MATHEMATICS 23a/E-23a, Fall 2018
Linear Algebra and Real Analysis I
Week 10 (Limits and continuity in $\mathbb{R}^n$, partial derivatives)

Author: Paul Bamberg
R scripts by Paul Bamberg
Last modified: November 16, 2018 by Paul Bamberg (rewrore workshop problem 1a)

Reading

- Hubbard, section 1.5, pages 92 through 99 (limits and continuity)
- Hubbard, section 1.6 up through page 112.
- Hubbard, Appendix A.3 (Heine-Borel)
- Hubbard, section 1.7 up through page 133.

Recorded Lectures

- Lecture 20 (Week 10, Class 1) (watch on November 13 or 14)
- Lecture 21 (Week 10, Class 2) (watch on November 15 or 16)

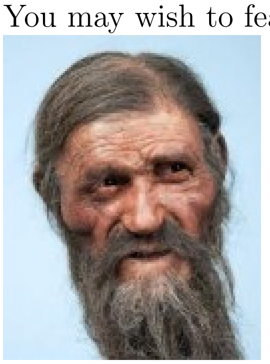
Proofs to present in section or to a classmate who has done them.

- 10.1 Let $X \subset \mathbb{R}^2$ be an open set, and consider $\mathbf{f} : X \to \mathbb{R}^2$. Let $\mathbf{x_0}$ be a point in X. Prove that $\mathbf{f}$ is continuous at $\mathbf{x_0}$ if and only if for every sequence $\mathbf{x_i}$ converging to $\mathbf{x_0}$,

$$\lim_{i\to\infty} \mathbf{f}(\mathbf{x_i}) = \mathbf{f}(\mathbf{x_0}).$$

- 10.2 Using the Bolzano-Weierstrass theorem, prove that a continuous real-valued function f defined on a compact subset $C \subset \mathbb{R}^n$ has a supremum M and that there is a point $\mathbf{a} \in C$ (a maximum) where $f(\mathbf{a}) = M$.

You may wish to feature Ötzi the Iceman as the protagonist of your proof.



R Scripts

- Script3.2A-LimitFunctionR2.R
  Topic 1 - Sequences that converge to the origin
  Topic 2 - Evaluating functions along these sequences

- Script 3.2B-AffineApproximation.R
  Topic 1 - The tangent-line approximation for a single variable
  Topic 2 - Displaying a contour plot for a function
  Topic 3 - The gradient as a vector field
  Topic 4 - Plotting some pathological functions

1 Executive Summary

1.1 Limits in $\mathbb{R}^n$

- To define $\lim_{\mathbf{x}\to\mathbf{x_0}} f(\mathbf{x})$, we need not require that $\mathbf{x_0}$ is in domain of f. We require only that $\mathbf{x_0}$ is in the closure of the domain of f. This requirement guarantees that for any $\delta > 0$ we can find an open ball of radius δ around $\mathbf{x_0}$ that includes points in the domain of f. There is no requirement that all points in that ball be in the domain.

- Limit of a function $\mathbf{f}$ from $\mathbb{R}^n$ to $\mathbb{R}^m$:
  We assume that the domain is a subset $X \subset \mathbb{R}^n$.
  Definition: Function $\mathbf{f} : X \to \mathbb{R}^m$ has the limit $\mathbf{a}$ at $\mathbf{x_0}$:
  $$\lim_{\mathbf{x}\to\mathbf{x_0}} \mathbf{f}(\mathbf{x}) = \mathbf{a}$$
  if $\mathbf{x_0}$ is in the closure of X and $\forall\epsilon > 0$, $\exists\delta > 0$ such that
  $\forall\mathbf{x} \in X$ that satisfy $|\mathbf{x} - \mathbf{x_0}| < \delta$, $|\mathbf{f}(\mathbf{x}) - \mathbf{a}| < \epsilon$.

- $\lim_{\mathbf{x}\to\mathbf{x_0}} \mathbf{f}(\mathbf{x}) = \mathbf{a}$ if and only if for all sequences with $\lim \mathbf{x_n} = \mathbf{x_0}$, $\lim \mathbf{f}(\mathbf{x_n}) = \mathbf{a}$. To show that a function $\mathbf{f}$ does not have a limit as $\mathbf{x} \to \mathbf{x_0}$, invent two different sequences, both of which converge to $\mathbf{x_0}$, for which the sequences of function values do not approach the same limit. Or just invent one sequence for which the sequence $\lim \mathbf{f}(\mathbf{x_n})$ does not converge!

- If $\lim_{\mathbf{x}\to\mathbf{x_0}} \mathbf{f}(\mathbf{x}) = \mathbf{a}$ and $\lim_{\mathbf{x}\to\mathbf{x_0}} \mathbf{f}(\mathbf{x}) = \mathbf{b}$, then $\mathbf{a} = \mathbf{b}$.

- Suppose $\mathbf{f}(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \end{pmatrix}$.
  $\lim_{\mathbf{x}\to\mathbf{x_0}} \mathbf{f}(\mathbf{x}) = \mathbf{a}$ if and only if $\lim_{\mathbf{x}\to\mathbf{x_0}} f_1(\mathbf{x}) = a_1$ and $\lim_{\mathbf{x}\to\mathbf{x_0}} f_2(\mathbf{x}) = a_2$.

- Properties of limits
  These are listed on p. 95 of Hubbard. The proofs are almost the same as for functions of one variable
  –Limit of sum = sum of limits.
  –Limit of product = product of limits.
  –Limit of quotient = quotient of limits if you do not have zero in the denominator.
  –Limit of dot product = dot product of limits. (proved on pages 95-96.)
  These last two useful properties involve a vector-valued function $\mathbf{f}(\mathbf{x})$ and a scalar-valued function $h(\mathbf{x})$, both with domain U.
  –If $\mathbf{f}$ is bounded and h has a limit of zero, then $h\mathbf{f}$ also has a limit of zero.
  –If h is bounded and $\mathbf{f}$ has a limit of zero, then $h\mathbf{f}$ also has a limit of zero.

1.2 Continuous functions in topology and in $\mathbb{R}^n$

- Function f is continuous at x_0 if, for any open set U in the codomain that contains $f(x_0)$, the preimage (inverse image) of U, i.e. the set of points x in the domain for which $f(x) \in U$, is also an open set.
- Here is the definition that lets us extend real analysis to n dimensions. $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ is continuous at x_0 if, for any open "codomain ball" of radius ϵ centered on $\mathbf{f}(\mathbf{x_0})$, we can find an open "domain ball" of radius δ centered on $\mathbf{x_0}$ such that if $\mathbf{x}$ is in the domain ball, $\mathbf{f}(\mathbf{x})$ is in the codomain ball.
- An equivalent condition (your proof 10.1): $\mathbf{f}$ is continuous at $\mathbf{x_0}$ if and only if every sequence that converges to $\mathbf{x_0}$ is a good sequence. We will need to prove this for $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$, but the proof is almost identical to the proof for $f : \mathbb{R} \to \mathbb{R}$, which we have already done.
- As was the case in $\mathbb{R}$, sums, products, compositions, etc. of continuous functions are continuous. If you can write a formula for a function of several variables that does appear to involve division by zero, the theorems on pages 98 and 99 will show that it is continuous.
- To show that a function is discontinuous, construct a bad sequence!

1.3 Compact subsets and Bolzano-Weierstrass

- A subset $X \in \mathbb{R}^n$ is *bounded* if there is some ball, centered on the origin, of which it is a subset. If a nonempty subset $C \in \mathbb{R}^n$ is closed as well as bounded, it is called *compact*.
- Bolzano-Weierstrass theorem in $\mathbb{R}^n$

  The theorem says that given any sequence of points $\mathbf{x_1}, \mathbf{x_2}, ...$ from a compact set C, we can extract a convergent subsequence whose limit is in C.

  Easy proof (Ross,section 13.5)
  In $\mathbb{R}^n$, using the theorem that we have proved for $\mathbb{R}$, extract a subsequence where the first components converge. Then extract a subsequence where the second components converge, continuing for n steps.

  Hubbard, theorem 1.6.3, offers an alternative but nonconstructive proof.

- Existence of a maximum

  The supremum M of function f on set C is the least upper bound of the values of f. The maximum, if it exists, is a point of evaluation: a point $a \in C$ such that $f(a) = M$. Infimum and minimum are defined similarly.

  A continuous real-valued function f defined on a compact subset $C \subset \mathbb{R}^n$ has a supremum M and that there is a point $\mathbf{a} \in C$ (a maximum) where $f(\mathbf{a}) = M$. The proof (your proof 10.2) is similar to the proof in $\mathbb{R}$.

1.4 The nested compact set theorem

$X_k \in \mathbb{R}^n$ is a decreasing sequence of nonempty compact sets: $X_1 \supset X_2 \supset \cdots$. For example, in $\mathbb{R}$, $X_n = [-1/n, 1/n]$. In $\mathbb{R}^2$. we can use nested squares.

$$\text{The theorem states that } \bigcap_{k=1}^{\infty} X_k \neq \emptyset.$$

If $X_k = (0, \frac{1}{k})$ (<u>not</u> compact!), the infinite intersection is the empty set.

The proof (Hubbard, Appendix A.3) starts by choosing a point $\mathbf{x}_k$ from each set X_k, then invokes the Bolzano-Weierstrass theorem to select a convergent subsequence $\mathbf{y}_i$ that converges to a point $\mathbf{a}$ that is contained in each of the X_k and so is also an element of their intersection $\bigcap_{m=1}^{\infty} X_m$.

1.5 The Heine-Borel theorem

The Heine-Borel theorem states that for a compact subset $X \in \mathbb{R}^n$, any open cover contains a finite subcover. In other words, if someone gives you a possibly infinite collection of open sets U_i whose union includes every point in X, you can select a finite number of them whose union still includes every point in X

$$X \subset \bigcup_{i=1}^{m} U_i.$$

The proof (Hubbard, Appendix A.3) uses the nested compact set theorem.

In general topology, where the sets that are considered are not necessarily subsets of $\mathbb{R}^n$, the statement "every open cover contains a finite subcover" is used as the *definition* of "compact set."

1.6 Partial derivatives

If U is an open subset of $\mathbb{R}^n$ and function $f : U \to \mathbb{R}$ is defined by a formula

$$f \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}$$

then its partial derivative with respect to the ith variable is

$$\frac{\partial f}{\partial x_i} = D_i f(\mathbf{a}) = \lim_{h \to 0} \frac{1}{h}(f \begin{pmatrix} a_1 \\ \cdots \\ a_i + h \\ a_n \end{pmatrix} - f \begin{pmatrix} a_1 \\ \cdots \\ a_i \\ a_n \end{pmatrix})$$

This does not give the generalization we want. It specifies a good approximation to f only along a line through $\mathbf{a}$, whereas we would like an approximation that is good in a ball around $\mathbf{a}$.

1.7 Directional derivative, Jacobian matrix, gradient

Let $\vec{\mathbf{v}}$ be the direction vector of a line through $\mathbf{a}$. Imagine a moving particle whose position as a function of time t is given by $\mathbf{a} + t\vec{\mathbf{v}}$ on some open interval that includes $t = 0$. Then $\mathbf{f}(\mathbf{a} + t\vec{\mathbf{v}})$ is a function of the single variable t. The derivative of this function with respect to t is the directional derivative.

More generally, we use h instead of t and define the directional derivative as

$$\nabla_{\vec{\mathbf{v}}} f(\mathbf{a}) = \lim_{h \to 0} \frac{f(\mathbf{a} + h\vec{\mathbf{v}}) - f(\mathbf{a})}{h}$$

If the directional derivative is a linear function of $\vec{\mathbf{v}}$, in which case f is said to be *differentiable* at $\mathbf{a}$, then the directional derivative can be calculated if we know its value for each of the standard basis vectors. Since

$$\nabla_{\vec{\mathbf{e}}_i} f(\mathbf{a}) = \lim_{h \to 0} \frac{f(\mathbf{a} + h\vec{\mathbf{e}}_i) - f(\mathbf{a})}{h} = D_i f(\mathbf{a})$$

we can write

$$\nabla_{\vec{\mathbf{v}}} f(\mathbf{a}) = D_1 f(\mathbf{a}) v_1 + D_2 f(\mathbf{a}) v_2 + \cdots + D_n f(\mathbf{a}) v_n.$$

For a more compact notation, we can make the partial derivatives into a $1 \times n$ matrix, called the *Jacobian matrix*

$$[\mathbf{J} f(\mathbf{a})] = [D_1 f(\mathbf{a}) D_2 f(\mathbf{a}) \cdots D_n f(\mathbf{a})],$$

whereupon

$$\nabla_{\vec{\mathbf{v}}} f(\mathbf{a}) = [\mathbf{J} f(\mathbf{a})] \vec{\mathbf{v}}.$$

Alternatively, we can make the partial derivatives into a column vector, the gradient vector

$$\text{grad } f(\mathbf{a}) = \begin{bmatrix} D_1 f(\mathbf{a}) \\ D_2 f(\mathbf{a}) \\ \cdots \\ D_n f(\mathbf{a}) \end{bmatrix},$$

so that

$$\nabla_{\vec{\mathbf{v}}} f(\mathbf{a}) = \text{grad } f(\mathbf{a}) \cdot \vec{\mathbf{v}}.$$

We now have, for differentiable functions (and we will soon prove that if the partial derivatives of f are continuous, then f is differentiable), a useful generalization of the tangent-line approximation of single variable calculus.

$$f(\mathbf{a} + h\vec{\mathbf{v}}) \approx f(\mathbf{a}) + [\mathbf{J} f(\mathbf{a})](h\vec{\mathbf{v}})$$

This sort of approximation (a constant plus a linear approximation) is called an"affine approximation."

2 Lecture outline

1. Limit of a function $\mathbf{f}$ from $\mathbb{R}^n$ to $\mathbb{R}^m$:

   We do not want to assume that the domain of $\mathbf{f}$ is all of $\mathbb{R}^n$. So we assume that the domain is a subset $X \subset \mathbb{R}^n$.

   Definition: Function $\mathbf{f} : X \to \mathbb{R}^m$ has the limit $\mathbf{a}$ at $\mathbf{x_0}$:

   $$\lim_{\mathbf{x} \to \mathbf{x_0}} \mathbf{f}(\mathbf{x}) = \mathbf{a}$$

   if $\mathbf{x_0}$ is in the closure of X and

   $\forall \epsilon > 0$, $\exists \delta > 0$ such that

   $\forall \mathbf{x} \in X$ that satisfy $|\mathbf{x} - \mathbf{x_0}| < \delta$

   $|\mathbf{f}(\mathbf{x}) - \mathbf{a}| < \epsilon$

   Draw a diagram to illustrate this definition for the case $m = n = 2$.

2. Theorems about limits of functions

   Propositions 1.5.21 and 1.5.22 in Hubbard are essentially repeats of the proofs we just did for sequences.

   - If $\lim_{\mathbf{x}\to\mathbf{x_0}} \mathbf{f}(\mathbf{x}) = \mathbf{a}$ and $\lim_{\mathbf{x}\to\mathbf{x_0}} \mathbf{f}(\mathbf{x}) = \mathbf{b}$, then

     How would you prove it?

   - Suppose $\mathbf{f}(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \end{pmatrix}$.

     If $\lim_{\mathbf{x}\to\mathbf{x_0}} \mathbf{f}(\mathbf{x}) = \mathbf{a}$, what is $\lim_{\mathbf{x}\to\mathbf{x_0}} f_1(\mathbf{x})$?

     How would you prove it?

     If $\lim_{\mathbf{x}\to\mathbf{x_0}} f_1(\mathbf{x}) = a_1$ and $\lim_{\mathbf{x}\to\mathbf{x_0}} f_2(\mathbf{x}) = a_2$, what can you say about $\lim_{\mathbf{x}\to\mathbf{x_0}} \mathbf{f}(\mathbf{x})$?

     How would you prove it?

3. Properties of limits

   These are boring to prove, and fundamentally the proofs are almost the same as for functions of one variable. They are listed on p. 95 of Hubbard.
   Limit of sum = sum of limits.
   Limit of product = product of limits.
   Limit of quotient = quotient of limits if the denominator is nonzero.
   Limit of dot product = dot product of limits.

   The last of these is proved on pages 95-96. The proof is tedious.

   The next two are a bit less obvious and are very useful. Both involve a vector-valued function $\mathbf{f}(\mathbf{x})$ and a scalar-valued function $h(\mathbf{x})$, both with domain U. The proof of the first one is a homework problem.

   If $\mathbf{f}$ is bounded and h has a limit of zero, then $h\mathbf{f}$ also has a limit of zero.
   If h is bounded and $\mathbf{f}$ has a limit of zero, then $h\mathbf{f}$ also has a limit of zero.

4. Continuous functions in topology and in $\mathbb{R}^n$

   - It is possible to define continuity using only the concept of "open set." Function f is continuous at x_0 if, for any open set U in the codomain that contains $f(x_0)$, the preimage (inverse image) of U, i.e. the set of points x in the domain for which $f(x) \in U$, is also an open set.
     As an application, show that the constant function that maps every page of our standard six-page Web site into the one and only page of a one-page Web site Y is continuous.

   - Here is the definition that lets us extend real analysis to n dimensions. $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ is continuous at x_0 if, for any open "codomain ball" of radius ϵ centered on $\mathbf{f}(\mathbf{x_0})$, we can find an open "domain ball" of radius δ centered on $\mathbf{x_0}$ such that if $\mathbf{x}$ is in the domain ball, $\mathbf{f}(\mathbf{x})$ is in the codomain ball.

   - An equivalent condition (your proof 10.1):
     $\mathbf{f}$ is continuous at $\mathbf{x_0}$ if and only if every sequence that converges to $\mathbf{x_0}$ is a good sequence. We will need to prove this for $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$, but the proof is almost identical to the proof for $f : \mathbb{R} \to \mathbb{R}$, which we have already done.

   - As was the case in $\mathbb{R}$, sums, products, compositions, etc. of continuous functions are continuous. If you can write a formula for a function of several variables that does appear to involve division by zero, the theorems on pages 98 and 99 will show that it is continuous.

   - To show that a function is discontinuous, construct a bad sequence!

5. Proof 10.1, part 1
   If function $\mathbf{f}$ is continuous, every sequence is good.
   Given that function $f : \mathbb{R}^k \to \mathbb{R}^m$ is continuous at $\mathbf{x_0}$, prove that every sequence such that $\mathbf{x_n} \to \mathbf{x_0}$ is a "good sequence" in the sense that $\mathbf{f}(\mathbf{x_n})$ converges to $\mathbf{f}(\mathbf{x_0})$.

6. Proof 10.1, part 2
   If function $\mathbf{f}$ is discontinuous, there exists a bad sequence.
   Given that function $f : \mathbb{R}^k \to \mathbb{R}^m$ is discontinuous at $\mathbf{x_0}$, show how to construct a "bad sequence" such that $\mathbf{x_i} \to \mathbf{x_0}$ but $\mathbf{f}(\mathbf{x_i})$ does not converge to $\mathbf{f}(\mathbf{x_0})$.

Warning: Don't try to prove this theorem using the topological definition of continuity – you cannot! The proof needs the additional assumption that if a set S contains the limit points of all sequences in S, then it is a closed set (its complement is open). For details take Math 131.

7. When you evaluate a limit, you must consider all sequences

$$\text{The simplest example is the function } f\begin{pmatrix} x \\ y \end{pmatrix} = \frac{x^2 - y^2}{x^2 + y^2}$$

Here are some wrong ways to try to evaluate the (non-existent) limit at $x = y = 0$. All of are similar in spirit to Ross's $\lim_{x \to a_S} f(x)$. The problem is that in $\mathbb{R}^n$, there are many more choices for the set S.

- Let S be the x-axis. Set $y = 0$, and then evaluate $\lim_{x\to 0} f\begin{pmatrix} x \\ 0 \end{pmatrix}$ or equivalently, $\lim f\begin{pmatrix} 1/n \\ 0 \end{pmatrix}$.

- Let S be the y-axis. Set $x = 0$, and then evaluate $\lim_{y\to 0} f\begin{pmatrix} 0 \\ y \end{pmatrix}$ or equivalently, $\lim f\begin{pmatrix} 0 \\ 1/n \end{pmatrix}$.

- Let S be the line $y = x$. Set $y = x = t$ and then evaluate $\lim_{t\to 0} f\begin{pmatrix} t \\ t \end{pmatrix}$ or equivalently, $\lim f\begin{pmatrix} 1/n \\ 1/n \end{pmatrix}$.

All these calculations are correct, but the limit does not exist! If you think of this function on the plane in terms of polar coordinates and note that it equals $\cos 2\theta$, it is clear that in any ball (disk) around the origin, no matter how small, the function assumes all values from -1 to 1 and so has no limit at the origin.

Alternatively, construct a sequence $\begin{pmatrix} x_n \\ y_n \end{pmatrix}$ that converges to the origin for which $\lim f\begin{pmatrix} x_n \\ y_n \end{pmatrix}$ does not exist

8. Using proof 10.1 to test for continuity

   We now have a practical technique for showing that a function is discontinuous. Just invent one bad sequence. The best choice is a sequence for which the value of the function is independent of i, because then it is easy to take the limit of the sequence of function values!

   Example: Show that the function defined by $f\begin{pmatrix}0\\0\end{pmatrix} = 0$ and

   $$f\begin{pmatrix}x\\y\end{pmatrix} = \frac{|y|e^{-\frac{|y|}{x^2}}}{x^2} \text{ elsewhere is discontinuous at the origin.}$$

   - Let $x_i = \frac{1}{i}, y_i = \frac{1}{i^2}$. Show that this sequence converges to the origin.

   - Evaluate $f\begin{pmatrix}x_i\\y_i\end{pmatrix}$ and show that its limit is not zero.

   The theorem does not lead to a practical way of testing that a function is continuous, because it involves "for every sequence $\mathbf{x_i}$ converging to $\mathbf{x_0}$." Fortunately, in $\mathbb{R}^2$ and $\mathbb{R}^3$, you can often find a way of evaluating the function at an arbitrary point in a ball of radius h.

   Example: Show that the function defined by

   $$f\begin{pmatrix}0\\0\end{pmatrix} = 0 \text{ and } f\begin{pmatrix}x\\y\end{pmatrix} = \frac{x^2y^2}{x^2+y^2} \text{ is continuous at the origin}$$

   - Set $x = r\cos\theta, y = r\sin\theta$ and evaluate f.

   - I want you to make $f\begin{pmatrix}x\\y\end{pmatrix} < \epsilon$ in a ball of radius r around the origin. Show that $r = \sqrt{\epsilon}$ does the job.

9. Continuity and discontinuity in $\mathbb{R}^3$

(a) Define

$$F\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{xyz}{x^2+y^2+z^2}, F\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = 0.$$

Prove that F is continuous at the origin.

(b) Define

$$g\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{xy+xz+yz}{x^2+y^2+z^2}, g\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = 0.$$

Prove that g is discontinuous at the origin.

10. Compact subsets

    A subset $X \in \mathbb{R}^n$ is bounded if there is some ball, centered on the origin, of which it is a subset. For example, the set $X \in \mathbb{R}^3$ of all financial predictions for 2014, where x_1 = U.S. budget surplus, x_2 = number of grams of carbon dioxide in the atmosphere, x_3 = number of spam emails sent, is bounded. So is the open rectangular region obtained by setting bounds on each component like

    $-5 \times 10^{11} < x_1 < 2 \times 10^{11}$.

    If a nonempty subset $C \in \mathbb{R}^n$ is closed as well as bounded, it is called compact. This is not the general definition of "compact" in topology, but for our purposes it is equivalent. In appendix A.3 ia a more general definition of "compact" that requires only the concept of open set.

11. Pigeonhole principle

    The usual version, due to Dirichlet, says that if n pigeons inhabit m pigeonholes and $n > m$, then at least one pigeonhole contains more than one pigeon.

    We need the version that says that if a countable infinity of pigeons inhabit a finite number of pigeonholes, then at least one pigeonhole contains an infinite number of pigeons.

    Give the easy proof by contradiction for each case.

12. Bolzano-Weierstrass theorem

The theorem says that given any sequence of points $\mathbf{x_1}, \mathbf{x_2}, ...$ from a compact set C, we can extract at least one convergent subsequence whose limit $\mathbf{b}$ is in C. We have already proved this theorem in $\mathbb{R}$ using Ross's ingenious "dominant term" approach.

One proof for $\mathbb{R}^n$ goes as follows:

- Extract a subsequence for which the sequence of first components converges to $\mathbf{b}_1$.
- From this, extract a subsequence for which the sequence of second components converges to $\mathbf{b}_2$.
- $\cdots$
- From this, extract a subsequence for which the sequence of nth components converges to $\mathbf{b}_n$.

An alternative is the proof from page 107 of Hubbard. For ease of visualization, choose subset C to be a subset of $\mathbb{R}^2$ that lies inside the closed disk of radius 10.

Now break the square that extends from -10 to 10 along each axis into 400 unit squares. At least one of these 400 squares contains infinitely many elements of the subsequence. Without loss of generality we can assume that it is one where both components are positive. So we might, for example, extract an infinite subsequence of points, all of the form $\begin{pmatrix} 6.xxxx \\ 3.xxxx \end{pmatrix}$

Now break this square into 100 subsquares. At least one of these contains infinitely many elements of the subsequence. By choosing a specific one of these subsqares, we can extract an infinite subsequence of points, for example, points that are all of the form $\begin{pmatrix} 6.4xxx \\ 3.7xxx \end{pmatrix}$

By induction on the number of digits m after the decimal point, we may conclude that it is possible to extract an infinite subsequence whose elements all agree in their first m digits, for example, points that are all of the form $\begin{pmatrix} 6.4259...7xxx \\ 3.7012...3xxx \end{pmatrix}$

We need to know that C is closed to be sure that the point to which the sequence converges is actually an element of C. If C were the open disk of radius 10, that the one and only convergent sequence might consist of points whose first components are 9.0 9.90, 9.990, 9.9990,...and whose other components are all zero. All points in that sequence lie in the open disk of radius 10, but their limit does not, since its first component is 10.

This page left blank for a good diagram to illustrate the Bolzano-Weierstrass theorem.

13. Proof 10.2: on a compact set, a continuous function has a maximum. The proof is the same as in $\mathbb{R}$. Here is a fanciful version,

    A continuous real-valued function f defined on a compact subset $C \subset \mathbb{R}^n$ has a supremum M and $\exists$ point $\mathbf{a} \in C$ (a maximum) where $f(\mathbf{a}) = M$.

    Ötzi the Iceman, whose mummy is the featured exhibit at the archaeological museum in Bolzano, Italy, has a goal of camping at the greatest altitude M on the Tyrol, a compact subset of the earth's surface on which altitude is a continuous function f of latitude and longitude.

    (a) Assume that there is no supremum M. Then Ötzi can select a sequence of campsites in C such that $f(\mathbf{x}_1) > 1$, $f(\mathbf{x}_2) > 2$,... $f(\mathbf{x}_n) > n, \cdots$. Show how to use Bolzano-Weierstrass to construct a "bad sequence," in contradiction to the assumption that f is continuous.

    (b) On night n, Ötzi chooses a campsite whose altitude exceeds $M - 1/n$. From this sequence, extract a convergent subsequence, and call its limit $\mathbf{a}$. Show that $f(\mathbf{a}) = M$, so $\mathbf{a}$ is a maximum, and M is not merely a supremum but a maximum value.

14. An application to football: why "compact set" is important

    A school playground is a compact subset $C \subset \mathbb{R}^2$. Two aspiring quarterbacks are playing catch with a football, and they want to get as far apart as possible. Show that if $\sup |\mathbf{x} - \mathbf{y}| = D$ for any two points in C, they can find a pair of points $\mathbf{x_0}$ and $\mathbf{y}_0$ such that $|\mathbf{x}_0 - \mathbf{y}_0| = D$. Then invent simple examples to show that this cannot be done if the playground is unbounded or is not closed.

15. Cauchy sequences in $\mathbb{R}^n$

- Prove that every Cauchy sequence of vectors $\vec{\mathbf{a}}_1, \vec{\mathbf{a}}_2, \cdots \in \mathbb{R}^n$ is bounded: i.e. $\exists M$ such that $\forall n, |\vec{\mathbf{a}}_n| < M$.
  Hint: $\vec{\mathbf{a}}_n = \vec{\mathbf{a}}_n - \vec{\mathbf{a}}_m + \vec{\mathbf{a}}_m$. When showing that a sequence is bounded, you can ignore the first N terms.
- Prove that if a sequence $\mathbf{a}_1, \mathbf{a}_2, \cdots \in \mathbb{R}^n$ converges to $\mathbf{a}$, it is a Cauchy sequence. Hint: $\mathbf{a}_m - \mathbf{a}_n = \mathbf{a}_m - \mathbf{a} + \mathbf{a} - \mathbf{a}_n$. Use the triangle inequality.
- Prove that every convergent sequence of vectors $\vec{\mathbf{a}}_1, \vec{\mathbf{a}}_2, \cdots \in \mathbb{R}^n$ is bounded (very easy, given the preceding results.)

16. Nested compact sets

    You have purchased a nice chunk of Carrara marble from which to carve the term project for your GenEd course on Italian Renaissance sculpture. On day 1 the marble occupies a compact subset X_1 of the space in your room. You chip away a bit every evening, hoping to reveal the masterpiece that is hidden in the marble, and you thereby create a decreasing sequence of nonempty compact sets: $X_1 \supset X_2 \supset \cdots$.
    Your understanding instructor gives you an infinite extension of time on the project. Prove that there is a point $\mathbf{a}$ that forever remains in the marble, no matter how much you chip away; i.e. that

$$\bigcap_{k=1}^{\infty} X_k \neq \emptyset.$$

17. Heine-Borel theorem (proved in $\mathbb{R}^2$, but the proof is the same for $\mathbb{R}^n$.)

    Suppose that you need security guards to guard a compact subset $X \in \mathbb{R}^2$.

    Heine-Borel Security, LLC proposes that you should hire an infinite number of their guards, each of whom will patrol an open subset U_i of $\mathbb{R}^2$. These guards protect all of X: the union of their patrol zones is an "open cover."

    Prove that you can fire all but a finite number m of the security guards (not necessarily the first m) and your property will still be protected:

    $$X \subset \bigcup_{i=1}^{m} U_i.$$

    Break up the part of the city where your property lies into closed squares, each 1 kilometer on a side. There will exist a square B_0 that needs infinitely many guards (the "infinite pigeonhole principle").

    Break up this square into 4 closed subsquares: again, at least one will need infinitely many guards. Choose one subsquare and call it B_1. Continue this procedure to get a decreasing sequence B_i of nested compact sets, whose intersection includes a point $\mathbf{a}$.

    Now show that any guard whose *open* patrol zone includes $\mathbf{a}$ can replace all but a finite number of other guards.

18. Converse of Heine-Borel in $\mathbb{R}$

The converse of Heine-Borel says that if the U.S government is hiring Heine-Borel security to guard a subset X of the road from Mosul to Damascus and wants to be sure that they do not have to pay an infinite number of guards, then X has to be closed and bounded.

(a) What happens if Heine-Borel assigns guard k to patrol the open interval $(-k, k)$?

(b) What happens if Heine-Borel selects a point x_0 that is not in X and assigns guard k to patrol the interval $(x_0 - 1/k, x_0 + 1/k)$.?

19. Directional derivative, Jacobian matrix, gradient

Let $\vec{\mathbf{v}}$ be the direction vector of a line through $\mathbf{a}$. Imagine a moving particle whose position as a function of time t is given by $\mathbf{a}+t\vec{\mathbf{v}}$ on some open interval that includes $t = 0$. Then $\mathbf{f}(\mathbf{a} + t\vec{\mathbf{v}})$ is a function of the single variable t. The derivative of this function with respect to t is the directional derivative.

More generally, use h instead of t and define the directional derivative as

$$\nabla_{\vec{\mathbf{v}}} f(\mathbf{a}) = \lim_{h \to 0} \frac{f(\mathbf{a} + h\vec{\mathbf{v}}) - f(\mathbf{a})}{h}$$

If the directional derivative is a linear function of $\vec{\mathbf{v}}$, in which case f is said to be *differentiable* at $\mathbf{a}$, then the directional derivative can be calculated if we know its value for each of the standard basis vectors. Since

$$\nabla_{\vec{\mathbf{e}}_i} f(\mathbf{a}) = \lim_{h \to 0} \frac{f(\mathbf{a} + h\vec{\mathbf{e}}_i) - f(\mathbf{a})}{h} = D_i f(\mathbf{a})$$

we can write

$$\nabla_{\vec{\mathbf{v}}} f(\mathbf{a}) = D_1 f(\mathbf{a}) v_1 + D_2 f(\mathbf{a}) v_2 + \cdots + D_n f(\mathbf{a}) v_n.$$

For a more compact notation, we can make the partial derivatives into a $1 \times n$ matrix, called the *Jacobian matrix*

$$[\mathbf{J} f(\mathbf{a})] = [D_1 f(\mathbf{a}) D_2 f(\mathbf{a}) \cdots D_n f(\mathbf{a})],$$

whereupon

$$\nabla_{\vec{\mathbf{v}}} f(\mathbf{a}) = [\mathbf{J} f(\mathbf{a})]\vec{\mathbf{v}}.$$

Alternatively, we can make the partial derivatives into a column vector, the gradient vector

$$\text{grad } f(\mathbf{a}) = \begin{bmatrix} D_1 f(\mathbf{a}) \\ D_2 f(\mathbf{a}) \\ \cdots \\ D_n f(\mathbf{a}) \end{bmatrix},$$

so that

$$\nabla_{\vec{\mathbf{v}}} f(\mathbf{a}) = \text{grad } f(\mathbf{a}) \cdot \vec{\mathbf{v}}.$$

We now have, for differentiable functions (and we will soon prove that if the partial derivatives of f are continuous, then f is differentiable), a useful generalization of the tangent-line approximation of single variable calculus.

$$f(\mathbf{a} + h\vec{\mathbf{v}}) \approx f(\mathbf{a}) + [\mathbf{J} f(\mathbf{a})](h\vec{\mathbf{v}})$$

This sort of approximation (a constant plus a linear approximation) is called an "affine approximation."

20. Constructing an affine approximation to a nonlinear function

Let $f\begin{pmatrix} x \\ y \end{pmatrix} = \sqrt{xy^3}$.

- Evaluate the Jacobian matrix of f at $\begin{pmatrix} 4 \\ 1 \end{pmatrix}$ and use it to find the best affine approximation to $f(\begin{pmatrix} 4 \\ 1 \end{pmatrix} + t\begin{pmatrix} 2 \\ 1 \end{pmatrix})$ for small t.

- Test the approximation for $t = 0.1$ and for $t = 0.01$.
From a calculator: $f(\begin{pmatrix} 4.2 \\ 1.1 \end{pmatrix} = 2.364436...$ $f(\begin{pmatrix} 4.02 \\ 1.01 \end{pmatrix} = 2.0351437...$

- By defining $g(t) = f(\begin{pmatrix} 4 \\ 1 \end{pmatrix} + t\begin{pmatrix} 2 \\ 1 \end{pmatrix})$, you can convert this problem to one in single-variable calculus. Show that using the tangent-line approximation near $t = 0$ leads to exactly the same answer.

21. A cautionary tale about partial derivatives, which are a concept of single-variable calculus.

Let

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \frac{x^2 y}{x^4 + y^2}.$$

f is defined to be 0 at $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Show that both partial derivatives are zero at $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ but that the function is not continuous there.

22. A clever application of the gradient vector

The Cauchy-Schwarz inequality says that

grad $f \cdot \mathbf{v} \leq |\text{grad} f||\mathbf{v}|$, with equality when grad f and $\mathbf{v}$ are proportional.

If $\mathbf{v}$ is a unit vector, the maximum value of the directional derivative occurs when $\mathbf{v}$ is a multiple of grad f.

Suppose that the temperature T in a open subset of the plane is given by $T\begin{pmatrix} x \\ y \end{pmatrix} = 25 + 0.1x^2y^3$. If you are at $x = 1, y = 2$, along what direction should you walk to have temperature increase most rapidly?

3 Seminar Topics

Your section instructor will either have emailed a list of topics to prepare or will have posted a signup list of appointments on the Calendar tab of Canvas. Either way, there will be one of the following topics that you should be prepared to present.

Practice your presentation so that it takes about 8 minutes. The text of the presentation will be projected onto a screen so that you need not recopy it. To save time, avoid writing long sentences on the chalkboard. You may use notes, but be discreet about it.

1. (Proof 10.1)
   - Given that function $f : \mathbb{R}^k \to \mathbb{R}^m$ is continuous at $\mathbf{x_0}$, prove that every sequence such that $\mathbf{x_n} \to \mathbf{x_0}$ is a "good sequence" in the sense that $\mathbf{f}(\mathbf{x_n})$ converges to $\mathbf{f}(\mathbf{x_0})$.
   - Given that function $f : \mathbb{R}^k \to \mathbb{R}^m$ is discontinuous at $\mathbf{x_0}$, show how to construct a "bad sequence" such that $\mathbf{x_i} \to \mathbf{x_0}$ but $\mathbf{f}(\mathbf{x_i})$ does not converge to $\mathbf{f}(\mathbf{x_0})$.
2. (Proof 10.2) Using the Bolzano-Weierstrass theorem, prove that a continuous real-valued function f defined on a compact subset $C \subset \mathbb{R}^n$ has a supremum M and that there is a point $\mathbf{a} \in C$ (a maximum) where $f(\mathbf{a}) = M$.
3. Define what is meant by a Cauchy sequence of vectors in $\mathbb{R}^n$ and prove that any convergent sequence is Cauchy.
4. State (but do not prove) the nested compact set theorem for $\mathbb{R}^n$.
   Then explain what is meant by an "open cover" of a subset $X \subset \mathbb{R}^n$ and state (but do not prove) the Heine-Borel theorem. Show that in $\mathbb{R}$, for the closed but unbounded set $[0, \infty]$ you can invent an open cover that does not have a finite subcover, then do the same for the set $(0, 1]$, which is bounded but not closed.
5. For a function $f : \mathbb{R}^n \to \mathbb{R}$, define the directional derivative
   $$\nabla_{\vec{\mathbf{v}}} f(\mathbf{a}),$$
   the partial derivative $D_i f(\mathbf{a})$, and the gradient vector grad $f(\mathbf{a})$.

   Then prove that if the directional derivative is a linear function of $\vec{\mathbf{v}}$,
   $$\nabla_{\vec{\mathbf{v}}} f(\mathbf{a}) = \text{grad } f(\mathbf{a}) \cdot \vec{\mathbf{v}}.$$

4 Workshop Problems

1. Theorems related to Heine-Borel

   (a) Converse of Heine-Borel in the one-dimensional case

   As the best topologist in the U.S. Army, you have been deployed to the border between Arizona and Mexico, where Congress has authorized funds to build a wall that includes a certain subset X of the border. The Compact Border Security Act does not specify X (for obvious national security reasons), but it includes a provision that any proposal to construct the wall on a day-by-day schedule must cover all of X within a finite number of days. The contractor is the venerable German firm Heine-Borel Borders GmbH, best known for its construction of the Berlin Wall, which has been teaching topology to the leaders of Germany since the days of Frederick the Great.

   Agents of the local drug cartel are easily persuaded to reveal to you one randomly chosen point $\mathbf{x}_0$ that is not in the set X and pay you a handsome bribe to make sure that no wall is built there. You ask Heine-Borel for a proposal, specifying, "but stay clear of $\mathbf{x}_0$." The resulting proposal includes, in the first k days, no construction between $\mathbf{x}_0 - 1/k$ and $\mathbf{x}_0 + 1/k$.

   i. Prove that set X is closed.

   ii. A fellow soldier has been deployed to south Texas, where the Rio Grande wiggles back and forth so much that he suspects that its length may be infinite. Prove for him that any set X that complies with the Compact Border Security Act must be finite in extent.

   (b) The converse of the Heine-Borel theorem states that if every open cover of set $X \in \mathbb{R}^n$ contains a finite subcover, then X must be closed and bounded.

   i. By choosing as the open cover a set of open balls of radius $1, 2, \cdots$, prove that X must be bounded.

   ii. To show that X is closed, show that its complement X^c must be open. Hint: choose any $\mathbf{x_0} \in X^c$ and choose an open cover of X in which the kth set consists of points whose distance from $\mathbf{x_0}$ is greater than $\frac{1}{k}$. This open cover of X must have a finite subcover.

   If you need a further hint, look on pages 90 and 91 of Chapter 2 of Ross.

2. Limits and continuity in $\mathbb{R}^2$ and $\mathbb{R}^3$

(a) Define

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \frac{xy^3}{x^2+y^6}, f\begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0.$$

Show that the sequence $\begin{pmatrix} \frac{1}{i} \\ \frac{1}{i} \end{pmatrix}$ is "good" but that $\begin{pmatrix} \frac{1}{i^3} \\ \frac{1}{i} \end{pmatrix}$ is "bad."

(b) • Let

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \frac{xy(x^2-y^2)}{(x^2+y^2)^2}, f\begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0.$$

Invent a "bad sequence" of points $(\mathbf{a}_1, \mathbf{a}_2, \cdots)$ that converges to $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ for which

$$\lim_{i\to\infty} f(\mathbf{a}_i) \neq 0.$$

This bad sequence proves that f is discontinuous at $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

• Let

$$g\begin{pmatrix} x \\ y \end{pmatrix} = \frac{xy(x^2-y^2)}{x^2+y^2}, g\begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0.$$

By introducing polar coordinates, prove that g is continuous at $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

3. Using partial derivatives to find approximate function values

(a) Let $f\begin{pmatrix} x \\ y \end{pmatrix} = x^2y$. Evaluate the Jacobian matrix of f at $\begin{pmatrix} 2 \\ 0.5 \end{pmatrix}$ and use it to find the best affine approximation to $f\begin{pmatrix} 1.98 \\ 0.51 \end{pmatrix}$ and to $f\begin{pmatrix} 1.998 \\ 0.501 \end{pmatrix}$. Use a calculator or R, find the "remainder" (the difference between the actual function value and the best affine approximation) in each case. You should find that the remainder decreases by a factor that is much greater than 10.

(b) Let $f\begin{pmatrix} x \\ y \end{pmatrix} = y + \log(xy)$ (natural logarithm) for $x, y > 0$. Evaluate the Jacobian matrix of f at $\begin{pmatrix} 0.5 \\ 2 \end{pmatrix}$ and use it to find the best affine approximation (constant plus linear approximation) to $f\begin{pmatrix} 0.51 \\ 2.02 \end{pmatrix}$.

5 Homework

1. You are the mayor of El Dorado. Not all the streets are paved with gold – only the interval [0,1] on Main Street – but you still have a serious security problem, and you ask Heine-Borel Security LLC to submit a proposal for keeping the street safe at night. Knowing that the city coffers are full, they come up with the following pricey plan for meeting your requirements by using a countable infinity of guards:

   - Guard 0 patrols the interval $(-\frac{1}{N}, \frac{1}{N})$, where you may choose any value greater than 100 for the integer N. She is paid 200 dollars.
   - Guard 1 patrols the interval $(0.4, 1.2)$ and is paid 100 dollars.
   - Guard 2 patrols the interval $(0.2, 0.6)$ and is paid 90 dollars.
   - Guard 3 patrols the interval $(0.1, 0.3)$ and is paid 81 dollars.
   - Guard k patrols the interval $(\frac{0.8}{2^k}, \frac{2.4}{2^k})$ and is paid $100(0.9)^{k-1}$ dollars.

   (a) Calculate the total cost of hiring this infinite set of guards (sum a geometric series).

   (b) Show that the patrol regions of the guards form an "open cover" of the interval [0,1].

   (c) According to the Heine-Borel theorem, this infinite cover has a finite subcover. Explain clearly how to construct it. (Hint: look at the proof of the Heine-Borel theorem)

   (d) Suppose that you want to protect only the open interval (0,1), which is not a compact subset of Main Street. In what very simple way can Heine-Borel Security modify their proposal so that you are forced to hire infinitely many guards?

2. Hubbard, Exercise 1.6.6. You might want to work parts (b) and (c) before attempting part (a). The function $f(x)$ is defined for all of $\mathbb{R}$, which is not a compact set, so you will have to do some work before applying theorem 1.6.9. Notice that "a maximum" does not have to be unique: a function could achieve the same maximum value at more than one point.

3. Singular Point, California is a spot in the desert near Death Valley that is reputed to have been the site of an alien visit to Earth. In response to a campaign contribution from AVSIG, the Alien Visitation Special Interest Group, the government has agreed to survey the region around the site.

   In the vicinity, the altitude is given by the function

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \frac{2x^2y}{x^4+y^2}.$$

   A survey team that traveled through the Point going west to east declares that the altitude at the Point itself is zero. A survey team that went south to north would comment only that zero was perhaps a reasonable interpolation.

   (a) Suppose you travel through the Point along the line $y = mx$, passing through the point at time $t = 0$ and moving with a constant velocity such that $x = t$: in other words, $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} t \\ mt \end{pmatrix}$. Find a function $g(m, t)$ that gives your altitude as a function of time on this journey. Sketch graphs of g as a function of t for $m = 1$ and for $m = 3$. Is what happens for large m consistent with what happens on the y axis?

   (b) Find a sequence of points that converges to $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$, for which $x_n = \frac{1}{n}$ and $f\begin{pmatrix} x \\ y \end{pmatrix} = 1$ for every point in the sequence. Do the same for $f\begin{pmatrix} x \\ y \end{pmatrix} = -1$.

   (c) Is altitude a continuous function at Singular Point? Explain.

4. (a) Hubbard, exercise 1.7.12. This is good practice in approximating a function by using its derivative and seeing how fast the "remainder" goes to zero.

   (b) Hubbard, exercise 1.7.4. These are all problems in single-variable calculus, but they cannot be solved by using standard differentation formulas. You have to use the definition of the derivative as a limit.

5. Linearity of the directional derivative.

   Suppose that, near the point $\mathbf{a} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, the Celsius temperature is specified by the function $f\begin{pmatrix} x \\ y \end{pmatrix} = 20 + xy^2$.

   (a) Suppose that you drive with a constant velocity vector $\vec{\mathbf{v}}_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, passing through the point $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ at time $t = 0$. Express the temperature outside your car as a function $g(t)$ and use single-variable calculus to calculate $g'(0)$, the rate at which the reading on your car's thermometer is changing. You have calculated the directional derivative of f along the vector $\vec{\mathbf{v}}_1$ by using single-variable calculus.

   (b) Do the same for the velocity vector $\vec{\mathbf{v}}_2 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$.

   (c) As it turns out, the given function f is differentiable, and the directional derivative is therefore a linear function of velocity. Use this fact to determine the directional derivative of f along the standard basis vector $\vec{\mathbf{e}}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ from your earlier answers, and confirm that your answer agrees with the partial derivative $D_2 f(\mathbf{a})$.

   (d) Remove all the mystery from this problem by recalculating the directional derivatives using the formula $[Df(\mathbf{a})]\vec{\mathbf{v}}$.

6. Let $f\begin{pmatrix} x \\ y \end{pmatrix} = x\sqrt{y}$. Evaluate the Jacobian matrix of f at $\begin{pmatrix} 2 \\ 4 \end{pmatrix}$ and use it to find the best affine approximation to $f\begin{pmatrix} 1.98 \\ 4.06 \end{pmatrix}$.

   As you can confirm by using a calculator, $1.98\sqrt{4.06} = 3.989589452...$

7. (a) Hubbard, Exercise 1.7.22. This is a slight generalization of a topic that was presented in lecture. The statement is in terms of derivatives, but it is equivalent to the version that uses gradients.

   (b) An application: suppose that you are skiing on a mountain where the height above sea level is described by the function $f\begin{pmatrix} x \\ y \end{pmatrix} = 1 - 0.2x^2 - 0.4y^2$ (with the kilometer as the unit of distance, this is not unreasonable). You are located at the point $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Find a unit vector $\vec{\mathbf{v}}$ along the direction in which you should head if you want to head straight down the mountain and two unit vectors $\vec{\mathbf{w}}_1$ and $\vec{\mathbf{w}}_2$ that specify directions for which your rate of descent is only $\frac{3}{5}$ of the maximum rate.

   (c) Prove that in general, the unit vector for which the directional derivative is greatest is orthogonal to the direction along which the directional derivative is zero, and use this result to find a unit vector $\vec{\mathbf{u}}$ appropriate for a timid but lazy skier who wants to head neither down nor up.