Coláiste na Tríonóide, Baile Átha Cliath
Trinity College Dublin
Ollscoil Átha Cliath | The University of Dublin

Faculty of Science, Technology, Engineering and Mathematics

School of Mathematics

SF Mathematics **Michaelmas Term 2024**
SF Joint Honours
JS TP

Analysis in Several Real Variables - Sample Exam 2 - Solutions

Day **Place** **Time**

Dr. Adam Keilthy

Instructions to candidates:

Attempt any three questions. If you attempt all four questions, only your best three will be considered in your grade. All questions are worth 30 points

Unless stated otherwise, you may use all statements given lectures without proof, but must clearly justify that the assumptions of statement are fulfilled.

Additional instructions for this examination:

You may use a non-programmable calculator. Please indicate the make and model of your calculator on each answer book used.

You may not start this examination until you are instructed to do so by the Invigilator.

Question 1

Let $M_{2\times 2}(\mathbb{R})$ denote the set of (2×2)-matrices, and let $\mathrm{GL}_2(\mathbb{R})$ denote the subset of invertible matrices.

i) (8pts) Determine an explicit isomorphism of real vector spaces $\Psi : \mathbb{R}^4 \to M_{2\times 2}$.

ii) (8pts) Show that the map $\mathbb{R}^4 \to \mathbb{R}$ given by

$$\vec{v} \mapsto \det \Psi(\vec{v})$$

is continuous on $\mathbb{R}^4$.

iii) (6pts) Hence determine whether $\Psi^{-1}(\mathrm{GL}_2(\mathbb{R}))$ is open or closed or neither as a subset of of $\mathbb{R}^4$.

iv) (8pts) Prove that the map

$$\vec{v} \mapsto \Psi^{-1}\left(\Psi(\vec{v})^{-1}\right)$$

induced by matrix inversion is continuous on $\Psi^{-1}(\mathrm{GL}_2(\mathbb{R}))$.

Solution

i) The map

$$\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \mapsto \begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

is clearly linear and injective, giving an injective linear map between two 4 dimensional vector spaces, hence an isomorphism.

ii) Explicitly, this map is

$$\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \mapsto ad - bc$$

which is continuous as the coordinate maps are continuous, multiplication is continuous, and addition is continuous. Thus, it is a composition of continuous functions.

If we wanted to be overly explicit, the map is

$$\vec{v} \mapsto s(m(\pi_1(\vec{v}), \pi_4(\vec{v})), m(\pi_2(\vec{v}), \pi_3(\vec{v})))$$

where π_i is the i^{th} coordinate function, $m : \mathbb{R}^2 \to \mathbb{R}$ is multiplication and $s : \mathbb{R}^2 \to \mathbb{R}$ is addition.

iii) The general linear group is the set of invertible matrices, i.e. those with non-zero determinant. Denoting by ψ the map from part (ii), we have that

$$\begin{aligned}\Psi^{-1}(\mathrm{GL}_2(\mathbb{R})) &= \{\vec{v} \in \mathbb{R}^4 \mid \det(\Psi(\vec{v}) \neq 0\} \\ &= \{\vec{v} \in \mathbb{R}^4 \mid \psi(\vec{v}) \neq 0\} \qquad = \{(a,b,c,d) \in \mathbb{R}^4 \mid ad - bc \neq 0\}\end{aligned}$$

which is open as ψ is continuous.

iv) Explicitly, this is the map

$$\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \mapsto \frac{1}{ad - bc} \begin{pmatrix} d \\ -b \\ -c \\ a \end{pmatrix}$$

Recall that $r : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ given by $t \mapsto t^{-1}$ is continuous. Since $ad - bc \neq 0$ on $\Psi^{-1}(\mathrm{GL}_2(\mathbb{R}))$, the map

$$\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \mapsto \frac{1}{ad - bc}$$

is continuous on $\Psi^{-1}(\mathrm{GL}_2(\mathbb{R}))$. Thus

$$\vec{v} = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \mapsto \frac{1}{ad - bc} \pi_i(\vec{v})$$

is continuous on $\Psi^{-1}(\mathrm{GL}_2(\mathbb{R}))$ for each $i = 1, 2, 3, 4$. Thus the map

$$\vec{v} \mapsto \Psi^{-1}\left(\Psi(\vec{v})^{-1}\right)$$

is continuous, as each of its components are.

Question 2

1. (8pts) State, with all necessary hypotheses, the chain rule for differentiable functions of several real variables

2. (10pts) Let $g : \mathbb{R} \to \mathbb{R}^m$ and $f : \mathbb{R}^m \to \mathbb{R}$ be functions differentiable everywhere. Using the chain rule, show that

$$(f \circ g)'(t) = \sum_{k=1}^{m} \frac{\partial f}{\partial x_k}(g(t)) \frac{dg_k}{dt}(t)$$

for every $t \in \mathbb{R}$, where

$$g(t) = (g_1(t), \ldots, g_m(t)).$$

3. (12pts) Fix $(x_1, \ldots, x_m) \in \mathbb{R}^m$. Hence, or otherwise, show that there exists $\delta > 0$ and $\theta \in (0, 1)$ such that

$$\begin{aligned} & f(x_1 + a_1 h, x_2 + a_2 h, \ldots, x_m + a_m h) \\ = {} & f(x_1, \ldots, x_m) + h \sum_{k=1}^{m} a_k \frac{\partial f}{\partial x_k}(x_1 + a_1 \theta h, x_2 + a_2 \theta h, \ldots, x_m + a_m \theta h) \end{aligned}$$

for all $h \in (-\delta, \delta)$.

Solution

i) Let $X \subset \mathbb{R}^m$, $Y \subset \mathbb{R}^n$ be open sets, and let $\vec{p} \in X$. Let $\varphi : X \to Y$ be differentiable at $\vec{p}$ and let $\psi : Y \to \mathbb{R}^s$ be differentiable at $\varphi(\vec{p})$. Then the composition

$$(\psi \circ \varphi) : X \to \mathbb{R}^s$$

is differentiable at $\vec{p}$ with derivative

$$(\mathrm{D}(\psi \circ \varphi))_{\vec{p}} = (\mathrm{D}\,\psi)_{\varphi(\vec{p})} \circ (\mathrm{D}\,\phi)_{\vec{p}}.$$

ii) If a function is differentiable everywhere, the derivative is given by the Jacobian everywhere. Hence

$$(\mathrm{D}\,g)_t = \begin{pmatrix} \frac{dg_1}{dt}(t) \\ \vdots \\ \frac{dg_m}{dt}(t) \end{pmatrix}$$

and

$$(\mathrm{D}\,f)_{\vec{x}} = \begin{pmatrix} \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_m} \end{pmatrix}.$$

Thus, by chain rule,

$$\begin{aligned}(f\circ g)'(t) &= (\mathrm{D}(f\circ g))_t \\ &= \begin{pmatrix} \frac{\partial f}{\partial x_1}(g(t)) & \cdots & \frac{\partial f}{\partial x_m}(g(t)) \end{pmatrix} \begin{pmatrix} \frac{dg_1}{dt}(t) \\ \vdots \\ \frac{dg_m}{dt}(t) \end{pmatrix} \\ &= \sum_{k=1}^{m} \frac{\partial f}{\partial x_k}(g(t))\frac{dg_k}{dt}(t).\end{aligned}$$

iii) Taylor's theorem says that for any function

$$F : \mathbb{R} \to \mathbb{R}$$

differentiable at 0, there exists $\delta > 0$ and $\theta \in (0,1)$ such that

$$F(h) = F(0) + F'(\theta h)h$$

for all $h \in (-\delta, \delta)$. Consider the function

$$F(t) = f(x_1 + a_1 t, \ldots, x_m + a_m t).$$

By chain rule, this is differentiable at $t = 0$, with derivative

$$F'(t) = \sum_{k=1}^{m} a_k \frac{\partial f}{\partial x_k} f(x_1 + a_1 t, \ldots, x_m + a_m t).$$

By Taylor's theorem applied to F, we find that there exists $\delta > 0$ and $\theta \in (0,1)$ such that

$$\begin{aligned} f(x_1 + a_1 h, \ldots, x_m + a_m h) &= F(h) \\ &= F(0) + F'(\theta h)h \\ &= f(x_1, \ldots, x_m) \\ &+ h \sum_{k=1}^{m} a_k \frac{\partial f}{\partial x_k}(x_1 + a_1\theta h, x_2 + a_2\theta h, \ldots, x_m + a_m\theta h) \end{aligned}$$

for all $h \in (-\delta, \delta)$.

Question 3

Recall that a real symmetric $(m \times m)$-matrix A is called positive definite if $\langle \vec{x}, A\vec{x} \rangle > 0$ for all non-zero $\vec{x} \in \mathbb{R}^m$. Given a positive definite matrix A, we define the A-norm by

$$\|\vec{x}\|_A = \sqrt{\langle \vec{x}, A\vec{x} \rangle}.$$

1. (5pts) Show that every eigenvalue of A is positive.

2. (8pts) Hence show that the exist constants $c, C > 0$ such that

   $$c\|\vec{x}\| \leq \|\vec{x}\|_A \leq C\|\vec{x}\|$$

   for all $\vec{x} \in \mathbb{R}^m$.

3. (10pts) Show that $\|\vec{x} + \vec{y}\|_A \leq \|\vec{x}\|_A + \|\vec{y}\|_A$ for all $\vec{x}$, $\vec{y} \in \mathbb{R}^m$.

   Hint: Diagonalise A. Given $\vec{x}$, can you find $\vec{v}_x$ such that $\|\vec{x}\|_A = \|\vec{v}_x\|$? Is the mapping $\vec{x} \to \vec{v}_x$ linear?

4. (7pts) Prove that a sequence $\{\vec{x}_n\}$ of points in $\mathbb{R}^m$ converges to a point $\vec{p}$ with respect to the A-norm if and only if it converges with respect to the usual Euclidean norm.

Solution

i) Let $\vec{v}$ be an eigenvector of A with eigenvalue λ. Then, as $\vec{v} \neq 0$

$$0 < \langle \vec{v}, A\vec{v} \rangle = \langle \vec{v}, \lambda\vec{v} \rangle = \lambda\|\vec{v}\|^2.$$

Since $\|\vec{v}\|^2 > 0$, this implies $\lambda > 0$. As $\vec{v}$ was an arbitrary eigenvector, this implies all eigenvalues are positive.

ii) We will give two solutions. The first will use the extreme value theorem, but requires some linear algebra, and the second will use just linear algebra.

For the first argument, first note at the inequality is true for $\vec{x} = 0$. Next we note that $\|\mu\vec{x}\| = |\mu|\|\vec{x}\|$ for every $\mu \in \mathbb{R}$. Hence, it suffices to prove the inequality for vectors of unit norm, as the inequality (for non-zero $\vec{x}$) is equivalent to

$$c \leq \left\| \frac{\vec{x}}{\|\vec{x}\|} \right\|_A \leq C.$$

Next, we note that

$$\vec{x} \mapsto \|\vec{x}\|_A$$

is a continuous map. As the set $\{\vec{x} \in \mathbb{R}^m \mid \|\vec{x}\| = 1\}$ is closed and bounded, $\|\vec{x}\|_A$ achieves a maximum C and a minimum c on this set. In particular, there exists $\vec{x}_{min}$ and $\vec{x}_{max}$ of norm 1 such that

$$c = \sqrt{\langle \vec{x}_{min}, A\vec{x}_{min} \rangle} > 0, \quad C = \sqrt{\langle \vec{x}_{max}, A\vec{x}_{max} \rangle} > 0$$

as A is positive definite and both vectors are non-zero.

For the second argument, we note that as A is real symmetric, there exists an orthonormal basis of eigenvectors $\vec{v}_1, \ldots, \vec{v}_m$. Without loss of generality, assume that the associated eigenvalues satisfy

$$0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_m.$$

Then, for any vector $\vec{x} = a_1\vec{v}_1 + \cdots a_m\vec{v}_m$, we have that

$$\begin{aligned}\|\vec{x}\|_A &= \sqrt{\langle\sum_{i=1}^m a_i\vec{v}_i, A(\sum_{j=1}^m a_j\vec{v}_j)\rangle} \\ &= \sqrt{\sum_{i,j=1}^m a_ia_j\langle\vec{v}_i, A\vec{v}_j\rangle} \\ &= \sqrt{\sum_{i,j=1}^m a_ia_j\langle\vec{v}_i, \lambda_j\vec{v}_j\rangle} \\ &= \sqrt{\sum_i^m \lambda_i a_i^2}\end{aligned}$$

via orthonormality. As

$$\lambda_1 a_i^2 \leq \lambda_i a_i^2 \leq \lambda_n a_i^2$$

for each i, we have that

$$\sqrt{\lambda}_1\|\vec{x}\| = \sqrt{\lambda_1\sum_{i=1}^m a_i^2} \leq \|\vec{x}\|_A \leq \sqrt{\lambda_m\sum_{i=1}^m a_i^2} = \sqrt{\lambda_m}\|\vec{x}\|$$

as needed.

iii) In the orthonormal eigenbasis, if $\vec{x} = \sum_{i=1}^m a_i\vec{v}_i$

$$\|\vec{x}\|_A = \sqrt{\sum_{i=1}^m \lambda_i a_i^2}$$

Since $\lambda_i > 0$ for each i, we can take its square root and define

$$\vec{v}_{\vec{x}} = \begin{pmatrix}\sqrt{\lambda_1}a_1 \\ \sqrt{\lambda_2}a_2 \\ \vdots \\ \sqrt{\lambda_m}a_n\end{pmatrix}$$

It is easy to see that

$$\|\vec{v}_{\vec{x}} = \sqrt{\sum_{i=1}^m \lambda_i a_i^2} = \|\vec{x}\|_A.$$

Furthermore, the map $\vec{x} \mapsto \vec{v}_{\vec{x}}$ is clearly linear, so

$$\vec{v}_{\vec{x}+\vec{y}} = \vec{v}_{\vec{x}} + \vec{v}_{\vec{y}}.$$

Thus, for any $\vec{x}$, $\vec{y} \in \mathbb{R}^m$

$$\begin{aligned}\|\vec{x}+\vec{y}\|_A &= \|\vec{v}_{\vec{x}+\vec{y}}\| \\ &= \|\vec{v}_{\vec{x}} + \vec{v}_{\vec{y}}\| \\ &\leq \|\vec{v}_{\vec{x}}\| + \|\vec{v}_{\vec{y}}\| \\ &= \|\vec{x}\|_A + \|\vec{x}\|_A\end{aligned}$$

using the standard triangle inequality.

iv) Suppose $\{\vec{x}_n\}$ converges to $\vec{p}$ with respect to the usual norm. Then, for all $\varepsilon > 0$, there exists $N > 0$ such that

$$\|\vec{x}_n - \vec{p}\| < \frac{\varepsilon}{C}$$

for all $n \geq N$. Hence, for all $n \geq N$,

$$\|\vec{x}_n - \vec{p}\|_A \leq C\|\vec{x} - \vec{p}\| < \varepsilon$$

and $\{\vec{x}_n\}$ converges to $\vec{p}$ in the A-norm. Similarly, if $\{\vec{x}_n\}$ converges to $\vec{p}$ in the A-norm, then for all $\varepsilon > 0$ there exists $N > 0$ such that

$$\|\vec{x}_n - \vec{p}\|_A < c\varepsilon$$

for all $n \geq N$ and hence

$$\|\vec{x}_n - \vec{p}\| \leq \frac{1}{c}\|\vec{x} - \vec{p}\|_A < \varepsilon$$

for all $n \geq N$.

Question 4

i) (10pts) State, with all necessary hypotheses, the implicit function theorem.

ii) (8pts) Show that

$$x^2 + 7xy + 10y^2 - 1 = 0$$

defines y as an continuously differentiable function of x near $(1, 0)$.

iii) (8pts) Are there any points around which neither x or y can be expressed as a continuous function of the other?

iv) (4pts) Determine the maximal set containing $(1, 0)$ on which y can be expressed as a continuously differentiable function of x.

Solution

i) Let $X \subset \mathbb{R}^m$ be open, $f_1, \ldots, f_k : X \to \mathbb{R}$ be continuously differentiable functions on X, with $k < m$. Define

$$S = \{\vec{x} \in X \mid f_i(\vec{x}) = 0 \text{ for all } 1 \leq i \leq k\}$$

and let $\vec{p} \in S$. Suppose that the matrix

$$J(\vec{x}) = (\partial_j f_i(\vec{x}))$$

has rank at least k at $\vec{p}$. Without loss of generality, we assume that the first k columns are independent, reordering the variables $x_1, \ldots, x_m$ if needed. Then there exists open $V \subset X$ containing $\vec{p}$, and continuously differentiable functions $g_1, \ldots, g_k : U \to \mathbb{R}$ defined on some open set $U \subset \mathbb{R}^{m-k}$ containing $(p_{k+1}, p_{k+2} \ldots, p_m)$ such that

$$S \cap V = \{(g_1(\vec{x}')), \ldots, g_k(\vec{x}'), x_{k+1}, \ldots, x_m) \mid \vec{x}' = (x_{k+1}, \ldots x_m) \in U\}$$

ii) Letting $f(x, y) = x^2 + 7xy + 10y^2 - 1$, the assumptions of the implicit function theorem apply. Hence, we can write y as a continuously differentiable function of x near $(1, 0)$ if $\frac{\partial f}{\partial y} \neq 0$ at $(1, 0)$.

$$\frac{\partial f}{\partial y}(1, 0) = (7x + 20y)|_{(x,y)=(1,0)} = 7 \neq 0$$

and so $f(x, y) = 0$ defines y as a continuously differentiable function of x near $(1, 0)$.

iii) From the implicit function theorem, $f(x, y) = 0$ defines one of x and y as a continuous function of the other whenever

$$\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)$$

has rank 1. This can only have rank less than 1 if it has rank 0, i.e.

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0.$$

Explicitly, one of x or y is a continuous function of the other near all points of

$$S = \{(x, y) \in \mathbb{R}^2 \mid f(x, y) = 0\}$$

other than those where

$$2x + 7y = 7x + 10y = 0 \quad \Leftrightarrow \quad x = y = 0.$$

But $f(0, 0) = -1 \neq 0$, so $(0, 0) \notin S$ and the implicit function theorem applies at all points of S.

iv) We know that y can be expressed as a continuously differentiable function of x for all $(x, y) \in S$ for which $\frac{\partial f}{\partial y} \neq 0$. Thus the point points of S for which we might not be able to express y as a continuously differentiable function of x are those where

$$7x + 20y = 0.$$

As $(0, 0) \notin S$, the only such points are given by solutions of

$$f(\frac{-20y}{7}, y) = 0 \quad \Leftrightarrow \quad \frac{400y^2}{49} - 20y^2 + 10y^2 = 1.$$

This reduces to

$$-90y^2 = 1$$

which has no real solutions. Thus, we can always express y as a continuously differentiable function of x - $f(x, y) = 0$ defines a hyperbola. The diagram shown in class was a lie.