Real Analysis Notes

Stan Tuznik

November 1, 2020

Contents

Chapter 1

Sequences

We know the usual definition of sequences in $\mathbb{R}$. A sequence of vectors converges similarly, except that the norm used in the definition of convergence is now a vector norm, not simply absolute value. A sequence of vectors $\{\vec{a}_i\}$ in $\mathbb{R}^n$ converges to $\vec{a} \in \mathbb{R}^n$ if for any $\epsilon > 0$ there exists a $n \in \mathbb{N}$ such that

$$n > N \implies \|\vec{a}_i - \vec{a}\| < \epsilon$$

This definition can be related to the convergence of the vector components.

Prop 1.1. *Let $\{\vec{a}_i\}$ be a sequence of vectors (coordinate vectors) in $\mathbb{R}^n$. Then $\{\vec{a}_i\}$ converges to $\vec{a} \in \mathbb{R}^n$ if and only if each component $(\vec{a}_i)_j$ converges to $(\vec{a})_j$.*

Proof. Assume that we have the components converging. That is, for $\epsilon > 0$ we have some $N \in \mathbb{N}$ such that

$$n > N \implies \left|(\vec{a}_i)_j - (\vec{a})_j\right| < \epsilon/\sqrt{n}$$

We choose N to hold for every component i. This can be arranged by taking N to the the maximum of the N guaranteed for each individual component. Thus, we have, for $n > N$,

$$\begin{aligned}\|\vec{a}_i - \vec{a}\|^2 &= \sum_{j=1}^{n} \left|(\vec{a}_i)_j - (\vec{a})_j\right|^2 \\ &< \sum_{j=1}^{n} \epsilon^2/n \\ &= \epsilon\end{aligned}$$

so that $\{\vec{a}_i\}$ converges to $\vec{a}$.

Conversely, assume that $\{\vec{a}_i\}$ converges to $\vec{a}$. Let $\epsilon > 0$ be arbitrary and let N be the guaranteed natural number in the definition of vector sequence convergence.

The key is to note that for any $j = 1, 2, \ldots, n$ we have

$$\left|(\vec{a}_i)_j - (\vec{a})_j\right|^2 \leq \sum_{k=1}^{n} \left|(\vec{a}_i)_k - (\vec{a})_k\right|^2 < \epsilon^2$$

whenever $n > N$. Thus, we have in particular

$$n > N \implies \left|(\vec{a}_i)_j - (\vec{a})_j\right| < \epsilon$$

and so the component sequences converge as expected. ∎

Let's look at some interesting consequences.

Prop 1.2. *Let $\vec{x}_k \to \vec{a}$. Then $\|\vec{x}_k\| \to \|\vec{a}\|$ and $\vec{b} \cdot \vec{x}_k \to \vec{b} \cdot \vec{a}$.*

Proof. The first can be proven by writing $\vec{x} = [x_1, x_2, \ldots, x_n]$ and proceeding as usual with the usual 2-norm. More generally, though, we can use the (reverse) triangle inequality property of the norm to see that

$$|\|\vec{x}_k\| - \|\vec{a}\|| \leq \|\vec{x}_k - \vec{a}\|$$

which can be made smaller than arbitrary $\epsilon > 0$ by proper choice of N. Thus, the sequence of norms converges as expected.

Next, we could let $f(\vec{x}) = \vec{b}^T \vec{x}$ and use continuity of linear functions to prove the conclusion. Instead, just notice that, from a useful proposition in the next chapter, we have

$$\left\|\vec{b}^T \vec{x}\right\| \leq \left\|\vec{b}\right\| \|\vec{x}\|$$

so that we have

$$\left\|\vec{b}^T \vec{x}_k - \vec{b}^T \vec{a}\right\| = \left\|\vec{b}^T (\vec{x}_k - \vec{a})\right\| = \left\|\vec{b}\right\| \|\vec{x}_k - \vec{a}\|$$

Since $\vec{x}_k \to \vec{a}$, given arbitrary $\epsilon > 0$, there is some N such that $k > N \implies \|\vec{x}_k - \vec{a}\| < \epsilon / \left\|\vec{b}\right\|$. Thus, we have proven the statement.

∎

1.1 Subsequences

A subsequence of a sequence $\{\vec{x}_k\}_k$ is a subset $\{\vec{x}_{k_i}\}_{k_i}$ where $k_1 < k_2 < k_3 < \ldots$ is an increasing sequence of natural numbers.

Prop 1.3. *If a sequence converges, then any subsequence converges. If every subsequence of a sequence converges, then the sequence converges.*

Proof. The converse is trivial, since the sequence itself is a subsequence. Note that it is not the case that a single convergent subsequence implies convergence of the sequence.

Let $\{\vec{x}_{k_i}\}_{k_i}$ be a subsequence of convergent sequence $\{\vec{x}_k\}_k$ which has limit $\vec{a}$.

Then for arbitrary $\epsilon > 0$, there is an N such that

$$k > N \implies \|\vec{x}_k - \vec{a}\| < \epsilon$$

Since $\{\vec{x}_{k_i}\}_{k_i}$ is a subsequence, we then have

$$i > N \implies \|\vec{x}_{k_i} - \vec{a}\| < \epsilon$$

■

Chapter 2

Continuity

Prop 2.1. *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map represented by the matrix $A \in \mathbb{R}^{m\times n}$. Then we have*

$$\|T(h)\| = \|Ah\| \leq M \|h\|$$

for some constant M.

Proof. Note that this is just a statement about matrix-vector multiplication. Let $[Ah]_i$ denote the ith entry of the matrix product $Ah \in \mathbb{R}^m$. That is,

$$[Ah]_i = \sum_{j=1}^{n} A_{ij} h_j$$

Then let

$$r = \max_{i,j} |A_{ij}|$$

Also note that since

$$\|h\|^2 = \sum_{j=1}^{n} |h_j|^2$$

we have

$$|h_j| \leq \|h\|$$

Taken together, we have

$$[Ah]_i = \sum_{j=1}^{n} A_{ij} h_j \leq \sum_{j=1}^{n} |A_{ij}| |h_j| \leq rn \|h\|$$

so that

$$\|Ah\|^2 = \sum_{i=1}^{m} [Ah]_i^2 \leq \sum_{i=1}^{m} r^2 n^2 \|h\|^2 = mr^2 n^2 \|h\|^2$$

and so we have

$$\|Ah\| \leq rn\sqrt{m} \|h\|$$

Letting $M = rn\sqrt{m}$, the conclusion holds. ∎

Lemma 2.2. *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map represented by the matrix $A \in \mathbb{R}^{m\times n}$. Then for $a \in \mathbb{R}^n$, T is continuous at a.*

Proof. We must show that

$$\lim_{h\to 0} \|T(a+h) - T(a)\| = 0$$

Note that, by linearity,

$$\|T(a+h) - T(a)\| = \|T(a) + T(h) - T(a)\| = \|T(h)\| \le M\|h\|$$

and the right-hand side clearly goes to 0 as $h \to 0$. ∎

The next seems to be an alternative approach. For arbitrary linear map $A \in \mathbb{R}^{m\times n}$, we can define a matrix norm

$$\|A\| := \sum_{i=1,\dots,m} \sum_{j=1,\dots,n} A_{i,j}^2$$

This is technically the square of the element-wise matrix p-norm:

$$\|A\|_p = \left(\sum_{\text{rows}} \sum_{\text{cols}} A_{i,j}^p \right)^{1/p}$$

Then we can look at bounds on matrix-vector and matrix-matrix products. In fact, the following proposition is a generalization of the one above.

Prop 2.3. *Let $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times p}$, and $\vec{b} \in \mathbb{R}^n$. Then*

$$\|A\vec{b}\| \le \|A\| \|\vec{b}\|$$

$$\|AB\| \le \|A\| \|B\|$$

Proof. For the first inequality, notice that the magnitude on the left is a vector norm. That is,

$$\|A\vec{b}\|^2 = \sum_{i=1}^{m} \left(A\vec{b}\right)_i^2$$

and that each entry in this vector is the inner product (usual Euclidean dot product) of $\vec{b}$ with a row of A:

$$\left(A\vec{b}\right)_i = A_{i*}\vec{b}$$

Since this is an inner product, we may apply the Cauchy-Schwarz inequality to it and obtain

$$\left(A\vec{b}\right)_i = A_{i*}\vec{b} \le \left|A_{i*}\vec{b}\right| \le \|A_{i*}\| \|\vec{b}\|$$

Taking this in the original vector magnitude, we have

$$\begin{aligned}\left\|A\vec{b}\right\|^2 &= \sum_{i=1}^{m}\left(A\vec{b}\right)_i^2\\ &\le \sum_{i=1}^{m}\|A_{i*}\|^2\left\|\vec{b}\right\|^2\\ &= \left(\sum_{i=1}^{m}\|A_{i*}\|^2\right)\left\|\vec{b}\right\|^2\\ &= \|A\|^2\left\|\vec{b}\right\|^2\end{aligned}$$

and so $\left\|A\vec{b}\right\| \le \|A\|\left\|\vec{b}\right\|$ as we wanted to prove.

The next inequality is even more general and follows from the matrix-vector product case. Note that every column of the matrix product AB is of the form $A\vec{b}$ where $\vec{b} = B_{*j}$. Thus

$$\begin{aligned}\|AB\|^2 &= \sum_{i=1}^{m}\sum_{j=1}^{n}(AB)_{i,j}^2\\ &= \sum_{j=1}^{n}|AB_{*j}|^2\\ &\le \sum_{j=1}^{n}\|A\|^2\|B_{*j}\|^2\\ &= \|A\|^2\left(\sum_{j=1}^{n}\|B_{*j}\|^2\right)\\ &= \|A\|^2\|B\|^2\end{aligned}$$

and so

$$\|AB\| \le \|A\|\,\|B\|$$

as desired.

■

The second inequality defines the matrix norm $\|\cdot\|$ (and more generally the matrix 2-norm) as a *submultiplicative* norm. With these general properties in hand, we can easily verify continuity of a linear function.

Proof. The linear function $T : \mathbb{R}^n \to \mathbb{R}^m$ will be continuous if we have

$$\lim_{x\to a} T(x) = T(a)$$

Let $\epsilon > 0$ be arbitrary. Then

$$|\vec{x} - \vec{a}| < \epsilon/\|A\|$$

implies that

$$\begin{aligned}\left\|T\left(\vec{x}\right)-T\left(\vec{a}\right)\right\| &= \left\|A\vec{x}-A\vec{a}\right\| \\ &= \left\|A\left(\vec{x}-\vec{a}\right)\right\| \\ &\leq \left\|A\right\|\left\|\vec{x}-\vec{a}\right\| \\ &< \epsilon\end{aligned}$$

Note that linearity of T allowed us to express its action as a matrix multiplication, and the associative (linearity) property of matrix multiplication allowed us to "pull it off" of the $\vec{x}$ and $\vec{a}$ and onto their vector difference. Then, the previous proposition allowed us to bound this matrix-vector product by the norm of the vector and a quantity which measures in some sense the size of the matrix A. ■

Chapter 3

The Derivative

Let
$$f : \mathbb{R}^n \to \mathbb{R}^m$$
be an arbitrary function. Then the derivative of f at $a \in \mathbb{R}^n$ is the unique linear map
$$\lambda : \mathbb{R}^n \to \mathbb{R}^m$$
such that
$$\lim_{h \to 0} \frac{\| f(a+h) - f(a) - \lambda(h) \|}{\| h \|} = 0$$
The norms are required since all quantities are vectors. Note that since λ is a linear map, it has a matrix representation, written $Df(a)$ and called the derivative of f at a. The action of λ on h is
$$\lambda(h) = Df(a) h$$
i.e., the usual matrix multiplication.

Prop 3.1. *If $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $a \in \mathbb{R}^n$, then this derivative is unique.*

Proof. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be differentiable at $a \in \mathbb{R}^n$ and assume there are two derivatives λ and μ. Then
$$\| \lambda(h) - \mu(h) \| \leq \| \lambda(h) - f(a+h) + f(a) \| + \| \mu(h) - f(a+h) + f(a) \|$$
and upon dividing by $\|h\|$ and taking the limit $h \to 0$, we obtain
$$\lim_{h \to 0} \frac{\| \lambda(h) - \mu(h) \|}{\| h \|} = 0$$
Choosing arbitrary $x \in \mathbb{R}^n \setminus \{0\}$, note that $tx \to 0$ when $t \to 0$ ($t \in \mathbb{R}$). Then
$$0 = \lim_{t \to 0} \frac{\| \lambda(tx) - \mu(tx) \|}{\| tx \|} = \lim_{t \to 0} \frac{|t| \, \| \lambda(x) - \mu(x) \|}{|t| \, \| x \|} = \frac{\| \lambda(x) - \mu(x) \|}{\| x \|}$$

and so we have $\lambda(x) = \mu(x)$ for arbitrary $x \in \mathbb{R}^n$. Thus, the derivative is unique. ∎

Prop 3.2. *If $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $a \in \mathbb{R}^n$, then it is continuous at a.*

Proof.

$$\|f(a+h) - f(a)\| \leq \|f(a+h) - f(a) - \lambda(h)\| + \|\lambda(h)\|$$

so that

$$\|f(a+h) - f(a)\| \leq \|h\| \frac{\|f(a+h) - f(a) - \lambda(h)\|}{\|h\|} + \|\lambda(h)\|$$

By taking the limit of both sides as $h \to 0$, the first term on the right-hand side is 0, and the second term, containing the linear λ, goes to 0 as well. Specifically, from a property of linear maps proved in the previous chapter, we have

$$\|\lambda(h)\| \leq M \|h\|$$

for some constant M which certainly tends toward 0. Specifically, we have shown that

$$\|\lambda(h)\| \leq \|Df(a)\| \|h\|$$

where $\|Df(a)\|$ is the squared entry-wise matrix norm of the derivative (matrix) of f at a, which certainly exists (is finite). Thus, we have

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

establishing continuity of f at a. ∎