LECTURE 5

Date of Lecture: January 29, 2020

Let $p \in [1, \infty]$. For $\boldsymbol{x} \in \mathbf{K}^n$ and $r > 0$ we set $B_p(\boldsymbol{x}, r)$ equal to the ball of radius r centred at $\boldsymbol{x}$ in the $\| \, \|_p$ norm.

1. Compactness

1.1. **The Heine-Borel Theorem.** In Proposition 2.1.3 of the last lecture, we showed that if $\mathscr{Q}$ is an collection of open cubes in $\mathbf{K}^n$ which cover a closed bounded set $K = J_1 \times \cdots \times J_n$ of $\mathbf{K}^n$, where the J_i are closed bounded intervals in $\mathbf{R}$, $i = 1, \ldots, n$, then $\mathscr{Q}$ has a finite subcover. In fact, as we will see below, K is compact.

Theorem 1.1.1. (The Heine-Borel Theorem in $\mathbf{K}^n$) *A subset of* $\mathbf{K}^n$ *is compact with respect to the* $\| \, \|_2$ *norm if and only if it is closed and bounded.*

Proof. According to Proposition 1.1.7 of Lecture 4, a compact subset of a metric space is necessarily closed and bounded. So we have to show that a closed and bounded set in $\mathbf{K}^n$ is compact.

Let S be a closed and bounded in $(\mathbf{K}^n, \| \, \|_2)$. Then it is closed and bounded in $(\mathbf{K}^n, \| \, \|_\infty)$. Hence there is a closed box $K = J_1 \times \cdots \times J_n$ such that S is a subset of K. By Proposition 1.1.4 of Lecture 4, if K is compact, so is S. So it is enough to show that K is compact.

Suppose $\mathscr{U} = \{U_\lambda \mid \lambda \in \Lambda\}$ is an open cover of K. For each $\boldsymbol{x} \in K$, we have $\lambda(\boldsymbol{x}) \in \Lambda$ such that $\boldsymbol{x} \in U_{\lambda(\boldsymbol{x})}$. Since $U_{\lambda(\boldsymbol{x})}$ is open in $\mathbf{K}^n$ we have an open cube $Q_{\boldsymbol{x}}$ such that $\boldsymbol{x} \in Q_{\boldsymbol{x}} \subset U_{\lambda(\boldsymbol{x})}$. Indeed, we can find $B_2(\boldsymbol{x}, r) \subset U_{\lambda(\boldsymbol{x})}$, and if we set $Q_{\boldsymbol{x}} = B_\infty(\boldsymbol{x}, \frac{r}{\sqrt{n}})$ then $Q_{\boldsymbol{x}} \subset B_2(\boldsymbol{x}, r)$. Let $\mathscr{V} = \{Q_{\boldsymbol{x}} \mid \boldsymbol{x} \in K\}$. Clearly $\mathscr{V}$ is an open cover of K by open cubes. By Proposition 2.1.3 of Lecture 4, we can find $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m \in K$ such that $\{Q_{\boldsymbol{x}_1}, \ldots, Q_{\boldsymbol{x}_m}\}$ is an open cover of K. Thus

$$K \subset \bigcup_{i=1}^{m} Q_{\boldsymbol{x}_i} \subset \bigcup_{i=1}^{m} U_{\lambda(\boldsymbol{x})},$$

i.e. $\{U_{\lambda(\boldsymbol{x}_1)}, \ldots, U_{\lambda(\boldsymbol{x}_m)}\}$ is a finite subcover of $\mathscr{U}$, showing that K is compact. $\square$

Corollary 1.1.2. *All norms on* $\mathbf{K}^n$ *are equivalent.*

Proof. This was done in your tutorial on January 23 using the Heine-Borel theorem for $\mathbf{K}^n$. $\square$

Corollary 1.1.3. *Let* $\| \, \|$ *be any norm on* $\mathbf{K}^n$. *A subset* S *of* $\mathbf{K}^n$ *is compact with respect to* $\| \, \|$ *if and only if it is closed and bounded with respect to* $\| \, \|$.

Proof. This is immediate from Theorem 1.1.1 and Corollary 1.1.2. $\square$

Remark 1.1.4. For $p \in [1, \infty]$ it is clear that $\| \, \|_p$ has all the properties of a norm on $\mathbf{K}^n$ except possibly the triangle inequality, i.e. the inequality $\|\boldsymbol{x} + \boldsymbol{y}\|_p \leq \|\boldsymbol{x}\|_p + \|\boldsymbol{y}\|_p$ for $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{K}^n$. According to Problem 2(b) of Quiz 2 and Problem

7 of Homework 3, this inequality holds for $p \in (1, \infty)$. Thus $\| \, \|_p$ is a norm for $p \in (1, \infty)$. Let us prove this when $p \in \{1, \infty\}$. Once again, the only non-trivial matter is the triangle inequality. For $(x_1, \dots, x_n)$ and $(y_1, \dots, y_n)$ in $\mathbf{K}^n$ we have

$$\begin{aligned}
\|(x_1, \dots, x_n) + (y_1, \dots, y_n)\|_1 &= \sum_{i=1}^{n} |x_i + y_i| \\
&\leq \sum_{i=1}^{n} \left(|x_i| + |y_i|\right) \\
&= \|(x_1, \dots, x_n)\|_1 + \|(y_1, \dots, y_n)\|_1,
\end{aligned}$$

showing that $\| \, \|_1$ is a norm. Similarly

$$\begin{aligned}
\|(x_1, \dots, x_n) + (y_1, \dots, y_n)\|_\infty &= \max_{i=1,\dots,n} |x_i + y_i| \\
&\leq \max_{i=1,\dots,n} \left(|x_i| + |y_i|\right) \\
&\leq \max_{i=1,\dots,n} |x_i| + \max_{i=1,\dots,n} |y_i| \\
&= \|(x_1, \dots, x_n)\|_\infty + \|(y_1, \dots, y_n)\|_\infty,
\end{aligned}$$

proving that $\| \, \|_\infty$ is a norm.

2. The derivative at a point for a function of several variables

2.1. **Norms on linear transformations.** Let $L(\mathbf{K}^m, \mathbf{K}^m)$ be the space of $\mathbf{K}$-linear transformations from $\mathbf{K}^m$ to $\mathbf{K}^n$. Let

$$T \colon \mathbf{K}^m \to \mathbf{K}^n$$

be a $\mathbf{K}$-linear transformation. From Corollary 1.1.2 Problem 4 of Homework 2, we know that T is continuous on $\mathbf{K}^m$ with respect to any norm on $\mathbf{K}^m$ and $\mathbf{K}^n$. For definiteness, let us give the Euclidean norm to both these spaces, and write $\| \, \|$ instead of $\| \, \|_2$ for both spaces. Since the unit sphere $S = \{\boldsymbol{x} \in \mathbf{K}^m \mid \|\boldsymbol{x}\| = 1\}$ is closed and bounded it is compact. Since T is continuous and S is compact, $\sup_{\boldsymbol{x} \in S} \|T\boldsymbol{x}\|$ is finite. Set

$$\|T\|_L = \sup_{\boldsymbol{x} \in S} \|T\boldsymbol{x}\| < \infty.$$

We thus have a map

$$\| \, \|_L \colon L(\mathbf{K}^m, \mathbf{K}^m) \longrightarrow [0, \infty) \tag{2.1.1}$$

It is straightforward to check that $\| \, \|_L$ is a norm on $L(\mathbf{K}^m, \mathbf{K}^n)$ (this was done in class, but the proof is straightforward). When the context is clear, we will drop the subscript L from $\| \, \|_L$ and simply write $\| \, \|$.

2.2. **Derivative at a point.** If U is an open set in $\mathbf{R}$, a a point in U, and $f \colon U \to \mathbf{R}$ a function which is differentiable at a, then the derivative $f'(a)$ can be regarded as a linear transformation $T \colon \mathbf{R} \to \mathbf{R}$, namely the linear transformation $T(x) = f'(a)x$, $x \in \mathbf{R}$. We are simply using the fact that $\mathbf{R}$ can be identified with 1×1 real matrices, and therefore $f'(a) \in \mathbf{R}$ can be regarded as a 1×1 real matrix. This is the idea that is generalised to many variables.

Definition 2.2.1. Let U be open in $\mathbf{R}^m$, $\boldsymbol{f}\colon U \to \mathbf{R}^n$ a map, and $\boldsymbol{a}$ a point in U. The map $\boldsymbol{f}$ is said to be differentiable at $\boldsymbol{a}$ if there exists $T \in L(\mathbf{R}^m,\, \mathbf{R}^n)$ such that

$$\lim_{\boldsymbol{h}\to\mathbf{0}} \frac{\|\boldsymbol{f}(\boldsymbol{a}+\boldsymbol{h}) - \boldsymbol{f}(\boldsymbol{a}) - T\boldsymbol{h}\|}{\|\boldsymbol{h}\|} = 0.$$

If this happens, we call T the derivative of $\boldsymbol{f}$ at $\boldsymbol{a}$. We will show that the derivative is well defined when it exists. If $\boldsymbol{f}$ is differentiable at $\boldsymbol{a}$, we write $\boldsymbol{f}'(\boldsymbol{a})$ for its derivative at $\boldsymbol{a}$.

We will see that if $f_i\colon U \to \mathbf{R}$, is the map $\pi_i \circ \boldsymbol{f}$, $i = 1, \ldots, n$, so that $\boldsymbol{f} = (f_1, \ldots, f_n)$, and if $\boldsymbol{f}$ is differentiable at $\boldsymbol{a}$, then the matrix of $\boldsymbol{f}'(\boldsymbol{a})$ in the standard basis for $\mathbf{R}^m$ and $\mathbf{R}^n$ is:

$$\begin{bmatrix} \frac{\partial f_1(\boldsymbol{a})}{\partial x_1} & \cdots & \frac{\partial f_1(\boldsymbol{a})}{\partial x_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n(\boldsymbol{a})}{\partial x_1} & \cdots & \frac{\partial f_n(\boldsymbol{a})}{\partial x_m} \end{bmatrix}$$

We will in general not distinguish between $L(\mathbf{R}^m,\, \mathbf{R}^n)$ and the space of $m \times n$ matrices $M_{\mathbf{R}}(n, m)$.

Remark 2.2.2. For any two sets A and B, let $\mathrm{Hom}(A,\, B)$ be the set of maps from A to B. A typical element of $\mathrm{Hom}(A,\, B)$ is sometimes identified with a subset f of $A \times B$ such that for each $a \in A$, there is exactly one $b \in B$ with $(a, b) \in f$. (In other words we identify a function with its "graph".)[1] Let $A, B_1, \ldots, B_n$ be sets. It is easy to see that there is a natural identification of $\mathrm{Hom}(A, \prod_{i=1}^n B_i)$ with $\prod_{i=1}^n \mathrm{Hom}(A,\, B_i)$. If one uses the somewhat pedantic "definition" of elements in $\mathrm{Hom}(A,\, B)$ described above, then the natural identification is the one-to-one and onto map induced by the map which sends $(a, (b_1, \ldots, b_n))$ to $((a,\, b_1), \ldots, (a, b_n))$. This one-to-one onto map

$$\mathrm{Hom}(A, \prod_{i=1}^{n} B_i) \longrightarrow \prod_{i=1}^{n} \mathrm{Hom}(A,\, B_i) \tag{2.2.2.1}$$

is a "functorial" identification (what is called a natural transformation) in all the "arguments", namely, in A, $B_1, \ldots, B_n$, and hence one treats such an identification as an equality. It is in this sense that we regard a map $\boldsymbol{f}\colon U \to \mathbf{R}^n$, where $U \subset \mathbf{R}^m$, as an n-tuple of maps and write $\boldsymbol{f} = (f_1, \ldots, f_n)$. For those smitten by such matters, an isomorphism in the "category of sets" is a one-to-one onto map, and in any category, natural transformations which are isomorphisms are often treated as equalities. This jumping through hoops is only required if you follow the pedantic route described above for the definition of a map. Thus (2.2.2.1) can (and will, for this and most courses in mathematics) be written as

$$\mathrm{Hom}(A, \prod_{i=1}^{n} B_i) = \prod_{i=1}^{n} \mathrm{Hom}(A,\, B_i). \tag{2.2.2.2}$$

---

[1] When A or B is empty one has take care with one's definitions, but we will pass over that in silence.

About these notes. These course notes are a reasonably faithful record of the lectures given at the Chennai Mathematical Institute (CMI) in the January-April semester of 2019-20. The course is Analysis II, a core course for first year undergraduates at CMI. For more material related to this course, visit `https://www.cmi.ac.in/~pramath/teaching.html#ANA2`. If you have comments on these notes, or on related course material, please send an email to pramath@cmi.ac.in.