 0$. This contradicts with $f(0) = 1$ above! $\Rightarrow\Leftarrow$

From the 2 cases, we see that only case 1 is valid. K is therefore closed.

2. WTS $K = \{f_n : n \in \mathbb{N}\}$ is bounded.

For any $n \in \mathbb{N}$,

$$|f_n|_{C^0} = |\cos(nt)|_{C^0} = 1$$

so K is bounded.

3. WTS K is not equicontinuous. Suppose that it is.

Then let $\varepsilon = 1/2$, there exists $\delta > 0$ such that

$$|s - t| < \delta \Rightarrow |f_n(s) - f_n(t)| < \varepsilon$$

for all $n \in \mathbb{N}$.

There exists N sufficiently large such that $\frac{\pi}{2N} < \delta$.

Then if $t = 0, s = \frac{\pi}{2N}$ then $|t - s| < \delta$ but

$$|f_N(0) - f_N\left(\frac{\pi}{2N}\right)| = |1 - 0| = 1 \not< 1/2$$

It follows that K is not equicontinuous. From Theorem 18, Chapter 4, K is therefore not compact.

(d)

$$T(K) = \{g_n = T(f_n) = \sin(nx)/n : n \in \mathbb{N}\}$$

1. WTS $T(K)$ is not compact. Suppose that it is.

Then select sequence $(g_n) \subseteq T(K)$ itself.

Let $g_0 \equiv 0$ be the zero function. $g_0 \in C^0$. Then

$$d_{sup}(g_0, g_n) \le \sup\{\sin(nx)/n\} \le 1/n \xrightarrow{n\to\infty} 0$$

so $g_n \xrightarrow{n\to\infty} g_0$. Any of its subsequence would therefore also converge to g_0.

But $g_0 \notin T(K) \Rightarrow T(K)$ is not compact.

2. WTS $E = \overline{T(K)}$ is compact, by showing that it is closed, bounded and equicontinuous.

2.1. E is closed in C^0 because $E = \overline{T(K)}$.

2.2. WTS E is bounded. First, notice that any $g_n \in T(K)$ is bounded:

$$|g_n| \le |\sin(nx)| = 1$$

Let $h \in E, h = \lim_{n\to\infty} h_n$ for some $(h_n) \subseteq T(K)$. Then choose $\varepsilon = 1$, then there exists $N \in \mathbb{N}$ such that $d_{sup}(h, h_N) < 1 \Rightarrow |h| < 1 + |h_N| \leq 2$. Therefore E is bounded.

2.3. WTS E is equicontinuous, i.e., for all $\varepsilon > 0$, there exists $\delta > 0$ such that

$$|s - t| < \delta \Rightarrow |h(s) - h(t)| < \varepsilon$$

for all $h \in E$.

Notice that each $g_n \in T(K)$ is 1-Lipschitz, because

$$|g_n'| = |\cos(nx)| \leq 1$$

It follows that $T(K)$ is equicontinuous, because for all $\varepsilon > 0$, one can choose $\delta = \varepsilon$, then

$$|s - t| < \delta \Rightarrow |g_n(s) - g_n(t)| < 1\delta = \varepsilon$$

Fix $\varepsilon > 0$ and $h = \lim_{n\to\infty} h_n$, for some $(h_n) \subseteq E$.

Since $T(K)$ is equicontinuous and $(h_n) \subseteq T(K)$, there exists $\delta_1 > 0$ such that

$$|s - t| < \delta \Rightarrow |h_n(s) - h_n(t)| < \varepsilon/3 \; \forall \, n \in \mathbb{N}$$

Since $h_n \xrightarrow{n\to\infty} h$, there exists $N \in \mathbb{N}$ such that

$$d_{sup}(h_n, h) < \varepsilon/3$$

Then, choose $\delta = \min\{\varepsilon/3, \delta_1\}$, then we have if $|s - t| < \delta$ then

$$|h(s) - h(t)| \leq |h(s) - h_N(s)| + |h_N(s) - h_N(t)| + |h_N(t) - h(t)| < 3\varepsilon/3 = \varepsilon$$

It follows that E is equicontinuous.

From **2.1, 2.2, 2.3**, it follows that E is compact (Theorem 18, Section 4). □

Problem 6.3 (5.15 done)

Show that both partial derivatives of the function

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{if } (x, y) = (0, 0) \end{cases}$$

exist at the origin but the function is not differentiable there.

Solution

Let $p = (0, 0)$.

$$\begin{aligned}\frac{\partial f}{\partial x}(0) &= \lim_{t\to 0} \frac{f(p+(t,0)) - f(p)}{t} \\ &= \lim_{t\to 0} \frac{\frac{0t}{t^2+0} - 0}{t} \\ &= \lim_{t\to 0} 0 = 0\end{aligned}$$

Similarly,

$$\begin{aligned}\frac{\partial f}{\partial y}(0) &= \lim_{t\to 0} \frac{f(p+(0,t)) - f(p)}{t} \\ &= \lim_{t\to 0} \frac{\frac{0t}{t^2+0} - 0}{t} \\ &= \lim_{t\to 0} 0 = 0\end{aligned}$$

So both partial derivatives exist at (0, 0).

Now WTS f is not differentiable. Suppose it is. Then

$$Jf_{(0,0)} = \begin{pmatrix} \frac{\partial f}{\partial x}(0,0) & \frac{\partial f}{\partial y}(0,0) \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix}$$

Then for $v \in \mathbb{R}^2$,

$$f(v) = f(0) + Df_{(0,0)}(v) + R(v) = R(v)$$

which implies

$$\lim_{v\to 0} \frac{|f(v)|}{|v|} = 0$$

which means for all $\varepsilon > 0$, there exists $\delta > 0$ such that $|v| < \delta \Rightarrow \frac{|f(v)|}{|v|} < \varepsilon$.

However, one can let $r = \min\{\delta/2, 1/(2\varepsilon)\}$ then there exists $v = (r/\sqrt{2}, r/\sqrt{2})$ that has $|v| = r < \delta$ while

$$\frac{|f(v)|}{|v|} = \frac{1}{2r} \geq \varepsilon, \Rightarrow\Leftarrow$$

It follows that f is not differentiable at (0, 0). □

Problem 6.4 (5.16a-c done)

Let $f : \mathbb{R}^2 \to \mathbb{R}^3$ and $g : \mathbb{R}^3 \to \mathbb{R}$ be defined by $f = (x, y, z)$ and $g = w$ where

$$\begin{cases} w &= w(x,y,z) = xy + yz + zx \\ x &= x(s,t) = st \\ y &= y(s,t) = s\cos t \\ z &= z(s,t) = s\sin t \end{cases}$$

(a) Find the matrices that represent the linear transformations $(Df)_p$ and $(Dg)_q$ where $p = (s_0, t_0) = (1, 0)$ and $q = f(p)$.

(b) Use the Chain Rule to calculate the 1×2 matrix $[\partial w/\partial s, \partial w/\partial t]$ that represents $(D(g \circ f))_p$.

(c) Plug the functions $x = x(s,t), y = y(s,t)$ and $z = z(s,t)$ directly into $w = w(x,y,z)$, and recalculate $[\partial w/\partial s, \partial w/\partial t]$, verifying the answer given in (b).

Solution

(a)

$$\begin{aligned} Jf_p &= \begin{pmatrix} -Jx_p- \\ -Jy_p- \\ -Jz_p- \end{pmatrix} \\ &= \begin{pmatrix} t_0 & s_0 \\ \cos t_0 & -s\sin t_0 \\ \sin t_0 & s\cos t_0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

Now when $p = (s_0, t_0) = (1, 0) \Rightarrow x_0 = 0, y_0 = 1, z_0 = 0$.

$$\begin{aligned} Jg_q &= \begin{pmatrix} \frac{\partial g}{\partial x}(q) & \frac{\partial g}{\partial y}(q) & \frac{\partial g}{\partial z}(q) \end{pmatrix} \\ &= \begin{pmatrix} y_0 + z_0 & z_0 + x_0 & x_0 + y_0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \end{pmatrix} \end{aligned}$$

(b)

$$\begin{aligned} D(g \circ f)_p &= (Dg)_q (Df)_p \\ &= \begin{pmatrix} 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 2 \end{pmatrix} \end{aligned}$$

(c)

$$\begin{aligned} w &= xy + yz + zx \\ &= st(s\cos t) + (s\cos t)(s\sin t) + (s\sin t)(st) \\ &= s^2 t(\cos t + \sin t) + s^2 \sin(2t)/2 \end{aligned}$$

then

$$\frac{\partial w}{\partial s}(p) = 2s_0 t_0(\cos t_0 + \sin t_0) + s_0 \sin(2t) = 0$$

and

$$\frac{\partial w}{\partial t}(p) = s_0^2[t_0(\cos t_0 - \sin t_0) + (\cos t_0 + \sin t_0)] + s_0 \cos(2t_0) = 2$$

which verifies

$$Dw_p = D(g \circ f)_p = \begin{pmatrix} 0 & 2 \end{pmatrix}$$

as required. □

Problem 6.5 (5.17a done)

Let $f : U \to \mathbb{R}^m$ be differentiable, $[p, q] \subset U \subset \mathbb{R}^n$, and ask whether the direct generalization of the one-dimensional Mean Value Theorem is true: Does there exist a point $\theta \in [p, q]$ such that

$$f(q) - f(p) = (Df)_\theta(q - p)?$$

Take $n = 1, m = 2$, and examine the function

$$f(t) = (\cos t, \sin t)$$

for $\pi \leq t \leq 2\pi$. Take $p = \pi$ and $q = 2\pi$. Show that there is no $\theta \in [p, q]$ which satisfies the equation above.

Solution

$$f(t) = (\cos t, \sin t)$$

$f : \mathbb{R} \to \mathbb{R}^2$ is differentiable as each of its coordinates is differentiable (Theorem 10):

$$Jf_\theta = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$$

Take $p = \pi, q = 2\pi$ then

$$f(q) - f(p) = (1, 0) - (-1, 0) = (2, 0); q - p = \pi$$

Suppose there exists $\theta \in [p, q]$ such that

$$f(q) - f(p) = (Df)_\theta(q - p)$$

Then

$$\begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix} \pi = \begin{pmatrix} -\pi\sin\theta \\ \pi\cos\theta \end{pmatrix}$$

which implies $\cos\theta = 0, \sin\theta = -2/\pi$. But $\cos\theta = 0$ implies $\sin\theta = \pm 1 \neq -2/\pi$. $\Rightarrow\Leftarrow$

Therefore there doesn't exist such a θ. □

Problem 6.6 (5.18 done)

The **directional derivative** of $f : U \to \mathbb{R}^m$ at $p \in U$ in the direction u is the limit, if it exists,

$$\nabla_p f(u) = \lim_{t \to 0} \frac{f(p + tu) - f(p)}{t}.$$

(Often one requires that $|u| = 1$.)

(a) If f is differentiable at p, why is it obvious that the directional derivative exists in each direction u?

(b) Show that the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x,y) = \begin{cases} \frac{x^3y}{x^4+y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$$

has $\nabla_{(0,0)} f(u) = 0$ for all u but is not differentiable at (0, 0).

Solution

(a) If f is differentiable at p, from Theorem 5,

$$\nabla_p f(u) = \lim_{t\to 0} \frac{f(p+tu) - f(p)}{t} = Df_p(u)$$

exists.

(b) For $u = 0 \in \mathbb{R}^2$,

$$\nabla_{(0,0)} f(u) = \lim_{t\to 0} \frac{f(tu) - f((0,0))}{t} = \lim_{t\to 0} 0/t = 0$$

For $u = (v,w) \in \mathbb{R}^2 \backslash \{0\}$,

$$\begin{aligned} \nabla_{(0,0)} f(u) &= \lim_{t\to 0} \frac{f(tu) - f((0,0))}{t} \\ &= \lim_{t\to 0} \frac{\frac{t^3v^3tw}{t^4v^4+t^2w^2}}{t} \\ &= \lim_{t\to 0} \frac{tv^3w}{t^2v^4+w^2} = \frac{0}{w^2} = 0 \end{aligned}$$

Therefore

$$\nabla_{(0,0)} f(u) = 0 \; \forall \; u \in \mathbb{R}^2$$

WTS f is not differentiable at $p = (0,0)$. Suppose that it is, then

$$Df_p(u) = \nabla_p(u) = 0 \; \forall \; u \in \mathbb{R}^2$$

Let $v \in \mathbb{R}^2$ then

$$f(v) = f((0,0)) + Df_p(v) + R(v) \Rightarrow f(v) = R(v)$$

Therefore

$$\lim_{v\to 0} \frac{|f(v)|}{|v|} = 0$$

i.e., for all $\varepsilon > 0$, there exists $\delta > 0$ such that $|v| < \delta \Rightarrow \frac{|f(v)|}{|v|} < \varepsilon$

But for all $\delta > 0$, let $r = \min\{1, \delta/2\}$ then there exists $v = (r, r^2)$.

v has norm: $|v| = \sqrt{r^4 + r^2} \leq \sqrt{r^2 + r^2} = r\sqrt{2} < \delta$. $(r \leq 1 \Rightarrow r^4 \leq r^2)$

However,

$$\frac{|f(v)|}{|v|} = \frac{\frac{r^5}{r^4+r^4}}{|v|} \geq \frac{r}{2|v|} \geq \frac{r}{2\sqrt{2}r} = \frac{1}{2\sqrt{2}} \not< \varepsilon$$

for sufficiently small ε, $\Rightarrow\Leftarrow$.

It follows that f is not differentiable at $(0,0)$. □

Problem 6.7 (5.20 done)

Assume that U is a connected open subset of $\mathbb{R}^n$ and $f : U \to \mathbb{R}^m$ is differentiable everywhere on U. If $(Df)_p = 0$ for all $p \in U$, show that f is constant.

Solution

Fix $p \in U$. Let $A = \{u \in U : f(u) = f(p)\}$. We WTS A is clopen in U.

1. WTS A is open. Let $u \in A$. Then $u \in U$, U is open so there exists $B_{\mathbb{R}^n}(u,r) \subseteq U$. WTS $B_{\mathbb{R}^n}(u,r) \subseteq A$, which would imply $B_U(u,r) \subseteq A$.

Indeed, pick any $v \in B_{\mathbb{R}^n}(u,r)$. Then $[u,v] \subset B_{\mathbb{R}^n}(u,r) \subset U$.

Since $(Df_p) = 0 \; \forall \, p \in U$, f is C^1. One can therefore apply the C^1 MVT:

$$f(v) - f(u) = \int_0^1 (Df_{u+t(v-u)})(v-u)dt = 0 \Rightarrow f(v) = f(u) = f(p) \Rightarrow v \in A$$

Therefore $B_U(u,r) \subseteq A \Rightarrow A$ is open.

2. WTS A is closed in U. Take sequence $(u_n) \subseteq A$ such that $u_n \xrightarrow{n\to\infty} u \in U$. Since f is continuous, $f(u_n) \xrightarrow{n\to\infty} f(u)$. But $f(u_n) = 0 \; \forall \, n \in \mathbb{N} \Rightarrow f(u) = 0 \Rightarrow u \in A$. Therefore A is closed in U.

From 1 and 2, it follows that A is clopen in U. It is non empty, since $p \in A$, so $A = U$. It follows that f is constant in U. □

Problem 6.8 (III done)

For each of the examples below, draw an approximate picture of $f(B(p,\varepsilon))$ for the indicated function f and point p, for ε very small.

(a) $f(t) = (t, t^2)$ at $p = 1$.

(b) $f(x,y) = (e^{2x^2+2x+y}, \sin(3x) - \cos(x+y)); p = (0,0)$.

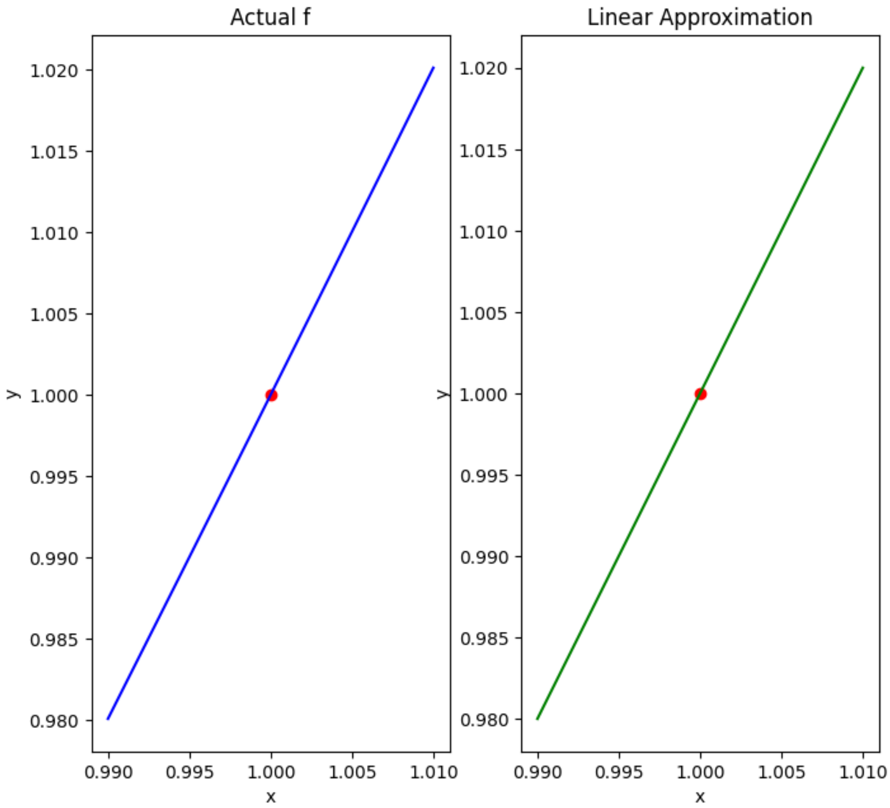
(c) $f(x,y) = (x^2+y, 2xy, 3\log(y) - x); p = (1,1)$.

(d) $f(x,y,z) = (x+y+\sin(xy), y+z, x+y+z); p = (0,0,0)$.

Solution

(a)

$$Jf_1 = \begin{pmatrix} 1 \\ 2t \end{pmatrix}_{t=0} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$



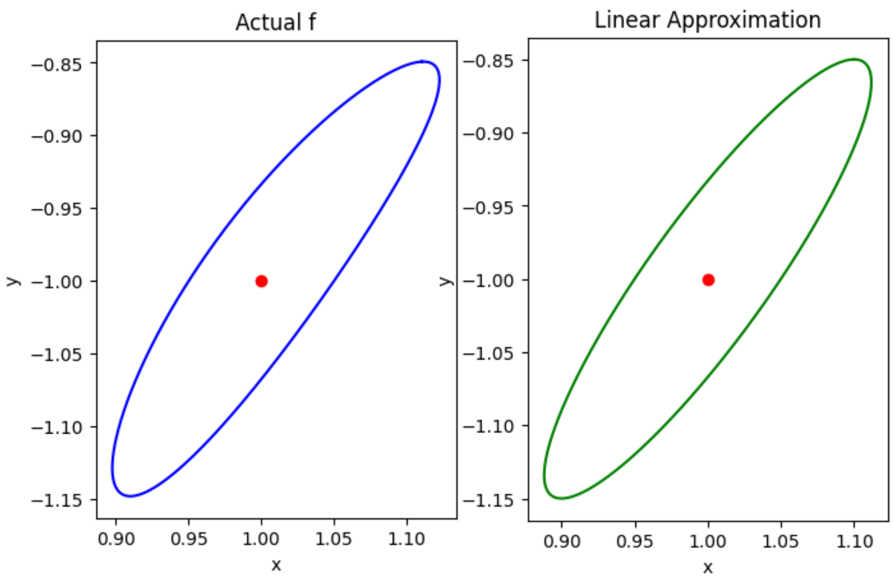
which stretches twice in y as it does in x.

(b)

$$Jf_{(0,0)} = \begin{pmatrix} e^{2x^2+2x+y}(4x+2) & e^{2x^2+2x+y} \\ 3\cos(3x)+\sin(x+y) & \sin(x+y) \end{pmatrix}_{(x,y)=(0,0)} = \begin{pmatrix} 2 & 1 \\ 3 & 0 \end{pmatrix}$$

Its singular values are $\sqrt{2\sqrt{10}+7} \approx 3.65$, $\sqrt{7-2\sqrt{10}} \approx 0.82$.

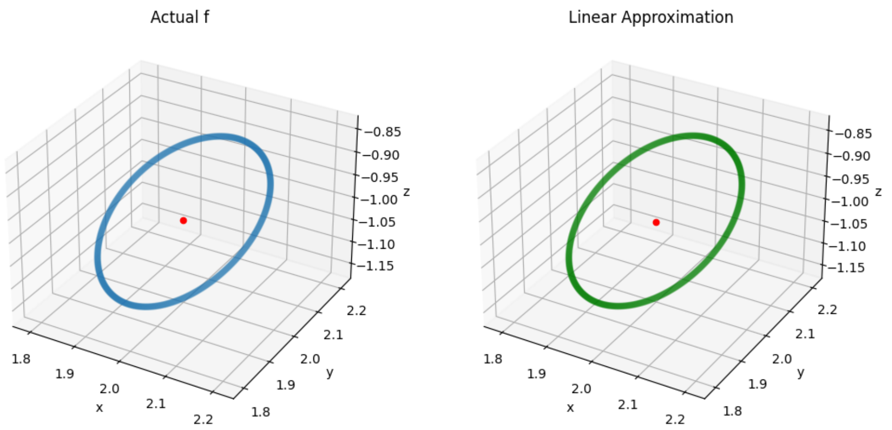
And $3.65/0.82 \approx 4.44$



(c)

$$Jf_{(1,1)} = \begin{pmatrix} 2x & 1 \\ 2y & 2x \\ -1 & 3/y \end{pmatrix}_{(x,y)=(1,1)} = \begin{pmatrix} 2 & 1 \\ 2 & 2 \\ -1 & 3 \end{pmatrix}$$

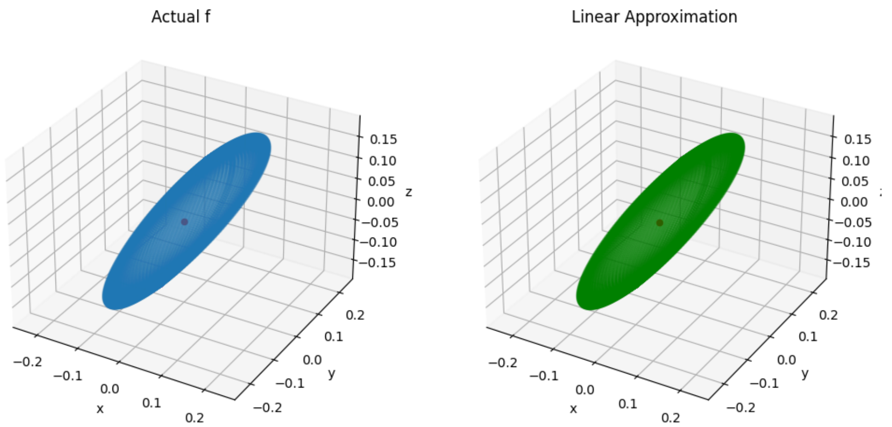
Its singular values are $\frac{\sqrt{2}\sqrt{\sqrt{61}+23}}{2} \approx 3.92$, $\frac{\sqrt{2}\sqrt{23-\sqrt{61}}}{2} \approx 2.75$. $3.92/2.75 \approx 1.42$



(d)

$$Jf_{(0,0,0)} = \begin{pmatrix} 1+y\cos(xy) & 1+x\cos(xy) & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}_{(x,y,z)=(0,0,0)} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Its singular values are $\sqrt{2\sqrt{2}+3} \approx 2.41, 1, \sqrt{3-2\sqrt{2}} \approx 0.41$.



$\square$

Problem 6.9 (IV done)

Let $U \subset \mathbb{R}^n$, and let $f : U \to \mathbb{R}$ be a function. For $v, w \in \mathbb{R}^n$ (viewed as $n \times 1$ matrices), denote by $\langle v, w \rangle = v \cdot w = v^t w$ the standard Euclidean dot product. Recall that $|v| = \langle v, v \rangle^{1/2}$.

(a) Suppose that f is differentiable at $p \in U$, with derivative $Df_p \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$. Show that there is a unique vector $w \in \mathbb{R}^n$ such that, for all $v \in \mathbb{R}^n$:

$$\langle v, w \rangle = Df_p(v).$$

The vector w is called the gradient vector of f at p and is denoted $grad_p(f)$.

(b) Show that, with respect to the standard basis $e_1, \ldots, e_n$ of $\mathbb{R}^n$, we have

$$grad_p(f) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \ldots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$

(c) Show that $\|Df_p\| = |grad_p(f)|$, and that the maximum value of the function $g : S^{n-1} \to \mathbb{R}$ on the $(n-1)-$sphere defined by

$$g(v) = Df_p(v)$$

is attained at $v = grad_p(f)/|grad_p(f)|$. Thus the direction of the gradient vector field $f(p) = grad_p(f)$ is the direction of steepest ascent (increase) for the function f, and the magnitude $|f(p)|$ is the greatest (to first order) rate of increase of the function f at p.

(d) Show that if $\gamma : (-\varepsilon, \varepsilon) \to U$ is a differentiable curve through p along which f is constant (i.e. if $\gamma(0) = p$ and $f(\gamma(t)) = f(\gamma(0))$ for all $t \in (-\varepsilon, \varepsilon)$), then $\gamma'(0)$ is perpendicular to $grad_p(f)$; i.e.

$$\langle grad_p(f), \gamma'(0) \rangle = 0.$$

In words, this means that the gradient of f is everywhere perpendicular to the level sets $f^{-1}(c), c \in \mathbb{R}$ of f.

(e) For the function $f(x, y) = \sin(2x+y)+e^{\cos(xy)}$, plot, using a computer program, the graph of the function f (in 3D), and the level sets of f in the region $[-2, 2]\times[-2, 2]$. Then plot the gradient vector field for f. Wolfram alpha works, for example. Think about steepest ascent in terms of climbing the graph and level sets in terms of walking at a fixed elevation on the graph. The graph of the level sets is like a map of the mountain, and the gradient vector field points at the shortest, most exhausting, path up. You don't have to write down your thoughts, but do supply pics of the graphs.

(f) *Explain how to generalize this notion of gradient to any inner product on $\mathbb{R}^m$. In this way, one can use a different measure of increase of the function (if you're a budding economist or physicist, this could be useful...). The notion of gradient generalizes even more: you can let the inner product $\langle \cdot, \cdot \rangle_p$ depend on the point p! (This is the beginning of Riemannian geometry).

Solution

(a) f is differentiable at $p \in U$, let J be the Jacobian matrix of Df_p, using the standard basis of $\mathbb{R}^n$ and $\mathbb{R}$. J is a $1 \times n$ matrix, $J \in \mathcal{M}_{1\times n}$.

We want to show that $w = J^t \in \mathcal{M}_{n\times 1}$, or alternatively $\in (\mathbb{R}^n)$ satisfies $\langle v, w\rangle = Df_p(v)$. Indeed,

$$\langle v, w\rangle = w^t v = Jv = Df_p(v)\ \forall\ v \in U$$

and we're done.

To prove uniqueness, suppose there exists w_1, w_2 that satisfy

$$\langle v, w_1\rangle = \langle v, w_2\rangle = Df_p(v)\ \forall\ v \in U$$

then

$$T_1(v) = T_2(v) = Df_p(v)\ \forall\ v \in U$$

where T_1, T_2 are linear transformations $U \to \mathbb{R}$ that has w_1^t and w_2^t respectively as their Jacobian matrices. Therefore T_1 and T_2 are linear transformations that satisfy the properties of the total derivative of f. But Df_p is the unique linear transformation that satisfies those properties, so $T_1 \equiv T_2 \equiv Df_p \Rightarrow w_1 = w_2$. $\square$

(b) Since

$$J = \begin{pmatrix} \frac{\partial p}{\partial x_1} & \frac{\partial p}{\partial x_2} & \cdots & \frac{\partial p}{\partial x_n} \end{pmatrix}$$

and we've shown above that

$$grad_p(f) = J^t = \begin{pmatrix} \frac{\partial p}{\partial x_1} \\ \frac{\partial p}{\partial x_2} \\ \cdots \\ \frac{\partial p}{\partial x_3} \end{pmatrix} \quad \square$$

(c) Let $v \in \mathbb{R}^n$. Then using Cauchy-Schwarz:

$$|Df_p(v)| = |\langle J^t, v\rangle| \leq |J^t||v| = |grad_p(f)||v| \Rightarrow \frac{|Df_p(v)|}{|v|} \leq |grad_p(f)|$$

with equality achievable when v and $grad_p(f)$ are linearly dependent, meaning

$$v = \lambda grad_p(f)$$

for some $\lambda \in \mathbb{R}$.

Therefore

$$\|Df_p\| = \sup_{v\in\mathbb{R}^n, v\neq 0} \left\{ \frac{|Df_p(v)|}{|v|} \right\} = |grad_p(f)|$$

since the max is achieved in $\mathbb{R}^n$.

Furthermore, with g defined for $v \in S^{n-1}$:

$$g(v) = Df_p(v)$$

The maximum, as aforementioned, is achieved when

$$v = \lambda grad_p(f)$$

Find λ:

$$|v| = \lambda |grad_p(f)| \Rightarrow \lambda = 1/|grad_p(f)|$$

so the max is achieved when

$$v = \lambda grad_p(f) = grad_p(f)/|grad_p(f)|$$

as required. $\square$

(d) What we WTS is equivalent to:

$$\langle grad_p(f), \gamma'(0)\rangle = 0 \Leftrightarrow J(\gamma'(0)) = 0 \Leftrightarrow Df_p(\gamma'(0)) = 0$$

Since f is differentiable at p:

$$f(\gamma(t)) = (f(p)) + Df_p(\gamma(t) - p) + R(\gamma(t) - p)$$

where

$$\lim_{\gamma(t)\to p} \frac{|R(\gamma(t) - p)|}{|\gamma(t) - p|} = 0$$

Then:

$$\frac{1}{t} Df_p(\gamma(t) - p) + \frac{1}{t} R(\gamma(t) - p) = \frac{1}{t}[f(\gamma(t)) - f(p)] = 0 \tag{1}$$

We try to take the limit as $t \to 0$.

1. Since γ is differentiable,

$$\gamma(t) = \gamma(0) + D\gamma_0(t) + R(t)$$

which implies

$$\lim_{t\to 0}\left(\frac{|\gamma(t) - \gamma(0)|}{|t|}\right) = \lim_{t\to 0}\left(\frac{|D\gamma_0(t) + R(t)|}{|t|}\right)$$

which is squeezed:

$$\lim_{t\to 0}\left(\frac{|D\gamma_0(t)| - |R(t)|}{|t|}\right) \le \lim_{t\to 0}\left(\frac{|D\gamma_0(t) + R(t)|}{|t|}\right) \le \lim_{t\to 0}\left(\frac{|D\gamma_0(t)| + |R(t)|}{|t|}\right)$$

therefore

$$|D\gamma_0(1)| \le \lim_{t\to 0}\left(\frac{|D\gamma_0(t)| - |R(t)|}{t}\right) \le |D\gamma_0(1)| \tag{2}$$

so the limit evaluates to $M = |D\gamma_0(1)| \in \mathbb{R}$.

2. Since γ is differentiable, and therefore continuous at 0, $t \to 0 \Rightarrow \gamma(t) \to p$. Therefore

$$\lim_{t\to 0} \frac{|R(\gamma(t) - p)|}{|\gamma(t) - p|} = \lim_{\gamma(t)-p} \frac{|R(\gamma(t) - p)}{|\gamma(t) - p|} = 0 \tag{3}$$

From (2) and (3), the following limit exists:

$$\lim_{t\to 0}\frac{|R(\gamma(t)-p)|}{|t|} = \left(\lim_{t\to 0}\frac{|R(\gamma(t)-p)|}{|\gamma(t)-p|}\right)\left(\lim_{t\to 0}\frac{|\gamma(t)-p|}{|t|}\right) = 0M = 0$$

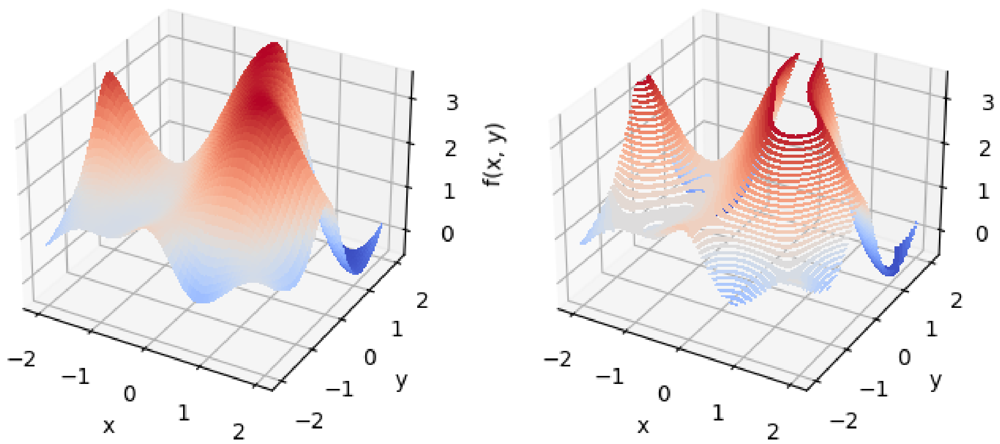
which implies

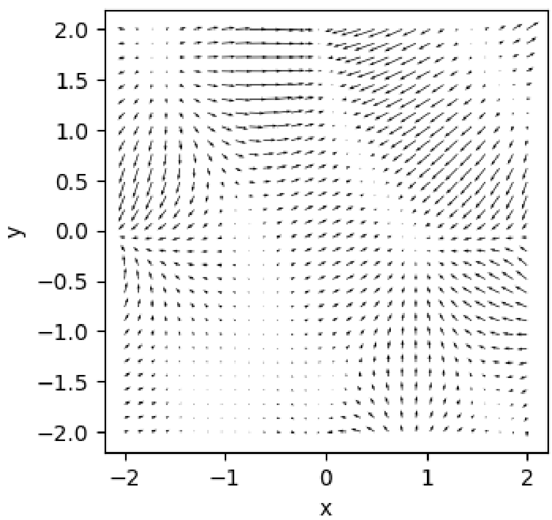
$$\lim_{t\to 0}\frac{R(\gamma(t)-p)}{t} = 0$$

Therefore, taking $t \to 0$ in (1), we have:

$$\begin{aligned}
0 &= \lim_{t\to 0}\frac{Df_p(\gamma(t)-p)}{t} \\
&= Df_p\left(\lim_{t\to 0}\frac{\gamma(t)-\gamma(0)}{t}\right) \quad \text{(since } Df_p, \gamma \text{ continuous)} \\
&= Df_p(\gamma'(0))
\end{aligned}$$

(e)





(f) To generalize this notion of gradient to any inner product on $\mathbb{R}^m$, e.g., $\langle \cdot, \cdot \rangle_p$, to find $grad_p(f)$ at each point p, we first find the level curve of f that goes through $f(p)$, i.e., the curve $\gamma : (-\varepsilon, \varepsilon) \to U$ such that $f(\gamma(t)) = f(\gamma(0)) = f(p)$.

Then $grad_p(f) \in U$ can be defined as the vector that is perpendicular to $\gamma'(0)$, i.e.

$$\langle grad_p(f), \gamma'(0) \rangle = 0$$

$\square$