Math 600 Lecture 36

Recall: If $f: E \to \mathbb{R}^m$, where $E \subset \mathbb{R}^n$ is open, then f is differentiable at $x \in E$ iff there exist a linear map $L: \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{h \to 0} \frac{\| f(x+h) - f(x) - L(h) \|}{\| h \|} = 0,$$

that is,

$$f(x+h) = f(x) + L(h) + o(\|h\|).$$

If such an L exist, it is called the derivative of f at x and written $Df(x)$.

Definition: If X and Y are vector spaces over $\mathbb{R}$, then we write $\mathcal{L}(X,Y)$ for the space of all linear map from X into Y. It is well known that $\mathcal{L}(X,Y)$ is also a vector space over $\mathbb{R}$.

Definition: The operator norm on $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ is defined by

$$\|L\| = \max \{ \|Lx\| : x \in \mathbb{R}^n, \|x\| = 1 \}.$$

Note that $\|L\|$ is well defined since $\{x \in \mathbb{R}^n \mid \|x\| = 1\}$ is compact and $f: \mathbb{R}^n \to \mathbb{R}$ defined by $f(x) = \|Lx\|$ is continuous. It is straightforward to prove that $\|\cdot\|$ defines a norm on $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$.

Theorem:

1. If $L \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, then

$$\underbrace{\|Lx\|}_{\text{Euclidean norm in } \mathbb{R}^m} \le \underbrace{\|L\|}_{\text{Operator norm on } \mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)} \underbrace{\|x\|}_{\text{Euclidean norm in } \mathbb{R}^n} \quad \forall x \in \mathbb{R}^n$$

2. If $L \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, then

$$\|L\| = \min\{\beta \ge 0 : \|Lx\| \le \beta \|x\| \ \forall x \in \mathbb{R}^n\}.$$

3. If $L \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ and $M \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^n)$, then

$$\|LM\| \le \|L\| \|M\|.$$

Proof: 1. By definition,

$$\|Lx\| \le \|L\| \quad \forall x \in \mathbb{R}^n, \|x\| = 1$$

$$\Rightarrow \left\| L\left(\frac{x}{\|x\|}\right) \right\| \le \|L\| \quad \forall x \in \mathbb{R}^n, x \neq 0$$

$$\Rightarrow \frac{\|Lx\|}{\|x\|} \le \|L\| \quad \forall x \in \mathbb{R}^n, x \neq 0$$

$$\Rightarrow \|Lx\| \le \|L\| \|x\| \quad \forall x \in \mathbb{R}^n$$ (since this inequality obviously holds for $x = 0$).

2. Write $S = \{\beta \ge 0 : \|Lx\| \le \beta \|x\| \ \forall x \in \mathbb{R}^n\}$. Then, by #1, $\|L\| \in S$.

Suppose $\beta \in S$. Then

$$\|Lx\| \le \beta \quad \forall x \in \mathbb{R}^n, \|x\| = 1$$

$$\Rightarrow \sup\{\|Lx\| : x \in \mathbb{R}^n, \|x\| = 1\} \le \beta$$

$$\Rightarrow \|L\| \le \beta.$$

This proves #2.

3. We have

$$\|LMx\| = \|L(Mx)\| \le \|L\|\,\|Mx\| \le \|L\|\,\|M\|\,\|x\| \quad \forall x \in \mathbb{R}^p.$$

By #2, this implies that

$$\|LM\| \le \|L\|\|M\|. //$$

<u>Theorem</u>: Let $\Omega = \{A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) \mid A \text{ is invertible}\}$. Then Ω is open, in fact,

$$A \in \Omega \Rightarrow B_r(A) \subset \Omega \text{ for } r = \frac{1}{\|A^{-1}\|}.$$

<u>Proof:</u> Recall that $B \in \mathcal{L}(\mathbb{R}^n) = \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ belongs to Ω iff B is nonsingular $(Bx = 0 \Rightarrow x = 0)$.

Let $A \in \Omega$ and suppose

$$B \in \mathcal{L}(\mathbb{R}^n), \ \|B - A\| < \frac{1}{\|A^{-1}\|}.$$

Then if $x \in \mathbb{R}^n$ and $x \neq 0$, we have

$$\begin{aligned} Bx = Ax + (B-A)x \Rightarrow \|Bx\| &\ge \|Ax\| - \|(B-A)x\| \\ &\ge \|Ax\| - \|B-A\|\,\|x\| \\ &> \|Ax\| - \frac{\|x\|}{\|A^{-1}\|} \\ &= \|Ax\| - \frac{\|A^{-1}Ax\|}{\|A^{-1}\|} \ge \|Ax\| - \frac{\|A^{-1}\|\,\|Ax\|}{\|A^{-1}\|} = 0. \end{aligned}$$

Thus B is nonsingular and hence $B \in \Omega$. //

<u>Corollary</u>: $f : \Omega \to \Omega$ defined by $f(A) = A^{-1}$ is continuous and invertible, with $f^{-1} = f$.

<u>Proof:</u> Since $(A^{-1})^{-1} = A$, it is obvious that f is invertible and $f^{-1} = f$.

Now suppose $A \in \Omega$ and

$$\|B-A\| \le \frac{1}{2\|A^{-1}\|}.$$

Then $B \in \Omega$ and

$$B^{-1}-A^{-1} = B^{-1}AA^{-1}-B^{-1}BA^{-1} = B^{-1}(A-B)A^{-1}$$

$$\Rightarrow \|B^{-1}-A^{-1}\| \le \|B^{-1}\|\,\|A-B\|\,\|A^{-1}\| \le \frac{\|B^{-1}\|\,\|A^{-1}\|}{2\|A^{-1}\|} = \frac{\|B^{-1}\|}{2}$$

$$\Rightarrow \|B^{-1}\| \le \|A^{-1}\| + \frac{\|B^{-1}\|}{2}$$

$$\Rightarrow \frac{\|B^{-1}\|}{2} \le \|A^{-1}\|$$

$$\Rightarrow \|B^{-1}\| \le 2\|A^{-1}\|.$$

Now let $\varepsilon > 0$ be given and define

$$\delta = \min\left\{\frac{1}{2\|A^{-1}\|}, \frac{\varepsilon}{2\|A^{-1}\|^2}\right\}.$$

Then

$$\begin{aligned}\|B-A\| < \delta \Rightarrow \|B^{-1}-A^{-1}\| &\le \|B^{-1}\|\,\|A-B\|\,\|A^{-1}\| \\ &\le 2\|A^{-1}\|^2\|A-B\| \\ &< 2\|A^{-1}\|^2 \cdot \frac{\varepsilon}{2\|A^{-1}\|^2} = \varepsilon.\end{aligned}$$

Thus f is continuous at A. //

Theorem (the chain rule): Suppose $E \subset \mathbb{R}^n$ and $F \subset \mathbb{R}^m$ are open, $f: E \to \mathbb{R}^m$, with $R(f) \subset F$, and $g: F \to \mathbb{R}^k$. If f is differentiable at $x \in E$ and g is differentiable at $f(x) \in F$, then $h = g \circ f$ is differentiable at x, and

$$Dh(x) = Dg(f(x))Df(x).$$

Proof: We have

$$\begin{aligned} h(x+p) - h(x) &= g(f(x+p)) - g(f(x)) \\ &= g(f(x) + Df(x)p + o(\|p\|)) - g(f(x)) \\ &= \cancel{g(f(x))} + Dg(f(x))(Df(x)p + o(\|p\|)) + o(\|Df(x)p + o(\|p\|)\|) - \cancel{g(f(x))} \\ &= Dg(f(x))Df(x)p + Dg(f(x))\,o(\|p\|) + o(\|Df(x)p + o(\|p\|)\|) \end{aligned}$$

Now,

$$\|Dg(f(x))\,o(\|p\|)\| \leq \|Dg(f(x))\|\,\|o(\|p\|)\| = o(\|p\|) \quad \text{(since } \|Dg(f(x))\| \text{ is a constant in } \mathbb{R}\text{)}$$

and

$$\begin{aligned} \|Df(x)p + o(\|p\|)\| &\leq \|Df(x)\|\,\|p\| + \|o(\|p\|)\| = \left(\|Df(x)\| + \frac{\|o(\|p\|)\|}{\|p\|}\right)\|p\| \\ &\leq (\|Df(x)\| + 1)\|p\| \quad \forall p \text{ suff. small} \end{aligned}$$

Thus

$$\frac{o(\|Df(x)p + o(\|p\|)\|)}{\|Df(x)p + o(\|p\|)\|} \to 0 \text{ as } p \to 0$$

$$\Rightarrow \quad \frac{o\left(\|Df(x)p + o(\|p\|)\|\right)}{(\|Df(x)\|+1)\|p\|} \to 0 \quad \text{as } p \to 0$$

$$\Rightarrow \quad \frac{o\left(\|Df(x)p + o(\|p\|)\|\right)}{\|p\|} \to 0 \quad \text{as } p \to 0$$

$$\Rightarrow \quad o\left(\|Df(x)p + o(\|p\|)\|\right) = o(\|p\|).$$

Thus

$$h(x+p) - h(x) = Dg(f(x))Df(x)p + o(\|p\|),$$

which implies that

$$Dh(x) = Dg(f(x))Df(x),$$

as desired. //

Note that $Dg(f(x))Df(x)$ is the product (composition) of two linear maps, and this product is <u>not</u> commutative.