Matrix norm

$\mathcal{M}_{m,n} := \{ A : \ m\times n\text{-matrix of real entries} \} \simeq \mathbb{R}^{m\cdot n}$

Def For $A \in \mathcal{M}_{m,n}$ def: $\|A\| = \sup_{x \in \mathbb{R}^n \setminus \{0\}} \frac{\|Ax\|}{\|x\|}$ (Euclidean norms)

Lemma For all $A \in \mathcal{M}_{m,n}$: $\max_{i \le m} \max_{j \le n} |A_{ij}| \le \|A\| \le \left(\sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij}^2 \right)^{1/2}$

Moreover,

- $\forall A \in \mathcal{M}_{m,n} \ \forall \lambda \in \mathbb{R}$: $\|\lambda A\| = |\lambda| \, \|A\|$
- $\forall A, B \in \mathcal{M}_{m,n}$: $\|A+B\| \le \|A\| + \|B\|$

$\Big\}$ $\|\cdot\|$ is a norm

- $\forall A \in \mathcal{M}_{m,n} \ \forall B \in \mathcal{M}_{k,m}$: $\|BA\| \le \|B\| \, \|A\|$

Pf. $x \in \mathbb{R}^n$, $i = 1 \dots m$:

$$(Ax)_i = \sum_{j=1}^n A_{ij} x_j$$

$$|(Ax)_i|^2 = \Big(\sum_{j=1}^n A_{ij} x_j\Big)^2 \le \Big(\sum_{j=1}^n A_{ij}^2\Big) \underbrace{\sum_{j=1}^n x_j^2}_{=\|x\|^2}$$

Sum over i to get

$$\|Ax\| \le \Big(\sum_{i=1}^m \sum_{j=1}^n A_{ij}^2\Big)^{1/2} \|x\|$$

$$\Rightarrow \|A\| \le \text{—〃—}$$

LB: $(Ae_j)_i = A_{ij}$

$$\|Ae_j\| \ge |A_{ij}| \underset{\|e_j\|=1}{\Rightarrow} \|A\| \ge |A_{ij}|$$

Optimize over i,j.

For the rest:

$$\|(\lambda A)x\| = \|\lambda(Ax)\| = |\lambda| \|Ax\|$$

$$\|(A+B)x\| = \|Ax + Bx\| \le \|Ax\| + \|Bx\|$$

Note def of $\|A\|$ implies

$$\forall x \in \mathbb{R}^n: \ \|Ax\| \le \|A\| \|x\|$$

if $A \in M_{m,n}$, $B \in M_{k,m}$: then

$$\forall x \in \mathbb{R}^n: \ \|BAx\| \le \|B\| \|Ax\| \le \|B\| \|A\| \|x\|$$

$\boxtimes$

We have $\frac{1}{\sqrt{mn}}\Big(\sum_{i=1}^m \sum_{j=1}^n A_{ij}^2\Big)^{1/2} \le \|A\| \le \Big(\sum_{i=1}^m \sum_{j=1}^n A_{ij}^2\Big)^{1/2}$

Lemma Assume $f: \mathbb{R}$

Then $\exists u_x: \mathbb{R}^n \to M_{m,n}$

$\forall z \in \mathrm{Dom}(f): f(z) -$

AND $\lim_{z \to x} \|u_x(z)\|$

Conversely, if such u_x

then f is differentiable a

Pf $(u_x(z))_{ij} = \big[f_i(z) - f_i(x$

Then $\sum_{j=1}^n u_x(z)_{ij} (z_j - x_j) = \big[\text{—}$

x_j

$x_j\Big)^2 \le \Big(\sum_{j=1}^{n} A_{ij}^2\Big) \underbrace{\sum_{j=1}^{n} x_j^2}_{=\|x\|^2}$

$\Big(\sum_{j=1}^{n} A_{ij}^2\Big)^{1/2} \|x\|$

$\Rightarrow \|A\| \ge |A_{ij}|$

$\|e_j\| = 1$ optimize over i, j.

For the rest:

$\|(\lambda A)x\| = \|\lambda (Ax)\| = |\lambda| \, \|Ax\|$

$\|(A+B)x\| = \|Ax + Bx\| \le \|Ax\| + \|Bx\|$

Note def of $\|A\|$ implies

$\forall x \in \mathbb{R}^n$: $\|Ax\| \le \|A\| \|x\|$

if $A \in \mathcal{M}_{m,n}$, $B \in \mathcal{M}_{k,m}$: then

$\forall x \in \mathbb{R}^n$: $\|BAx\| \le \|B\| \|Ax\| \le \|B\| \|A\| \|x\|$

We have

$$\frac{1}{\sqrt{mn}} \Big(\sum_{i=1}^{m}\sum_{j=1}^{n} A_{ij}^2\Big)^{1/2} \le \|A\| \le \Big(\sum_{i=1}^{m}\sum_{j=1}^{n} A_{ij}^2\Big)^{1/2}$$

$\boxtimes$

Lemma Assume $f : \mathbb{R}^n \to \mathbb{R}^m$ is diff. at $x \in \operatorname{int} \operatorname{Dom}(f)$.

Then $\exists u_x : \mathbb{R}^n \to \mathcal{M}_{m,n}$, $\operatorname{Dom}(u_x) = \operatorname{Dom}(f)$ s.t.

$\forall z \in \operatorname{Dom}(f)$: $f(z) - f(x) = \big[Df(x) + u_x(z)\big](z - x)$

AND $\lim_{z \to x} \|u_x(z)\| = 0$

Conversely, if such $u_x(z)$, $Df(x) \in \mathcal{M}_{m,n}$ exist then f is differentiable at x.

Pf $(u_x(z))_{ij} = \Big[f_i(z) - f_i(x) - \sum_{k=1}^{n} \frac{\partial f_i}{\partial x_k}(x)(z_k - x_k)\Big] \dfrac{z_j - x_j}{\|z - x\|^2}$

Then $\sum_{j=1}^{n} u_x(z)_{ij}(z_j - x_j) = \big[\text{—— " ——}\big] \underbrace{\sum_{j=1}^{n} \frac{(z_j - x_j)^2}{\|z - x\|^2}}_{=1}$

So $u_x(z)(z-x) = f(z) - f(x) - Df(x)(z-x)$

Moreover,

$$\|u_x(z)\|^2 \le \sum_{i=1}^{m} \sum_{j=1}^{n} u_x(z)_{ij}^2 = \sum_{i=1}^{m} \left[\ldots \right]^2 \underbrace{\sum_{j=1}^{n} \frac{(z_j - x_j)^2}{\|z-x\|^4}}_{= \frac{1}{\|z-x\|^2}}$$

$$= \frac{\|f(z) - f(x) - Df(x)(z-x)\|^2}{\|z-x\|^2} = \frac{\|u_x(z)(z-x)\|^2}{\|z-x\|^2} \le \|u_x(z)\|^2$$

So $\|u_x(z)\| \to 0 \iff Df(x)$ exists.

$\boxtimes$

Rules of multivariate differentiation

Lemma Let $f, g : \mathbb{R}^n \to \mathbb{R}^m$ be diff. at x. Then so is $f+g$ and $D(f+g)(x) = Df(x) + Dg(x)$

Lemma Let $f : \mathbb{R}^n \to \mathbb{R}^m$, $g : \mathbb{R}^n \to \mathbb{R}^k$, $A \in \mathcal{M}_{m,k}$.
If f and g are diff. at x, then so is $z \mapsto f(z) \cdot A\, g(z)$

and

$$D(f \cdot Ag)(x) = \underbrace{g(x) \cdot A^T Df(x) + f(x) \cdot A\, Dg(x)}_{\text{RHS}}$$

Pf

$$f\cdot Ag(z) - f\cdot Ag(x) - (RHS)(z-x)$$

$$= f(z)\cdot A\left[\, g(z) - g(x) - Dg(x)(z-x) \,\right]$$

$$+\ g(x)\cdot A^T\left[\, f(z) - f(x) - Df(x)(z-x) \,\right]$$

Apply matrix norms and use the fact that f is continuous at x.

So $\exists M>0\ \exists \delta>0\ \forall z\in\mathbb{R}^n: \|z-x\|<\delta \Rightarrow \|f(z)\|\leq M$

Apply def of $Df(x)$, $Dg(x)$. $\boxtimes$

In coordinates: $h = f\cdot Ag$

$$h(z) := \sum_{i=1}^{m}\sum_{j=1}^{k} f_i(z)\, A_{ij}\, g_j(z)$$

$$\frac{\partial f}{\partial x_r} = \sum_{i=1}^{m}\sum_{j=1}^{k}\left[\frac{\partial f_i}{\partial x_r} A_{ij}\, g_j + f_i A_{ij}\frac{\partial g_j}{\partial x_r}\right]$$

Lemma (Chain rule)

Let $f:\mathbb{R}^n\to\mathbb{R}^m$, $g:\mathbb{R}^m\to\mathbb{R}^k$ be s.t.

- $x\in \operatorname{int}\operatorname{Dom}(f) \wedge f(x)\in\operatorname{int}\operatorname{Dom}(g)$
- f diff at x $\wedge$ g diff. at $f(x)$

Then $g\circ f$ is diff at x and

$$D(g\circ f)(x) = Dg(f(x))Df(x)$$

Pf $g\circ f(z) - g\circ f(x)$

$= [Dg(f(x)) + u$

$= [\ —\ "\ —$

$= [Dg(f(x))Df($

where

$\hat{u}_x(z) = [Dg(f(x))$

$\|u_x(z)\| \leq (\|Dg(f(x))\|$

Now use that $f(z)\to$

$\ldots) - (RHS)(z-x)$

$\ldots g(x) - Dg(x)(z-x)]$

$[f(z) - f(x) - Df(x)(z-x)]$

... and use the

...continuous at x

$\ldots \forall z \in \mathbb{R}^n : \|z-x\| < \delta \Rightarrow \|f(z)\| \leq M$

$\ldots), Dg(x).$ $\boxtimes$

In coordinates: $h = f \cdot Ag$

$$h(z) := \sum_{i=1}^{m} \sum_{j=1}^{k} f_i(z) A_{ij} g_j(z)$$

$$\frac{\partial f}{\partial x_r} = \sum_{i=1}^{m} \sum_{j=1}^{k} \left[\frac{\partial f_i}{\partial x_r} A_{ij} g_j + f_i A_{ij} \frac{\partial g_j}{\partial x_r} \right]$$

Lemma (Chain rule)

Let $f : \mathbb{R}^n \to \mathbb{R}^m$, $g : \mathbb{R}^m \to \mathbb{R}^k$ be s.t.

- $x \in \operatorname{int} \operatorname{Dom}(f) \wedge f(x) \in \operatorname{int} \operatorname{Dom}(g)$
- f diff at x $\wedge$ g diff. at $f(x)$

Then $g \circ f$ is diff at x and

$$D(g \circ f)(x) = Dg(f(x)) Df(x)$$

Pf

$$\begin{aligned} g \circ f(z) - g \circ f(x) &= [Dg(f(x)) + u_{f(x)}(f(z))][f(z) - f(x)] \\ &= [\text{ --- } " \text{ --- }][Df(x) + \tilde{u}_x(z)](z-x) \\ &= [Dg(f(x)) Df(x) + \hat{u}_x(z)](z-x) \end{aligned}$$

where

$$\hat{u}_x(z) = [Dg(f(x)) + u_{f(x)}(f(z))] \tilde{u}_x(z) + u_{f(x)}(f(z)) Df(x)$$

$$\|\hat{u}_x(z)\| \leq \left(\|Dg(f(x))\| + \|u_{f(x)}(f(z))\| \right) \|\tilde{u}_x(z)\| + \|u_{f(x)}(f(z))\| \|Df(x)\|$$

Now use that $f(z) \to f(x)$ as $z \to x$. $\boxtimes$

Inverse function rule needs:

<u>Lemma</u> $\forall\ A, B \in M_{n,n}$, A invertible:

$$\|B\|\,\|A^{-1}\| < 1 \Rightarrow A+B \text{ invertible}$$

$$\wedge\ \|(A+B)^{-1}\| \leq \frac{\|A^{-1}\|}{1 - \|B\|\,\|A^{-1}\|}$$

Moreover, we also get

$$\|(A+B)^{-1} - A^{-1}\| \leq \frac{\|A^{-1}\|^2}{1 - \|B\|\,\|A^{-1}\|}\,\|B\|$$

Pf:

$$
\begin{aligned}
\|(I+A^{-1}B)x\| &= \|x + A^{-1}Bx\| \\
&\geq \|x\| - \|A^{-1}Bx\| \\
&\geq \|x\| - \|A^{-1}\| \, \|B\| \, \|x\| \\
&= \|x\| \left(1 - \|A^{-1}\| \, \|B\|\right)
\end{aligned}
\Bigg\}
\quad
\begin{aligned}
&\|(I+A^{-1}B)^{-1}z\| \leq \frac{\|z\|}{1 - \|A^{-1}\| \, \|B\|} \\
&\Rightarrow \|(I+A^{-1}B)^{-1}\| \leq \frac{1}{1 - \|A^{-1}\| \, \|B\|}
\end{aligned}
$$

So $\|A^{-1}\| \, \|B\| < 1 \Rightarrow (I+A^{-1}B)$ invertible.

Then $(A+B) = A(I+A^{-1}B) \Rightarrow (A+B)^{-1} = (I+A^{-1}B)^{-1}A^{-1}$

$$\Rightarrow \|(A+B)^{-1}\| \leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \, \|B\|}$$

$$(A+B)^{-1} - A^{-1} = (A+B)^{-1}\big(\underbrace{A - (A+B)}_{-B}\big)A^{-1} = -(A+B)^{-1}BA^{-1} \qquad \boxtimes$$