LECTURE 7

Date of Lecture: February 5, 2020

As always, $\mathbf{K} \in \{\mathbf{R}, \mathbf{C}\}$.

The symbol ☡ is for flagging a cautionary comment or a tricky argument. It occurs in the margins and is Knuth's version of Bourbaki's "dangerous bend symbol".

An n-tuple $(x_1, \dots, x_n)$ of symbols (x_i not necessarily real or complex numbers) will also be written as a column vector when convenient. Thus

$$(x_1, \dots, x_n) = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

A map $\boldsymbol{f}$ from a set S to a product set $T_1 \times \cdots \times T_n$ will often be written as an n-tuple $\boldsymbol{f} = (f_1, \dots, f_n)$, with f_i a map from S to T_i, and hence, by the above convention, as a column vector

$$\boldsymbol{f} = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}.$$

(See Remark 2.2.2 of Lecture 5.)

The default norm on Euclidean spaces of the form $\mathbf{R}^n$ is the Euclidean norm $\| \ \|_2$ and we will simply denote it as $\| \ \|$.

☡ Note that $(x_1, \dots, x_n) \neq [x_1 \dots x_n]$. Each side is the transpose of the other.

1. Elementary but important results

There are a couple of results that we have talked about or used, but which have not been put into earlier notes due to oversight. These too have been put below with adequate reference to where they have been used earlier.

1. *If $f\colon X \to Y$ is a continuous map between metric spaces and K is a compact subset of X, then $f(K)$ is compact. In particular, if $Y \subset \mathbf{K}^n$, then $f(K)$ is closed and bounded, and if $Y \subset \mathbf{R}$, then $f|_K$ attains its maximum and minimum.*

This was used in the proof that all norms on a finite dimensional vector space over $\mathbf{K}$ are equivalent (done in a tutorial).

Proof. If $\mathscr{V} = \{V_\alpha\}$ is an open cover of $f(K)$, then $\{f^{-1}(V_\alpha)\}$ is an open cover of K, and hence there exist $\alpha_1, \dots, \alpha_k$ such that $\{f^{-1}(V_{\alpha_1}), \dots, f^{-1}(V_{\alpha_k})\}$ is an open cover of K. It follows that $\{V_{\alpha_1}, \dots, V_{\alpha_k}\}$ is an open cover of $f(K)$. The remaining assertions are obvious from the Heine-Borel theorem, and the fact that a closed and bounded set in $\mathbf{R}$ must necessarily contain its greatest lower bound and its least upper bound.

2. *The composite of continuous maps on metric spaces is continuous.*

Proof. Suppose $X \xrightarrow{f} Y \xrightarrow{g} Z$ is a pair of continuous maps on metric spaces. Let $\{x_n\}$ be a convergent sequence in X. Then

$$\begin{aligned}(g \circ f)(\lim_{n\to\infty} x_n) &= g(f(\lim_{n\to\infty} x_n)) \\ &= g(\lim_{n\to\infty} f(x_n)) \qquad \text{(since } f \text{ is continuous)} \\ &= \lim_{n\to\infty} g(f(x_n)) \qquad \text{(since } g \text{ is continuous)} \\ &= \lim_{n\to\infty} (g \circ f)(x_n).\end{aligned}$$

Thus $g \circ f$ is continuous.

3. *Let X be a metric space, $f_i \colon X \to \mathbf{K}$ continuous maps for $i = 1, \dots, n$, and $p \colon \mathbf{K}^n \to \mathbf{K}$ a polynomial. Then $p(f_1, \dots, f_n) \colon X \to \mathbf{K}$ is continuous.*

Proof. If $\{x_n\}$ and $\{y_n\}$ are convergent sequences in $\mathbf{K}$, say $\lim_{n\to\infty} x_n = x$ and $\lim_{n\to\infty} y_n = y$, then we know that $\{x_n + y_n\}$ and $\{x_n y_n\}$ converge to $x + y$ and and xy respectively. This means that the maps $\mathbf{K}^2 \xrightarrow{+} \mathbf{K}$ and $\mathbf{K}^2 \xrightarrow{\cdot} \mathbf{K}$ are continuous. By repeated application of this we see that $p \colon \mathbf{K}^n \to \mathbf{K}$ is continuous. By Problem 3 of Homework 2 (replacing the normed space X there by the metric space X here, without loss of generality) we see that the map $\boldsymbol{f} \colon X \to \mathbf{K}^n$, where $\boldsymbol{f} = (f_1, \dots, f_n)$ is continuous. It follows that $p(f_1, \dots, f_n)$ is continuous on X by item **2.** since $p(f_1, \dots, f_n) = p \circ \boldsymbol{f}$.

4. *Let $f \colon X \to \mathbf{K} \smallsetminus \{0\}$ be a continuous map from a metric space X. Then $\dfrac{1}{f} \colon X \to \mathbf{K} \smallsetminus \{0\}$ is also continuous.*

Proof. The map $I \colon \mathbf{K} \smallsetminus \{0\} \to \mathbf{K} \smallsetminus \{0\}$ given by $I(x) = \frac{1}{x}$ is continuous. Since $\dfrac{1}{f} = I \circ f$, we are done by **2.** above.

5. *If $A \colon V \to W$ is a linear transformation between finite dimensional normed linear spaces, and $\|A\|$ its operator norm, then $\|Ax\| \le \|A\|\|x\|$ for all $x \in V$.*

This was used implicitly in the proof of the chain rule, i.e. Theorem 2.1.7 of Lecture 6.

Proof. If $x = 0$, there is nothing to prove. Otherwise set $u = x/\|x\|$. By definition of $\|A\|$, $\|Au\| \le \|A\|$, and hence we done.

6. *If $A \colon V \to W$ is as above and there is a real number M such that $\|Ax\| \le M\|x\|$ for every $x \in V$, then $\|A\| \le M$. In fact $\|A\|$ is the infimum of such numbers M.*

Proof. Let $S = \{M \in \mathbf{R} \mid \|Ax\| \le M\|x\| \quad \text{for } x \in V\}$. Now, $\|A\| \in S$ by **5.** Moreover, if $M \in S$ then $\|A\| = \sup_{\|x\|=1}\|Ax\| \le \sup_{\|x\|=1} M\|x\| = M$. Thus $\|A\| = \inf_{M\in S} M$.

7. *If $U \xrightarrow{A} V \xrightarrow{B} W$ are a pair of linear transformations between finite dimensional normed vector spaces over $\mathbf{K}$ then $\|BA\| \leq \|B\|\|A\|$.*

Proof. If $u \in U$ is such that $\|u\| = 1$, then $\|BAu\| \leq \|B\|\|Au\| \leq \|B\|\|A\|$. Taking supremums we get the result.

8. *Suppose U, V, W are as above and for each $n \in \mathbf{N}$ we have a pair of maps $U \xrightarrow{A_n} V \xrightarrow{B_n} W$. Suppose further that $A_n \to A$ and $B_n \to B$ as $n \to \infty$. Then $B_nA_n \to BA$ as $n \to \infty$. In particular, if $L(U, V) \oplus L(V, W)$ is given any norm, say $\|(S, T)\| = \|S\| + \|T\|$, for $(S, T) \in L(U, V) \oplus L(V, W)$, then the map $(S, T) \mapsto TS$ is a continuous map.* (All limits are with respect to the respective the operator norms.)

Proof. Since $\{A_n\}$ is convergent, and since $L(V, W)$ is a complete, it is a bounded sequence, i.e. there exists $M \in \mathbf{R}$ such that $\|A_n\| \leq M$ for every $n \in \mathbf{N}$. Then

$$\begin{aligned}
\|B_nA_n - BA\| &= \|(B_n - B)A_n + B(A_n - A)\| \\
&\leq \|(B_n - B)A_n\| + \|B(A_n - A)\| \\
&\leq \|B_n - B\|\|A_n\| + \|B\|\|A_n - A\| \\
&\leq M\|B_n - B\| + \|B\|\|A_n - A\| \\
&\to 0
\end{aligned}$$

as $n \to \infty$. Thus $B_nA_n \to BA$ as $n \to \infty$.

9. *Let $M_{mn} = M_{mn}(\mathbf{K})$ be the space of $m \times n$ matrices over $\mathbf{K}$. We identify this with $L(\mathbf{K}^n, \mathbf{K}^m)$ in the usual way. The default norm on M_{mn} is the operator norm on $L(\mathbf{K}^n, \mathbf{K}^m)$ under the above identification. Let $k \leq \min\{m, n\}$. If $I \subset \{1, \ldots, m\}$ and $J \subset \{1, \ldots, n\}$ are subsets of cardinality k and if for any $A \in M_{mn}$, A_{IJ} is the $k \times k$ submatrix of A obtained by selecting the k rows given by I and the k columns given by J, then the map $\Delta_{IJ} \colon M_{mn} \to \mathbf{K}$ given by $\Delta_{IJ}(A) = \det A_{IJ}$, is continuous.*

Proof. Δ_{IJ} is a polynomial in the entries of $A \in M_{mn} \xrightarrow{\sim} \mathbf{K}^{mn}$, and hence by item **3.** above and the fact that all norms on a finite dimensional vector space over $\mathbf{K}$ are equivalent, we are done.

10. *Let $GL_n = GL_n(\mathbf{K})$ be the set of invertible $n \times n$ matrices over $\mathbf{K}$ and $M_n = M_{nn}$ (we write $M_n(\mathbf{K})$ for M_n if we wish to emphasise the role of the field $\mathbf{K}$). Let $\mathbf{Inv} \colon GL_n \to GL_n$ be the map $A \mapsto A^{-1}$. Then GL_n is an open subset of M_n and $\mathbf{Inv}$ is continuous.*

Proof. By **9.**, the map $\det \colon M_n \to \mathbf{K}$ is continuous. Now $GL_n = \det^{-1}(\mathbf{K} \smallsetminus \{0\})$. Since $\mathbf{K} \smallsetminus \{0\}$ is open in $\mathbf{K}$, we see that GL_n is open in M_n. By Cramer's rule, as well as **4.** and **9.** above, it is clear that $\mathbf{Inv}$ is continuous.

2. Infinite series representation for $(I - A)^{-1}$

2.1. **Infinite series in $L(V, W)$.** Let V and W be finite dimensional normed linear spaces over $\mathbf{K}$ and $A_1, \ldots, A_n, \ldots$ a sequence of elements in $L(V, W)$, the $\mathbf{K}$ vector space of K-linear transformations from V to W. We endow $L(V, W)$ with the operator norm. We say that the infinite series $\sum_{n=1}^{\infty} A_n$ *converges* if the sequence of partial sums $\{S_n\}$, $S_n = \sum_{k=1}^{n} A_k$, $n \in \mathbf{N}$, converges in $L(V, W)$.[1] Though the operator norm on $L(V, W)$ depends upon the norms on V and W, the notion of convergence of a sequence or a series in $L(V, W)$ is independent of these choices. More precisely, suppose for a certain choice of norms on V and W, the operator norm on $L(V, W)$ is $\| \, \|$ and for another choice of norms on V and W, the induced norm on $L(V, W)$ is $\| \, \|'$. Then, $L(V, W)$ being finite dimensional, the two norms $\| \, \|$ and $\| \, \|'$ are equivalent and hence $\sum_n A_n$ converges to A in the norm $\| \, \|$ if and only if it does so in $\| \, \|'$.

Recall that we write $L(V)$ for $L(V, V)$ for $\mathbf{K}$-vector spaces V.

Theorem 2.1.1. *Let V, be a finite dimensional normed linear spaces over K. Let $A \in L(V)$ be such that $\|A\| < 1$, where $\| \, \| \colon L(V) \to \mathbf{R}_+$ is the operator norm. Then*

(a) *$\sum_{n=0}^{\infty} A^n$ is convergent.*
(b) *$I - A$ is invertible, where I is the identity linear transformation. Moreover*

$$(I - A)^{-1} = \sum_{n=0}^{\infty} A^n.$$

Proof. Let $\{S_n\}$ be the sequence of partial sums of $\sum_{k=0}^{\infty} A^k$. Then for $m < n$ we have $\|S_n - S_m\| = \|\sum_{k=m+1}^{n} A^k\| \le \sum_{k=m+1}^{n} \|A\|^k$. Since $\|A\| < 1$, $\sum_{n=0}^{\infty} \|A\|^n < \infty$, whence given $\epsilon > 0$, there exists $N \in \mathbf{N}$ such that $\sum_{k=m+1}^{n} \|A\|^k < \epsilon$ for $n > m \ge N$. It follows that $\{S_n\}$ is a Cauchy sequence. Since $L(V)$ is complete, being a finite dimensional space,[2] it follows that $\sum_{n=0}^{\infty} A^n$ is convergent. This gives (a).

Using the same notations as above we have $S_n(I - A) - I = -A^{n+1}$ for all $n \in \mathbf{N}$, whence $\|S_n(I - A) - I\| = \|A^{n+1}\| \le \|A\|^{n+1} \to 0$ as $n \to \infty$, since $\|A\| < 1$. Thus $(\sum_{n=0}^{\infty} A^n)(I - A) = I$, giving (b). ☐

3. Mean Value Theorems

3.1. **Two versions of the mean value theorem.** Recall that if $\varphi \colon [a, b] \to \mathbf{R}$ is a continuous map which is differentiable on (a, b), then there is a point $\theta \in (a, b)$ such that $f(b) - f(a) = f'(\theta)(b - a)$. This is the so-called (Lagrange's) *Mean Value Theorem* in one variable.

Here are two versions in higher dimensions. First note that if $\boldsymbol{g} \colon (a, b) \to \mathbf{R}^n$ is differentiable, then for each $t \in (a, b)$, $\boldsymbol{g}'(t) \in L(\mathbf{R}, \mathbf{R}^n)$ is identified with an $n \times 1$ matrix, i.e. with a column vector, and hence as a point of $\mathbf{R}^n$. Moreover, clearly the operator norm on an $n \times 1$ matrix is the same as its norm as a vector in $\mathbf{R}^n$, for if $\boldsymbol{v}$ is an $n \times 1$ matrix, then for $x \in \mathbf{R}$, $\|\boldsymbol{v}x\| = \|\boldsymbol{v}\||x|$ forcing the operator norm on $\boldsymbol{v}$ to equal $\|\boldsymbol{v}\|$ by item **6**. in § 1. Recall that the default norm we are using on $\mathbf{R}^n$ is the Euclidean norm.

---

[1] Sometimes we may use another indexing set, for example $\sum_{n=0}^{\infty} A_n$, for an infinite series of matrices, and it is obvious what is meant by convergence in such cases.

[2] See Problem 6, Homework 1.

Theorem 3.1.1. (First Version) *Let* $\boldsymbol{g}\colon [a,b] \to \mathbf{R}^n$ *be a continuous function which is differentiable on* (a,b)*. Then there exists* $\theta \in (a,b)$ *such that*

$$\|\boldsymbol{g}(b) - \boldsymbol{g}(a)\| \le (b-a)\|\boldsymbol{g}'(\theta)\|.$$

Proof. Let $\boldsymbol{z} = \boldsymbol{g}(b) - \boldsymbol{g}(a)$. Let $\langle\ ,\ \rangle$ be the usual inner product on $\mathbf{R}^n$. Let $\varphi\colon [a,\,b] \to \mathbf{R}$ be the map given by $\varphi(t) = \langle \boldsymbol{z},\, \boldsymbol{g}(t)\rangle$. Note that

$$\varphi(b) - \varphi(a) = \langle \boldsymbol{z}, \boldsymbol{g}(b) - \boldsymbol{a}\rangle = \langle \boldsymbol{z}, \boldsymbol{z}\rangle = \|\boldsymbol{z}\|^2.$$

It is easy to see that φ is continuous and that on (a,b) it is differentiable with derivative $\varphi'(t) = \langle \boldsymbol{z},\, \boldsymbol{g}'(t)\rangle$. By the one variable mean value theorem, we therefore have $\theta \in (a,b)$ such that

$$\|\boldsymbol{z}\|^2 = \varphi(b) - \varphi(a) = (b-a)\varphi'(\theta) = (b-a)\langle \boldsymbol{z}, \boldsymbol{g}'(\theta)\rangle.$$

By Cauchy-Schwarz this means

$$\|\boldsymbol{z}\|^2 \le (b-a)\|\boldsymbol{z}\|\|\boldsymbol{g}'(\theta)\|.$$

The assertion in the theorem is trivial if $\boldsymbol{z} = \mathbf{0}$. If $\boldsymbol{z} \ne \mathbf{0}$ we can cancel $\|\boldsymbol{z}\|$ from the inequality above to get the theorem. ☐

Theorem 3.1.2. (Second Version) *Let* U *be a convex open subset of* $\mathbf{R}^n$ *and* $\boldsymbol{f}\colon U \to \mathbf{R}^m$ *a differentiable map. Suppose there is a real number* M *such that* $\|f'(\boldsymbol{x})\| \le M$ *for* $\boldsymbol{x} \in U$*. Then for any two points* $\boldsymbol{a}$ *and* $\boldsymbol{b}$ *in* U

$$\|\boldsymbol{f}(\boldsymbol{b}) - \boldsymbol{f}(\boldsymbol{a})\| \le M\|\boldsymbol{b} - \boldsymbol{a}\|.$$

Proof. Let $\boldsymbol{\gamma}\colon [0,1] \to \mathbf{R}^n$ be the map $\boldsymbol{\gamma}(t) = \boldsymbol{a} + t(\boldsymbol{b} - \boldsymbol{a})$. Since U is convex, $\boldsymbol{\gamma}(t) \in U$ for every $t \in [0,1]$. Let

$$\boldsymbol{g}\colon [0,1] \longrightarrow \mathbf{R}^m$$

be the map $\boldsymbol{g} = \boldsymbol{f} \circ \boldsymbol{\gamma}$. By Theorem 3.1.1 we see that there is a $\theta \in (0,1)$ such that

$$\|\boldsymbol{g}(0) - \boldsymbol{g}(1)\| \le \|\boldsymbol{g}'(\theta)\|.$$

Since $\boldsymbol{g}'(\theta) = \boldsymbol{f}'(\boldsymbol{\gamma}(\theta))(\boldsymbol{b} - \boldsymbol{a})$, we get

$$\|\boldsymbol{f}(\boldsymbol{b}) - \boldsymbol{f}(\boldsymbol{a})\| = \|\boldsymbol{g}(0) - \boldsymbol{g}(1)\| \le \|\boldsymbol{g}'(\theta)\| \le \|\boldsymbol{f}'(\boldsymbol{\gamma}(\theta))\|\|\boldsymbol{b} - \boldsymbol{a}\| \le M\|\boldsymbol{b} - \boldsymbol{a}\|.$$

☐

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.