MAU22203/33203 - Analysis in Several Real Variables

Exercise Sheet 1

Trinity College Dublin

Course homepage

Answers are due for October 4^{th}, 2pm.
The use of electronic calculators and computer algebra software is allowed.

Exercise 1 *Properties of sequences (40pts)*

Let $\{\vec{x}_n\}$ and $\{\vec{y}_n\}$ be two sequences of points in $\mathbb{R}^m$, and let $\lambda \in \mathbb{R}$ be a real number. Suppose that $\{\vec{x}_n\}$ converges to a point $\vec{p}$, and $\{\vec{y}_n\}$ converges to a point $\vec{q}$. By giving a formal ε-N proof, establish the following:

(10pts) The sequence $\{\lambda\vec{x}_n\}$ converges to $\lambda\vec{p}$,

Hint: consider $\lambda = 0$ as a separate case.

(10pts) The sequence $\{\vec{z}_n = \vec{x}_n + \vec{y}_n\}$ converges to $\vec{p} + \vec{q}$.

(20pts) The sequence of real numbers $\{a_n = \langle \vec{x}_n, \vec{y}_n \rangle\}$ given by the inner product of $\vec{x}_n$ with $\vec{y}_n$ converges to the inner product $\langle \vec{p}, \vec{q} \rangle$.

Hint: Using the following, apply the Cauchy-Schwarz inequality

$$\langle \vec{x}, \vec{y} \rangle - \langle \vec{p}, \vec{q} \rangle = \langle \vec{x} - \vec{p}, \vec{y} - \vec{q} \rangle + \langle \vec{p}, \vec{y} - \vec{q} \rangle + \langle \vec{x} - \vec{p}, \vec{q} \rangle.$$

Solution 1

1. First note that the claim is trivial if $\lambda = 0$, so we will assume otherwise. As $\{\vec{x}_n\}$ converges to $\vec{p}$, for every $\varepsilon > 0$, there exists $N > 0$ such that $\|\vec{x}_n - \vec{p}\| < \varepsilon$ for all $n \geq N$. Thus, for all $\varepsilon > 0$, there exists $N_\lambda > 0$ such that
$$\|\vec{x}_n - \vec{p}\| < \frac{\varepsilon}{|\lambda|}$$
for all $n \geq N_\lambda$. Hence, for all $n \geq N_\lambda$
$$\|\lambda\vec{x}_n - \lambda\vec{p}\| = |\lambda|\|\vec{x}_n - \vec{p}\| < |\lambda|\frac{\varepsilon}{|\lambda|} = \varepsilon$$
and so $\{\lambda\vec{x}_n\}$ converges to $\lambda\vec{p}$.

2. By the convergence of $\{\vec{x}_n\}$ and $\{\vec{y}_n\}$, we have that for every $\varepsilon > 0$ there exist $N_x, N_y > 0$ such that
$$\|\vec{x}_n - \vec{p}\| < \frac{\varepsilon}{2} \quad \text{and} \quad \|\vec{y}_n - \vec{q}\| < \frac{\varepsilon}{2}$$
and so, by the triangle inequality
$$\begin{aligned} \|\vec{x}_n + \vec{y}_n - \vec{p} - \vec{q} &= \|(\vec{x}_n - \vec{p}) + (\vec{y}_n - \vec{q})\| \\ &\leq \|\vec{x}_n - \vec{p}\| + \|\vec{y}_n - \vec{q}\| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$
for all $n \geq \max\{N_x, N_y\}$.

3. , As noted in the hint
$$\langle\vec{x},\vec{y}\rangle - \langle\vec{p},\vec{q}\rangle = \langle\vec{x}-\vec{p},\vec{y}-\vec{q}\rangle + \langle\vec{p},\vec{y}-\vec{q}\rangle + \langle\vec{x}-\vec{p},\vec{q}\rangle$$
and so
$$|\langle\vec{x},\vec{y}\rangle - \langle\vec{p},\vec{q}\rangle| \leq |\langle\vec{x}-\vec{p},\vec{y}-\vec{q}\rangle| + |\langle\vec{p},\vec{y}-\vec{q}\rangle| + |\langle\vec{x}-\vec{p},\vec{q}\rangle|.$$
Applying the Cauchy Schwarz inequality to each term individually, we obtain that
$$|\langle\vec{x},\vec{y}\rangle - \langle\vec{p},\vec{q}\rangle| \leq \|\vec{x}-\vec{p}\|\|\vec{y}-\vec{q}\| + \|\vec{p}\|\|\vec{y}-\vec{q}\| + \|\vec{x}-\vec{p}\|\|\vec{q}\|.$$

Applying this to $(\vec{x}_n, \vec{y}_n)$, this gives an upper bound for the difference of our inner products in terms of norms. From convergence, we know that for any $\eta > 0$ there exists $N_x, N_y > 0$ such that

$$\|\vec{x}_n - \vec{p}\| < \eta \quad \text{and} \quad \|\vec{y}_n - \vec{q}\| < \eta$$

for all $n \geq N = \max\{N_x, N_y\}$. Thus, for all $n \geq N$, we have that

$$|\langle \vec{x}_n, \vec{y}_n \rangle - \langle \vec{p}, \vec{q} \rangle| < \eta\,(\eta + \|\vec{p}\| + \|\vec{q}\|)$$

Any reasonable argument from here that, given $\varepsilon > 0$, we can find $\eta > 0$ such that

$$\eta\,(\eta + \|\vec{p}\| + \|\vec{q}\|) \leq \varepsilon$$

should be awarded the majority of the points. Something like, noting that

$$f(x) = x^2 + x(\|\vec{p}\| + \|\vec{q}\|)$$

is positive for positive x, satisfies $f(0) = 0$, and is continuous at 0, and so for any $\varepsilon > 0$, there exists $\eta > 0$ such that $f(\eta) < \varepsilon$ would do. I will instead note that taking

$$\eta = \min\left\{1, \frac{\varepsilon}{1 + \|\vec{p}\| + \|\vec{q}\|}\right\}$$

implies that

$$\eta\,(\eta + \|\vec{p}\| + \|\vec{q}\|) \leq \eta(1 + \|\vec{p}\| + \|\vec{q}\|) \leq \varepsilon$$

Thus, as there exists $N > 0$ such that

$$|\langle \vec{x}_n, \vec{y}_n \rangle - \langle \vec{p}, \vec{q} \rangle| < \eta\,(\eta + \|\vec{p}\| + \|\vec{q}\|) \leq \varepsilon$$

for such η, for any $\varepsilon > 0$, we conclude that the sequence of real numbers $\{a_n = \langle \vec{x}_n, \vec{y}_n \rangle\}$ given by the inner product of $\vec{x}_n$ with $\vec{y}_n$ converges to the inner product $\langle \vec{p}\vec{q} \rangle$.

Exercise 2 *Matrix norms (60pts)*

In the following, you may use any standard facts from your first year courses. Let $A : \mathbb{R}^m \to \mathbb{R}^s$ be a linear transformation represented by the $(s \times m)$-matrix $(A_{i,j})_{\substack{1\le i\le s \\ 1\le j\le m}}$ with respect to the standard bases. We define the Hilbert-Schmidt norm of T by

$$\|A\|_{HS} := \sqrt{\sum_{i=1}^{s}\sum_{j=1}^{m} A_{i,j}^2}.$$

(15 pts) Show that, for any $\vec{x} \in \mathbb{R}^m$,

$$\|A\vec{x}\| \le \|A\|_{HS}\|\vec{x}\|.$$

Hint: what are the components of $A\vec{x}$ and how could we bound them using the Cauchy-Schwarz inequality?

(15 pts) Show that, given linear transformations

$$A : \mathbb{R}^m \to \mathbb{R}^s \quad \text{and} B : \mathbb{R}^s \to \mathbb{R}^t,$$

the Hilbert Schmidt norm satisfies

$$\|BA\|_{HS} \le \|B\|_{HS}\|A\|_{HS}$$

(15 pts) Denoting by A^T the transpose of the matrix A, and by $\mathrm{tr}(M)$ the trace of a square matrix M, show that

$$\|A\|_{HS}^2 = \mathrm{tr}(A^T A)$$

(15 pts) Let $A : \mathbb{R}^3 \to \mathbb{R}^3$ be given by the below matrix, and define a sequence of points in $\mathbb{R}^3$ by $\vec{x}_n := A^{n-1}\vec{x}_1$, where $\vec{x}_1$ is given below. Prove that $\{\vec{x}_n\}$ converges to $\vec{0}$.

$$A = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{5} & \frac{1}{10} & 0 \end{pmatrix}, \quad \vec{x}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Solution 2

1. Letting $\vec{v} = A\vec{x}$, note that

$$v_i = A_{i,1}x_1 + A_{i,2}x_2 + \cdots A_{i,m}x_m$$

is given by the inner product of the i^{th} row of A with $\vec{x}$. Hence, by the Cauchy-Schwarz inequality, we have that

$$v_i^2 = (A_{i,1}x_1 + A_{i,2}x_2 + \cdots A_{i,m}x_m)^2 \leq \left(A_{i,1}^2 + \cdots + A_{i,m}^2\right) \|\vec{x}\|^2$$

. Summing over i, we get that

$$\|\vec{v}\|^2 = \sum_{i=1}^{s} v_i^2 \leq \sum_{i=1}^{s} \left(\sum_{j=1}^{m} A_{i,j}^2\right) \|\vec{x}\|^2$$

and so

$$\|A\vec{x}\|^2 = \|\vec{v}\|^2 \leq \left(\sum_{i=1}^{s}\sum_{j=1}^{m} A_{i,j}^2\right) \|\vec{x}\|^2 = \|A\|_{HS}^2\|\vec{x}\|^2$$

from which the claim follows.

2. Let $C = BA$. We have that

$$C_{i,j} = \sum_{k=1}^{s} B_{i,k}A_{k,j}$$

which we can interpret as an inner product between the i^{th} row of B and the j^{th} column of A. Hence, by Cauchy-Schwarz, we must have that

$$C_{i,j}^2 \leq \left(\sum_{k=1}^{s} B_{i,k}^2\right)\left(\sum_{\ell=1}^{s} A_{\ell,j}^2\right)$$

Summing over i and j, we get

$$\|BA\|_{HS}^2 = \|C\|_{HS}^2 \leq \|B\|_{HS}^2\|A\|_{HS}^2$$

from which the claim follows.

3. The matrix $M = A^T A$ is an $(s \times s)$ matrix, with diagonal entries

$$\begin{aligned} M_{i,i} &= \sum_{k=1}^{m} \left(A^T\right)_{i,k} A_{k,i} \\ &= \sum_{k=1}^{m} A_{k,i} A^{=}_{k,i} A^2_{k,i} \end{aligned}$$

and so

$$\operatorname{tr}(M) = \sum_{i=1}^{s} M_{i,i} = \sum_{i=1}^{s} \sum_{k=1}^{m} A^2_{k,i} = \|A\|^2_{HS}.$$

4. Note that we must have

$$\|\vec{x}_n\| = \|A^{n-1}\vec{x}_1\| \leq \|A^{n-1}\|_{HS}\|\vec{x}_1\| \leq \|A\|^{n-1}_{HS}\|\vec{x}_1\|$$

by the previous parts of the question. We can easily compute

$$\|A\|_{HS} \approx 0.87876\ldots < 1$$

and so

$$\|\vec{x}_n\| < (0.9)^{n-1}\|\vec{x}_1\|.$$

From first year analysis, we know that $(0.9)^n \to 0$, and hence the limit of

$$\|\vec{x}_n\| = \|\vec{x}_n - \vec{0}\|$$

is 0. Hence, we must have that $\{\vec{x}_n\}$ converges to $\vec{0}$.