LECTURE 9

Date of Lecture: February 12, 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. The Inverse Function Theorem

Throughout we assume U is a convex open set in $\mathbf{R}^n$, $\mathbf{0} \in U$ and $\boldsymbol{f} \colon U \to \mathbf{R}^n$ is a $\mathscr{C}^1$ map with $\boldsymbol{f}(\mathbf{0}) = \mathbf{0}$ and $\boldsymbol{f}'(\mathbf{0}) = I$, the identity linear transformation on $\mathbf{R}^n$. We will prove that there are open neighbourhoods V and W of $\mathbf{0}$, $V \subset U$, such that $\boldsymbol{f}$ is one-to-one on V, $\boldsymbol{f}(V) = W$, and the inverse map $\boldsymbol{\varphi} \colon W \to V$ to $\boldsymbol{f}|_V$ is in $\mathscr{C}^1$. Next lecture we will formally state the full fledged inverse function theorem.

1.1. $\boldsymbol{f}$ is one-to-one.

Since $\boldsymbol{f}'$ is continuous on U and the operator norm $\| \; \|$ is continuous on $L(\mathbf{R}^n, \mathbf{R}^m)$, therefore we have an open neighbourhood of $\mathbf{0}$, which we can take to be convex, on which $\|I - \boldsymbol{f}'(\boldsymbol{x})\| \leq \frac{1}{2}$. By replacing U by this neighbourhood, we assume that $\|I - \boldsymbol{f}'(\boldsymbol{x})\| \leq \frac{1}{2}$ for all $\boldsymbol{x} \in U$. Then

1. $\boldsymbol{f}'(\boldsymbol{x}) = I - (I - \boldsymbol{f}'(\boldsymbol{x}))$ is invertible, for $\boldsymbol{x} \in U$, by part (b) of Theorem 2.1.1 of Lecture 7. Indeed $\|I - \boldsymbol{f}'(\boldsymbol{x})\| < 1$ for $\boldsymbol{x} \in U$.
2. Let $\boldsymbol{g} \colon U \to \mathbf{R}^n$ be the map given by $\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{x}$. Then $\boldsymbol{g}$ is $\mathscr{C}^1$ and $\|\boldsymbol{g}'(\boldsymbol{x})\| = \|\boldsymbol{f}'(\boldsymbol{x}) - I\| \leq \frac{1}{2}$ for all $\boldsymbol{x} \in U$. In particular, by the version of the mean value theorem in Theorem 3.1.2 of Lecture 7 we get

$$\|\boldsymbol{g}(\boldsymbol{x}_1) - \boldsymbol{g}(\boldsymbol{x}_2)\| \leq \tfrac{1}{2}\|\boldsymbol{x}_1 - \boldsymbol{x}_2\| \qquad (\boldsymbol{x}_1, \boldsymbol{x}_2 \in U).$$

For $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ in U, from 2. above we get

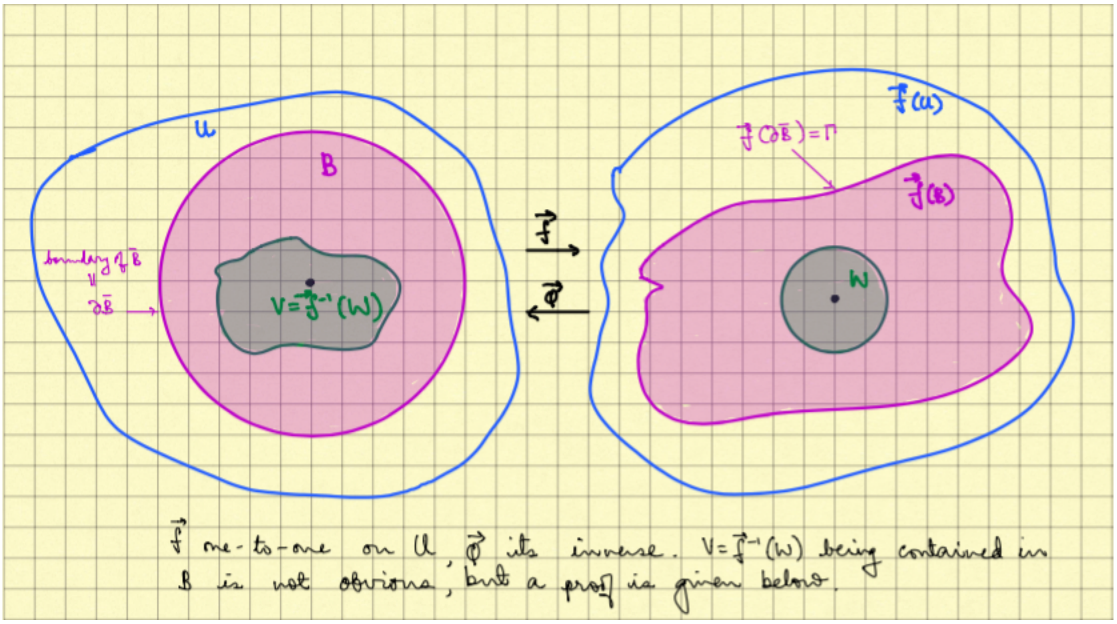
$$\begin{aligned}\|\boldsymbol{x}_1 - \boldsymbol{x}_2\| - \|\boldsymbol{f}(\boldsymbol{x}_1) - \boldsymbol{f}(\boldsymbol{x}_2)\| &\leq \|\boldsymbol{x}_1 - \boldsymbol{x}_2 - (\boldsymbol{f}(\boldsymbol{x}_1) - \boldsymbol{f}(\boldsymbol{x}_2))\| \\ &= \|\boldsymbol{g}(\boldsymbol{x}_1) - \boldsymbol{g}(\boldsymbol{x}_2)\| \\ &\leq \tfrac{1}{2}\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|\end{aligned}$$

whence

$$\tfrac{1}{2}\|\boldsymbol{x}_1 - \boldsymbol{x}_2\| \leq \|\boldsymbol{f}(\boldsymbol{x}_1) - \boldsymbol{f}(\boldsymbol{x}_2)\|. \tag{1.1.1}$$

From (1.1.1) it is clear that $\boldsymbol{f}(\boldsymbol{x}_1) = \boldsymbol{f}(\boldsymbol{x}_2)$ if and only if $\boldsymbol{x}_1 = \boldsymbol{x}_2$. Thus $\boldsymbol{f}$ is one-to-one.[1]

Let $\overline{B}$ be a closed ball in $\mathbf{R}^n$ centered at $\mathbf{0}$ such that $\overline{B} \subset U$. Such a closed ball always exists. Indeed, since U is open, we can find an open ball of radius $r > 0$ centred at $\mathbf{0}$ contained in U, and we can take $\overline{B}$ to be the closed ball of radius $s = r/2$ centred at $\mathbf{0}$. Let B be the open ball which is the interior of B. Let $\partial(\overline{B}) = \{\boldsymbol{x} \in \mathbf{R}^n \mid \|\boldsymbol{x}\| = s\}$ be the sphere which is the boundary $\overline{B}$ (which is the same as the boundary of B) and let $\Gamma = f(\partial(\overline{B}))$.



$\vec{f}$ one-to-one on U, $\vec{\phi}$ its inverse. $V = \vec{f}^{-1}(W)$ being contained in B is not obvious, but a proof is given below.

Note that $\boldsymbol{x} \neq 0$ for any $\boldsymbol{x} \in \partial(\overline{B})$, and since $\boldsymbol{f}$ is one-to-one, $\boldsymbol{f}(\boldsymbol{x}) \neq \mathbf{0}$ for any $\boldsymbol{x} \in \partial(\overline{B})$. Since the set $\partial(\overline{B})$ is compact, and since $\boldsymbol{f}$ is continuous, this means that if $\delta = \inf_{\boldsymbol{x} \in \partial(\overline{B})} \|\boldsymbol{f}(\boldsymbol{x})\|$, then $\delta > 0$. Note that δ is the distance from Γ to $\mathbf{0}$. Let

$$d = \tfrac{1}{2}\delta \quad \text{and} \quad W = B(\mathbf{0}, d).$$

Claim: $W \subset \boldsymbol{f}(B)$.

First note that if $\boldsymbol{y} \in W$ and $\boldsymbol{\gamma} \in \Gamma$, then $\|\boldsymbol{\gamma} - \boldsymbol{y}\| \geq \|\boldsymbol{\gamma}\| - \|\boldsymbol{y}\| \geq \delta - \|y\| > \delta - d = d$ since $\|\boldsymbol{y}\| < d$. In particular if

$$G\colon \overline{B} \longrightarrow \mathbf{R}$$

---

[1] Recall we have shrunk U so that $\|\boldsymbol{f}'(\boldsymbol{x}) - I\| < \frac{1}{2}$.

is given by $G(\boldsymbol{x}) = \|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{y}\|$ then

$$G(\boldsymbol{x}) > d \qquad (\boldsymbol{x} \in \partial(\overline{B})).$$

On the other hand $G(\mathbf{0}) = \|\boldsymbol{y}\| < d$ and hence G cannot attain its minimum on $\partial(\overline{B})$. Since $\overline{B}$ is compact and G is continuous, G must attain a minimum and we have just argued that that minimum occurs in B and not on the boundary. Consider the continuous map

$$F \colon B \longrightarrow \mathbf{R}$$

given by

$$F(\boldsymbol{x}) = G(\boldsymbol{x})^2 = \|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{y}\|^2 = \sum_{i=1}^{n} (f_i(\boldsymbol{x}) - y_i)^2 \tag{1.1.2}$$

where $\boldsymbol{f} = (f_1, \dots, f_n)$ and $\boldsymbol{y} = (y_1, \dots, y_n)$. Then the minimum of F occurs exactly where the minimum of G occurs. If $\boldsymbol{x}^*$ is this point of minimum, then using one variable calculus is it clear that

$$D_j F(\boldsymbol{x}^*) = 0, \qquad (j = 1, \dots, n). \tag{1.1.3}$$

Using (1.1.2) we see that (1.1.3) is equivalent to

$$2 \sum_{i=1}^{n} (f_i(\boldsymbol{x}^*) - y_i)(D_j f_i)(\boldsymbol{x}^*) = 0 \qquad (j = 1, \dots, n),$$

which in turn can be written as

$$\begin{bmatrix} f_1(\boldsymbol{x}^*) - y_1 & f_2(\boldsymbol{x}^*) - y_2 & \dots & f_n(\boldsymbol{x}^*) - y_n \end{bmatrix} \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\boldsymbol{x}^*) & \dots & \frac{\partial f_1}{\partial x_n}(\boldsymbol{x}^*) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(\boldsymbol{x}^*) & \dots & \frac{\partial f_n}{\partial x_n}(\boldsymbol{x}^*) \end{bmatrix} = 0$$

Since $\boldsymbol{f}'(\boldsymbol{x})$ is invertible on all point of U, and hence on all points of B, in particular at $\boldsymbol{x}^*$, therefore the above gives $f_i(\boldsymbol{x}^*) - y_i = 0$ for $i = 1, \dots, n$, i.e. $\boldsymbol{f}(\boldsymbol{x}^*) = \boldsymbol{y}$. Thus $\boldsymbol{y} \in \boldsymbol{f}(B)$ as claimed.

Let $V = \boldsymbol{f}^{-1}(W)$. Since $\boldsymbol{f}$ is one-to-one, from our claim it is clear that

$$V \subset B.$$

Let $\boldsymbol{\varphi} \colon W \to V$ be the inverse of $\boldsymbol{f}|_V$. Then the inequality (1.1.1) translates to

$$\|\boldsymbol{\varphi}(\boldsymbol{y}_1) - \boldsymbol{\varphi}(\boldsymbol{y}_2)\| \le 2\|\boldsymbol{y}_1 - \boldsymbol{y}_2\| \qquad (\boldsymbol{y}_1, \boldsymbol{y}_2 \in W).$$

It is immediate that $\boldsymbol{\varphi}$ is continuous. Indeed, if $\epsilon > 0$ is given, then $\|\boldsymbol{y}_1 - \boldsymbol{y}_2\| < \epsilon/2$ implies that $\|\boldsymbol{\varphi}(\boldsymbol{y}_1) - \boldsymbol{\varphi}(\boldsymbol{y}_2)\| < \epsilon$, giving continuity of $\boldsymbol{\varphi}$.

Let us prove $\boldsymbol{\varphi}$ is differentiable. To that end, let $\boldsymbol{y} \in W$ and let $\boldsymbol{x} = \boldsymbol{\varphi}(\boldsymbol{y}) \in V$. Let $A = \boldsymbol{f}'(\boldsymbol{x})$, and for $\boldsymbol{h}$ and $\boldsymbol{k}$ sufficiently small, let $\boldsymbol{\varepsilon(h)}$ and $\boldsymbol{\eta(k)}$ be given by

$$\begin{aligned} \boldsymbol{\varepsilon(h)} &= \boldsymbol{f(x+h)} - \boldsymbol{f(x)} - A\boldsymbol{h} \\ \boldsymbol{\eta(k)} &= \boldsymbol{\varphi(y+k)} - \boldsymbol{\varphi(k)} - A^{-1}\boldsymbol{k}. \end{aligned} \tag{1.1.4}$$

We know $\|\boldsymbol{\varepsilon(h)}\|/\|\boldsymbol{h}\| \to 0$ as $\boldsymbol{h} \to \mathbf{0}$. We wish to prove that $\|\boldsymbol{\eta(k)}\|/\|\boldsymbol{k}\| \to 0$ as $\boldsymbol{k} \to \mathbf{0}$. This will prove $\boldsymbol{\varphi}'(\boldsymbol{y})$ exists and is equal to A^{-1}.

For $\|\boldsymbol{k}\|$ sufficiently small, let $\boldsymbol{h} = \boldsymbol{\varphi}(\boldsymbol{y}+\boldsymbol{k}) - \boldsymbol{\varphi}(\boldsymbol{y})$. Then it is clear that $\boldsymbol{k} = \boldsymbol{f}(\boldsymbol{x}+\boldsymbol{h}) - \boldsymbol{f}(\boldsymbol{x})$. Moreover, $\boldsymbol{h} \to \mathbf{0}$ if $\boldsymbol{k} \to \mathbf{0}$, and $\boldsymbol{h} \neq 0$ if and only if $\boldsymbol{k} \neq \mathbf{0}$. Finally, (1.1.1) translates to

$$\frac{\|\boldsymbol{h}\|}{\|\boldsymbol{k}\|} \leq 2 \qquad (\boldsymbol{k} \neq \mathbf{0}). \tag{1.1.5}$$

For non-zero $\boldsymbol{k}$ of sufficiently small magnitude we have

$$\begin{aligned}
\frac{\|\boldsymbol{\eta}(\boldsymbol{k})\|}{\|\boldsymbol{k}\|} &= \frac{1}{\|\boldsymbol{k}\|}\|\boldsymbol{\varphi}(\boldsymbol{y}+\boldsymbol{k}) - \boldsymbol{\varphi}(\boldsymbol{y}) - A^{-1}\boldsymbol{k}\| \\
&= \frac{1}{\|\boldsymbol{k}\|}\|\boldsymbol{h} - A^{-1}\boldsymbol{k}\| \\
&= \frac{1}{\|\boldsymbol{k}\|}\|A^{-1}(A\boldsymbol{h} - \boldsymbol{k})\| \\
&\leq \frac{\|\boldsymbol{h}\|}{\|\boldsymbol{k}\|}\,\frac{\|A^{-1}\|\|A\boldsymbol{h} - \boldsymbol{k}\|}{\|\boldsymbol{h}\|} \\
&\leq 2\frac{\|A^{-1}\|\|A\boldsymbol{h} - \boldsymbol{k}\|}{\|\boldsymbol{h}\|} \qquad \text{(via (1.1.5))} \\
&= \frac{2\|A^{-1}\|\|\boldsymbol{\varepsilon}(\boldsymbol{h})\|}{\|\boldsymbol{h}\|} \qquad (\text{since } \boldsymbol{k} = \boldsymbol{f}(\boldsymbol{x}+\boldsymbol{h}) - \boldsymbol{f}(\boldsymbol{x})).
\end{aligned}$$

Note that we have used the fact that $\boldsymbol{h} \neq \mathbf{0}$ whenever $\boldsymbol{k} \neq \mathbf{0}$. The above chain of inequalities, and the fact that $\|\boldsymbol{\varepsilon}(\boldsymbol{h})\|/\|\boldsymbol{h}\| \to 0$ as $\boldsymbol{h} \to \mathbf{0}$ yields

$$\lim_{\boldsymbol{k}\to\mathbf{0}} \frac{\|\boldsymbol{\eta}(\boldsymbol{k})\|}{\|\boldsymbol{k}\|} = 0.$$

This proves that $\boldsymbol{\varphi}$ is differentiable on W and

$$\boldsymbol{\varphi}'(\boldsymbol{y}) = (\boldsymbol{f}'(\boldsymbol{\varphi}(\boldsymbol{y})))^{-1}$$

for $\boldsymbol{y} \in W$. Since $T \mapsto T^{-1}$ is continuous on $GL_n(\mathbf{R})$,[2] we see that $\boldsymbol{\varphi}'$ is continuous. Thus $\boldsymbol{\varphi}\colon W \to V$ is in $\mathscr{C}^1(W)$.

We will draw out the consequences of what we proved here in the next lecture.

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.

REFERENCES

[R] W. Rudin, *Principles of Mathematical Analysis*, (Third Edition), McGraw-Hill, New Delhi, 1976.

[S] M. Spivak, *Calculus on Manifolds (A modern approach to classical theorems of Advanced Calculus)*, Addison-Wesley, Reading, Massachusetts, 1965.

---

[2] See item 10 in §1 of Lecture 7.