1. Limits and Continuity

It is often the case that a non-linear function of n-variables $x = (x_1, \ldots, x_n)$ is not really defined on all of $\mathbf{R}^n$. For instance $f(x_1, x_2) = \frac{x_1 x_2}{x_1^2 - x_2^2}$ is not defined when $x_1 = \pm x_2$. However, I will adopt a convention from the vector calculus notes of Jones and write $F : \mathbf{R}^n \to \mathbf{R}^m$ regardless, meaning only that the source of F is some subset of $\mathbf{R}^n$. While a bit imprecise, this will not cause any big problems and will simplify many statements.

I will often distinguish between functions $f : \mathbf{R}^n \to \mathbf{R}$ that are *scalar-valued* and functions $F : \mathbf{R}^n \to \mathbf{R}^m$, $m \geq 2$ that are *vector-valued*, using lower-case letters to denote the former and upper case letters to denote the latter. Note that any vector-valued function $F : \mathbf{R}^n \to \mathbf{R}^m$ may be written $F = (F_1, \ldots, F_m)$ where $F_j : \mathbf{R}^n \to \mathbf{R}$ are scalar-valued functions called the *components* of F. For example, $F : \mathbf{R}^2 \to \mathbf{R}^2$ given by $F(x_1, x_2) = (x_1 x_2, x_1 + x_2)$ is a vector valued function with components $F_1(x_1, x_2) = x_1 x_2$ and $F_2(x_1, x_2) = x_1 + x_2$.

Definition 1.1. *Let $\mathbf{a} \in \mathbf{R}^n$ be a point and $r > 0$ be a positive real number. The* open ball *of radius r about $\mathbf{a}$ is the set*

$$B(\mathbf{a}, r) := \{\mathbf{x} \in \mathbf{R}^n : \|\mathbf{x} - \mathbf{a}\| < r\}.$$

I will also use $B^*(\mathbf{a}, r)$ to denote the set of all $\mathbf{x} \in B(\mathbf{a}, r)$ except $\mathbf{x} = \mathbf{a}$.

Proposition 1.2. *Let $\mathbf{a}, \mathbf{b} \in \mathbf{R}^n$ be points and $r, s > 0$ be real numbers. Then*

- $B(\mathbf{a}, r) \subset B(\mathbf{b}, s)$ *if and only if* $\|\mathbf{a} - \mathbf{b}\| \leq s - r$.
- $B(\mathbf{a}, r) \cap B(\mathbf{b}, s) = \emptyset$ *if and only if* $\|\mathbf{a} - \mathbf{b}\| \geq s + r$.

Proof. Exercise. Both parts depend on the triangle inequality. □

Extending my above convention, I will say that a function $F : \mathbf{R}^n \to \mathbf{R}^m$ is defined *near* a point $\mathbf{a} \in \mathbf{R}^n$ if there exists $r > 0$ such that $F(\mathbf{x})$ is defined for all points $\mathbf{x} \in B(\mathbf{a}, r)$, except possibly the center $\mathbf{x} = \mathbf{a}$.

Below I will need the following inequality relating the magnitude of a vector to the sizes of its coordinates.

Proposition 1.3. *For any vector* $\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \in \mathbf{R}^n$ *we have for each $1 \leq j \leq n$ that*

$$|v_j| \leq \|\mathbf{v}\| \leq \sqrt{n} \max_{1 \leq j \leq n} |v_j|.$$

Proof. Suppose that v_J is the coordinate of $\mathbf{v}$ with largest absolute value. Then for any index j, we have

$$v_j^2 \leq \sum_{i=1}^{n} v_i^2 \leq n v_J^2.$$

Taking square roots throughout gives the inequalities in the statement of the lemma. □

Now we come to the main point. The idea of a 'limit' is one of the most important in all of mathematics. In differential calculus, it is the key to relating non-linear (i.e. hard) functions to linear (i.e. easier) functions.

Definition 1.4. *Suppose that $F : \mathbf{R}^n \to \mathbf{R}^m$ is a function defined near a point $\mathbf{a} \in \mathbf{R}^n$. We say that $F(\mathbf{x})$ has* limit $\mathbf{b} \in \mathbf{R}^m$ *as $\mathbf{x}$ approaches $\mathbf{a}$, i.e.*

$$\lim_{\mathbf{x}\to\mathbf{a}} F(\mathbf{x}) = \mathbf{b} \in \mathbf{R}^m,$$

if for each $\epsilon > 0$ there exists $\delta > 0$ such that $0 < \|\mathbf{x} - \mathbf{a}\| < \delta$ implies $\|F(\mathbf{x}) - \mathbf{b}\| < \epsilon$.

Notice that the final phrase in this definition can be written in terms of balls instead of magnitudes: *for any $\epsilon > 0$ there exists $\delta > 0$ such that $\mathbf{x} \in B^*(\mathbf{a}, \delta)$ implies $F(\mathbf{x}) \in B(\mathbf{b}, \epsilon)$.*

A function might or might not have a limit as $\mathbf{x}$ approaches some given point $\mathbf{a}$, but it never has more than one.

Proposition 1.5 (uniqueness of limits)**.** *If $F : \mathbf{R}^n \to \mathbf{R}^m$ is defined near $\mathbf{a} \in \mathbf{R}^n$, then there is at most one point $\mathbf{b} \in \mathbf{R}^m$ such that $\lim_{\mathbf{x}\to\mathbf{a}} F(\mathbf{x}) = \mathbf{b}$.*

Proof. Suppose, in order to reach a contradiction, that $F(\mathbf{x})$ converges to two different points $\mathbf{b}, \tilde{\mathbf{b}} \in \mathbf{R}^m$ as $\mathbf{x}$ approaches $\mathbf{a}$. Then the quantity $\epsilon := \frac{1}{2}\left\|\tilde{\mathbf{b}} - \mathbf{b}\right\|$ is positive. So by the definition of limit, there exists a number $\delta_1 > 0$ such that $0 < \|\mathbf{x} - \mathbf{a}\| < \delta_1$ implies $\|F(\mathbf{x}) - \mathbf{b}\| < \epsilon$. Likewise, there exists a number $\delta_2 > 0$ such that $0 < \|\mathbf{x} - \mathbf{a}\| < \delta_2$ implies $\left\|F(\mathbf{x}) - \tilde{\mathbf{b}}\right\| < \epsilon$. So if I let $\delta = \min\{\delta_1, \delta_2\}$ be the smaller bound, then $0 < \|\mathbf{x} - \mathbf{a}\| < \delta$ implies that

$$\left\|\tilde{\mathbf{b}} - \mathbf{b}\right\| = \left\|(F(\mathbf{x}) - \mathbf{b}) - (F(\mathbf{x}) - \tilde{\mathbf{b}})\right\| \le \|F(\mathbf{x}) - \mathbf{b}\| + \left\|F(\mathbf{x}) - \tilde{\mathbf{b}}\right\| < \epsilon + \epsilon \le \left\|\tilde{\mathbf{b}} - \mathbf{b}\right\|.$$

Note that the '$\le$' in this estimate follows from the triangle inequality, and the '$<$' follows from my choice of ϵ. At any rate, no real number is smaller than itself, so I have reached a contradiction and conclude that F cannot have two different limits at $\mathbf{a}$. $\square$

Definition 1.6. *We say that a function $F : \mathbf{R}^n \to \mathbf{R}^m$ is* continuous *at $\mathbf{a} \in \mathbf{R}^n$ if F is defined near and at $\mathbf{a}$ and*

$$\lim_{\mathbf{x}\to\mathbf{a}} F(\mathbf{x}) = F(\mathbf{a}).$$

If F is continuous at all points in its domain, we say simply that F is continuous.

Now let us verify that many familiar scalar-valued functions are continuous.

Proposition 1.7. *The following are continuous functions.*

(a) *The constant function $f : \mathbf{R}^n \to \mathbf{R}$, given by $f(\mathbf{x}) = c$ for some fixed $c \in \mathbf{R}$ and all $\mathbf{x} \in \mathbf{R}^n$.*
(b) *The magnitude function $f : \mathbf{R}^n \to \mathbf{R}$ given by $f(\mathbf{x}) = \|\mathbf{x}\|$.*
(c) *The addition function $f : \mathbf{R}^2 \to \mathbf{R}$ given by $f(x_1, x_2) = x_1 + x_2$.*
(d) *The multiplication function $f : \mathbf{R}^2 \to \mathbf{R}$ given by $f(x_1, x_2) = x_1 x_2$.*
(e) *The reciprocal function $f : \mathbf{R} \to \mathbf{R}$ given by $f(x) = 1/x$.*

Proof. (a) Fix a point $\mathbf{a} \in \mathbf{R}^n$. Given $\epsilon > 0$, let $\delta > 0$ be any positive number, e.g. $\delta = 1$ (it won't matter). Then if $\|\mathbf{x} - \mathbf{a}\| < \delta$ it follows that

$$|f(\mathbf{x}) - f(\mathbf{a})| = |c - c| = 0 < \epsilon.$$

So the constant function $f(\mathbf{x}) = c$ is continuous at any point $\mathbf{a} \in \mathbf{R}^n$.

(b) Fix a point $\mathbf{a} \in \mathbf{R}^n$. Given $\epsilon > 0$, let $\delta = \epsilon$. Then if $\|\mathbf{x} - \mathbf{a}\| < \delta$, it follows that

$$|f(\mathbf{x}) - f(\mathbf{a})| = |\,\|\mathbf{x}\| - \|\mathbf{a}\|\,| \leq \|\mathbf{x} - \mathbf{a}\| < \delta = \epsilon.$$

The '$<$' here follows from Problem 1.2.17 (on Homework 1) in Shifrin. Hence $f(\mathbf{x}) = \|\mathbf{x}\|$ is continuous at any point $\mathbf{a} \in \mathbf{R}^n$.

(c) Fix a point $\mathbf{a} = (a_1, a_2) \in \mathbf{R}^2$. Given $\epsilon > 0$, let $\delta = \epsilon/2$. Then if $\|\mathbf{x} - \mathbf{a}\| < \delta$, it follows that

$$|f(\mathbf{x}) - f(\mathbf{a})| = |x_1 + x_2 - a_1 - a_2| \leq |x_1 - a_1| + |x_2 - a_2| < \delta + \delta = \epsilon.$$

Hence $f(x_1, x_2) = x_1 + x_2$ is continuous at any point $(a_1, a_2) \in \mathbf{R}^2$.

(d) Fix a point $\mathbf{a} = (a_1, a_2) \in \mathbf{R}^2$. Given $\epsilon > 0$, let $\delta = \min\{1, \epsilon(1 + |a_1| + |a_2|)^{-1}\}$. Then if $\|\mathbf{x} - \mathbf{a}\| < \delta$, it follows that

$$\begin{aligned} |f(\mathbf{x}) - f(\mathbf{a})| &= |x_1x_2 - a_1a_2| = |(x_1x_2 - x_1a_2) + (x_1a_2 - a_1a_2)| \\ &\leq |x_1x_2 - x_1a_2| + |x_1a_2 - a_1a_2| = |x_1||x_2 - a_2| + |a_2||x_1 - a_1| \\ &< \delta(|x_1| + |a_2|) < \delta(|a_1| + 1 + |a_2|) \leq \epsilon. \end{aligned}$$

Notice that the final '$<$' follows from the fact that $|x_1 - a_1| < \delta \leq 1$. Hence $f(x_1, x_2) = x_1x_2$ is continuous at any point $\mathbf{a} = (a_1, a_2) \in \mathbf{R}^2$.

(e) Homework exercise.

$\square$

Theorem 1.8. *Linear transformations $T : \mathbf{R}^n \to \mathbf{R}^m$ are continuous.*

This one requires a little warm-up. If A is an $m \times n$ matrix, let us define the *magnitude* of A to be the quantity

$$\|A\| := \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2}.$$

That is, we are measuring the size of A as if it were a vector in $\mathbf{R}^{mn}$. Be aware that elsewhere in mathematics (including Shifrin), there are other notions of the magnitude of a matrix.

Lemma 1.9. *Given any matrix $A \in \mathcal{M}_{m\times n}$ and any vector $\mathbf{x} \in \mathbf{R}^n$, we have $\|A\mathbf{x}\| \leq \|A\| \|\mathbf{x}\|$.*

Proof. Let $\mathbf{row}_i$ denote the ith row of A. Then on the one hand, we have

$$\|A\|^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2 = \sum_{i=1}^{m} \|\mathbf{row}_i\|^2.$$

But on the other hand, the ith entry of $A\mathbf{x}$ is $\mathbf{row}_i \cdot \mathbf{x}$. Hence from the Cauchy-Schwartz inequality, we obtain

$$\|A\mathbf{x}\|^2 = \sum_{i=1}^{m} (\mathbf{row}_i \cdot \mathbf{x})^2 \leq \sum_{i=1}^{m} \|\mathbf{row}_i\|^2 \|\mathbf{x}\|^2 = \|\mathbf{x}\|^2 \sum_{i=1}^{m} \|\mathbf{row}_i\|^2 = \|A\|^2 \|\mathbf{x}\|^2.$$

$\square$

Proof of Theorem 1.8. Let $A \in \mathcal{M}_{m\times n}$ be the standard matrix of the linear transformation $T : \mathbf{R}^n \to \mathbf{R}^m$ and $\mathbf{a} \in \mathbf{R}^n$ be any point. Given $\epsilon > 0$, I choose $\delta = \frac{\epsilon}{\|A\|}$. Then if $\|\mathbf{x} - \mathbf{a}\| < \delta$, it follows that

$$\|T(\mathbf{x}) - T(\mathbf{a})\| = \|T(\mathbf{x} - \mathbf{a})\| = \|A(\mathbf{x} - \mathbf{a})\| \leq \|A\| \, \|\mathbf{x} - \mathbf{a}\| < \|A\| \, \delta = \epsilon.$$

Hence T is continuous at any point $\mathbf{a} \in \mathbf{R}^n$ □

Questions about limits of vector-valued functions can always be reduced to questions about scalar-valued functions.

Proposition 1.10. *Suppose that $F : \mathbf{R}^n \to \mathbf{R}^m$ is a vector-valued function $F = (F_1, \dots, F_m)$ defined near $\mathbf{a} \in \mathbf{R}^n$. Then the following are equivalent.*

(a) $\lim_{\mathbf{x}\to\mathbf{a}} F(\mathbf{x}) = \mathbf{b} \in \mathbf{R}^m$.
(b) $\lim_{\mathbf{x}\to\mathbf{a}} \|F(\mathbf{x}) - \mathbf{b}\| = 0$.
(c) $\lim_{\mathbf{x}\to\mathbf{a}} F_j(\mathbf{x}) = b_j$ *for* $1 \leq j \leq m$.

Proof of Proposition 1.10.

(a) $\implies$ (b) Suppose that $\lim_{\mathbf{x}\to\mathbf{a}} F(\mathbf{x}) = \mathbf{b}$ and set $f(\mathbf{x}) := \|F(\mathbf{x}) - \mathbf{b}\|$. Given $\epsilon > 0$, the definition of limit gives me $\delta > 0$ such that $0 < \|\mathbf{x} - \mathbf{a}\| < \delta$ implies that $\|F(\mathbf{x}) - \mathbf{b}\| < \epsilon$. But this last inequality can be rewritten $|f(\mathbf{x}) - 0| < \epsilon$. Thus $\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x}) = 0$, i.e. (b) holds.

(b) $\implies$ (a) Similar.

(a) $\implies$ (c) Suppose again that $\lim_{\mathbf{x}\to\mathbf{a}} F(\mathbf{x}) = \mathbf{b}$. Fix an index j between 1 and m and let $\epsilon > 0$ be given. Since $\lim_{\mathbf{x}\to\mathbf{a}} F(\mathbf{x}) = \mathbf{b}$, there exists $\delta > 0$ such that $0 < \|\mathbf{x} - \mathbf{a}\| < \delta$ implies that $\|F(\mathbf{x}) - \mathbf{b}\| < \epsilon$. Then by Proposition 1.3, I also have that

$$|F_j(\mathbf{x}) - b_j| \leq \|F(\mathbf{x}) - \mathbf{b}\| < \epsilon.$$

So $\lim_{\mathbf{x}\to\mathbf{a}} F_j(\mathbf{x}) = b_j$.

(c) $\implies$ (a) Suppose that $\lim_{\mathbf{x}\to\mathbf{a}} F_j(\mathbf{x}) = b_j$ for each $1 \leq j \leq m$. Then given any $\epsilon > 0$, there exist real numbers $\delta_j > 0$ such that $0 < \|\mathbf{x} - \mathbf{a}\| < \delta_j$ implies that $|F_j(\mathbf{x}) - b_j| < \frac{\epsilon}{\sqrt{m}}$. Taking $\delta = \min\{\delta_1, \dots, \delta_m\}$, I infer that $0 < \|\mathbf{x} - \mathbf{a}\| < \delta$ implies $|F_j(\mathbf{x}) - b_j| < \frac{\epsilon}{\sqrt{m}}$ for *all* indices j. Thus by the lemma,

$$\|F(\mathbf{x}) - \mathbf{b}\| < \sqrt{m}\frac{\epsilon}{\sqrt{m}} = \epsilon.$$

So $\lim_{\mathbf{x}\to\mathbf{a}} F(\mathbf{x}) = \mathbf{b}$. □

The following theorem is sometimes paraphrased by saying that limits commute with continuous functions.

Theorem 1.11 (composite limits)**.** *Let $F : \mathbf{R}^n \to \mathbf{R}^m$ and $G : \mathbf{R}^m \to \mathbf{R}^p$ be functions and $\mathbf{a} \in \mathbf{R}^n$, $\mathbf{b} \in \mathbf{R}^m$ be points such that $\lim_{\mathbf{x}\to\mathbf{a}} F(\mathbf{x}) = \mathbf{b}$ and G is continuous at $\mathbf{b}$. Then*

$$\lim_{\mathbf{x}\to\mathbf{a}} G \circ F(\mathbf{x}) = G(\lim_{\mathbf{x}\to\mathbf{a}} F(\mathbf{x})) = G(\mathbf{b}).$$

Proof. Let $\epsilon > 0$ be given. By continuity of G at $\mathbf{b}$, there exists a number $\epsilon' > 0$ such that $\|\mathbf{y} - \mathbf{b}\| < \epsilon'$ implies $\|G(\mathbf{y}) - G(\mathbf{b})\| < \epsilon$. Likewise, since $\lim_{\mathbf{x}\to\mathbf{a}} F(\mathbf{x}) = \mathbf{b}$, there exists $\delta > 0$ such that $0 < \|\mathbf{x} - \mathbf{a}\| < \delta$ implies that $\|F(\mathbf{x}) - \mathbf{b}\| < \epsilon'$. Putting these two things together, I see that $0 < \|\mathbf{x} - \mathbf{a}\| < \delta$ further implies that $\|G(F(\mathbf{x})) - G(\mathbf{b})\| < \epsilon$. So $\lim_{\mathbf{x}\to\mathbf{a}} G(F(\mathbf{x})) = G(\mathbf{b})$. □

Corollary 1.12. *Let $F : \mathbf{R}^n \to \mathbf{R}^m$ and $G : \mathbf{R}^m \to \mathbf{R}^p$ be continuous functions. Then $G \circ F$ is continuous.*

Proof. Note that $G \circ F$ is defined at $\mathbf{a} \in \mathbf{R}^n$ precisely when F is defined at $\mathbf{a}$ and G is defined at $F(\mathbf{a})$. Then Since both functions are continuous wherever they are defined, we have

$$\lim_{\mathbf{x}\to\mathbf{a}} G(F(\mathbf{x})) = G(\lim_{\mathbf{x}\to\mathbf{a}} F(\mathbf{x})) = G(F(\mathbf{a})).$$

Hence $G \circ F$ is continuous at any point $\mathbf{a} \in \mathbf{R}^n$ where it is defined. □

Corollary 1.13. *Let $f, g : \mathbf{R}^n \to \mathbf{R}$ be functions with limits $\lim_{x\to a} f(x) = b$ and $\lim_{x\to a} g(x) = c$ at some point $a \in \mathbf{R}^n$. Then*

(a) $\lim_{x\to a} |f(x)| = |b|$.
(b) $\lim_{x\to a} f(x) + g(x) = b + c$;
(c) $\lim_{x\to a} f(x)g(x) = bc$;
(d) $\lim_{x\to a} \frac{1}{f(x)} = \frac{1}{b}$, *provided* $b \neq 0$.

Hence a sum or product of continuous functions is continuous, as is the reciprocal of a continuous function.

Actually, the corollary extends to dot products, magnitudes and sums of vector-valued functions $F, G : \mathbf{R}^n \to \mathbf{R}^m$, too. I'll let you write down the statements of these facts.

Proof. I prove (c). The other parts are similar. Let $F : \mathbf{R} \to \mathbf{R}^2$ be given by $F(x) := (f(x), g(x))$ and $m : \mathbf{R}^2 \to \mathbf{R}$ be the multiplication function $m(y_1, y_2) := y_1 y_2$. Recall that m is continuous (Proposition 1.7). Moreover, Proposition 1.10 and our hypotheses about f and g imply that $\lim_{x\to a} F(x) = (b, c)$. Hence I infer from Theorem 1.11 that

$$\lim_{x\to a} f(x)g(x) = \lim_{x\to a} m(F(x)) = m(\lim_{x\to a} F(x)) = m(b, c) = bc.$$

□

When used with the fact that functions can't have more than one limit at a given point, Theorem 1.11 leads to a useful criterion for establishing that a limit *doesn't* exist.

Definition 1.14. *A parametrized curve is a continuous function $\gamma : \mathbf{R} \to \mathbf{R}^n$.*

Corollary 1.15. *Given a function $F : \mathbf{R}^n \to \mathbf{R}^m$ defined near a point $\mathbf{a} \in \mathbf{R}^n$, suppose that $\gamma_1, \gamma_2 : \mathbf{R} \to \mathbf{R}^n$ are parametrized curves such that $\gamma_1(t) = \gamma_2(t) = a$ if and only if $t = 0$. If the limits $\lim_{t\to 0} F \circ \gamma_1(t)$ and $\lim_{t\to 0} F \circ \gamma_2(t)$ are not equal, then $\lim_{\mathbf{x}\to\mathbf{a}} F(\mathbf{x})$ does not exist.*

Proof. I will prove the contrapositive statement: suppose that $\lim_{\mathbf{x}\to\mathbf{a}} F(\mathbf{x}) = \mathbf{b}$ exists and $\gamma_1, \gamma_2 : \mathbf{R} \to \mathbf{R}^n$ are parametrized curves with initial points $\gamma_j(0) = a$ but $\gamma_j(t) \neq a$ for $t \neq a$. Then the limits $\lim_{\mathbf{x}\to\mathbf{a}} F(\mathbf{x})$ and $\lim_{t\to 0} F \circ \gamma_j(t)$ do *not* concern the value of F at a. So I may assume with no loss of generality that $F(\mathbf{a}) = \mathbf{b}$, i.e. that F is actually continuous at $\mathbf{a}$.

Theorem 1.11 then tells me that

$$\lim_{t\to 0} F \circ \gamma_1(t) = F(\lim_{t\to 0} \gamma_1(t)) = F(\gamma_1(0)) = F(\mathbf{a}) = \mathbf{b}.$$

The second equality holds because γ_1 is continuous. Likewise, $\lim_{t\to 0} F\circ\gamma_2(t) = \mathbf{b}$. In particular, the two limits are the same. $\square$

I remark that if $\gamma : \mathbf{R} \to \mathbf{R}^n$ is a continuous curve and $F : \mathbf{R}^n \to \mathbf{R}^m$ is a function, then the composite function $F\circ\gamma : \mathbf{R} \to \mathbf{R}^m$ is sometimes called the *restriction* of F to γ.

One last fact about limits that will prove useful for us is the following.

Theorem 1.16 (The Squeeze Theorem)**.** *Suppose that* $F : \mathbf{R}^n \to \mathbf{R}^m$ *and* $g : \mathbf{R}^n \to \mathbf{R}$ *are functions defined near* $\mathbf{a} \in \mathbf{R}^n$*. Suppose there exists* $r > 0$ *such that*

- $\|F(\mathbf{x})\| \le |g(\mathbf{x})|$ *for all* $\mathbf{x} \in B^*(\mathbf{a}, r)$*;*
- $\lim_{\mathbf{x}\to\mathbf{a}} g(\mathbf{x}) = 0$.

Then $\lim_{\mathbf{x}\to\mathbf{a}} F(\mathbf{x}) = \mathbf{0}$.

Proof. Exercise. $\square$