MATHEMATICS 23a/E-23a, Fall 2018
Linear Algebra and Real Analysis I
Fortnight 1 (Fields, Vectors, and Matrices)

Authors: Paul Bamberg and Kate Penner
R scripts by Paul Bamberg
Last modified: September 12, 2018 by Paul Bamberg

Reading

- Hubbard, Sections 0.1 through 0.4
- Hubbard, Sections 1.1, 1.2, and 1.3
- Lawvere and Schanuel, Conceptual Mathematics
  See the main page of the website for temporary access. At a minimum, read the following:
  Article I (Sets, maps, composition – definition of a category)
  Session 2
  This is very easy reading.

  DO NOT PURCHASE THIS BOOK! We will be using only a little bit of it.

Recorded Lectures In 2015, when these lectures were recorded, the first class met on the Thursday before Labor Day, and there were three lectures before the Thursday or Friday after students registered for the course. This year, there would have been four meetings of a Tuesday-Thursday class. I have spread the three lectures over two weeks, and we are going to try to have class meetings on Sept. 5 and 6, before course registration is complete, to go over the first half of the material.

- Lecture 1 (Fortnight 1, Class 1) (watch on September 4 or 5 – 70 minutes)
- Lecture 2 (Fortnight 1, Class 2) (watch first 58 minutes on September 5 or 6)
- Lecture 2 (Fortnight 1, Class 2) (watch last 20 minutes, starting from 58 minute mark, on September 10 or 11)
- Lecture 3 (Fortnight 1, Class 3) (watch on September 11 or 12)

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

- 1.1 Suppose that a and b are two elements of a field F. Using only the axioms for a field, prove the following:
  - $\forall a \in F$, $0a = 0$.
  - If $ab = 0$, then either a or b must be 0.
  - The additive inverse of a is unique.
- 1.2(Generalization of Hubbard, proposition 1.2.9) A is an $n \times m$ matrix. The entry in row i, column j is $a_{i,j}$
  B is an $m \times p$ matrix.
  C is an $p \times q$ matrix.
  The entries in these matrices are all from the same field F. Using summation notation, prove that matrix multiplication is associative:
  that $(AB)C = A(BC)$. Include a diagram showing how you would lay out the calculation in each case so the intermediate results do not have to be recopied.
- 1.3 (Hubbard, proposition 1.3.14) Suppose that linear transformation $T : F^n \to F^m$ is represented by the $m \times n$ matrix $[T]$.
  - Suppose that the matrix $[T]$ is invertible. Prove that the linear transformation T is one-to-one and onto (injective and surjective), hence invertible.
  - Suppose that linear transformation T is invertible. Prove that its inverse S is linear and that the matrix of S is $[S] = [T]^{-1}$

  Note: Use * to denote matrix multiplication and $\circ$ to denote composition of linear transformations. You may take it as already proved that matrix multiplication represents composition of linear transformations. Do not assume that $m = n$. That is true, but we are far from being able to prove it, and you do not need it for the proof.

R Scripts

- Script 1.1A-Finite Fields.R
  Topic 1 - Why the real numbers form a field
  Topic 2 - Making a finite field, with only five elements
  Topic 3 - A useful rule for finding multiplicative inverses

- Script 1.1B-PointsVectors.R
  Topic 1 - Addition of vectors in $\mathbb{R}^2$
  Topic 2 - A diagram to illustrate the point-vector relationship
  Topic 3 - Subtraction and scalar multiplication

- Script 1.1C-Matrices.R
  Topic 1 - Matrices and Matrix Operations in R
  Topic 2 - Solving equations using matrices
  Topic 3 - Linear functions and matrices
  Topic 4 - Matrices that are not square
  Topic 5 - Properties of the determinant

- Script 1.1D-MarkovMatrix
  Topic 1 - A game of volleyball
  Topic 2 - traveling around on ferryboats

- Script 1.1L-LinearMystery
  Topic 1 - Define a mystery linear function $fMyst : \mathbb{R}^2 \to \mathbb{R}^2$

1 Executive Summary

- Quantifiers and Negation Rules
  The "universal quantifier" $\forall$ is read "for all."
  The "existential quantifier" $\exists$ is read "there exists." It is usually followed by "s.t," a standard abbreviation for "such that."
  The negation of "$\forall x, P(x)$ is true" is "$\exists x, P(x)$ is not true."
  The negation of "$\exists x, P(x)$ is true" is "$\forall x, P(x)$ is not true."
  The negation of "P and Q are true" is "either P or Q is not true."
  The negation of "either P or Q is true" is "both P and Q are not true."

- Functions
  A function f needs two sets: its domain X and its codomain Y.
  f is a rule that, to any element $x \in X$, assigns a specific element $y \in Y$. We write $y = f(x)$.
  f must assign a value to every $x \in X$, but not every $y \in Y$ must be of the form $f(x)$. The subset of the codomain consisting of elements that are of the form $y = f(x)$ is called the *image* of f. If the image of f is all of the codomain Y, f is called *surjective* or *onto*.
  f need not assign different of elements of Y to different elements of X. If $x_1 \neq x_2 \implies f(x_1) \neq f(x_2)$, f is called *injective* or *one-to-one*
  If f is both surjective and injective, it is *bijective* and has an inverse f^{-1}.

- Categories
  A category $\mathscr{C}$ has objects (which might be sets) and arrows (which might be functions)
  An arrow f must have a specific domain object X and a specific codomain object Y; we write $f : X \to Y$ or $X \xrightarrow{f} Y$.
  If arrows $f : X \to Y$ and $g : Y \to Z$ are in the category, then the composition arrow $g \circ f : X \to Z$ is in the category.
  For every object X there must be an identity arrow $I_X : X \to X$
  Identity laws: Given $f : X \to Y$, $f \circ I_X = f$ and $I_Y \circ f = f$.
  Associative law: given $X \xrightarrow{f} Y \xrightarrow{g} Z \xrightarrow{h} W$, $h \circ (g \circ f) = (h \circ g) \circ f$
  Given an arrow $f : X \to Y$, an arrow $g : Y \to X$ such that $g \circ f = I_X$ is called a *retraction*.
  Given an arrow $f : X \to Y$, an arrow $g : Y \to X$ such that $f \circ g = I_Y$ is called a *section*.
  If, for arrow f, arrow g is both a retraction and a section, then g is the inverse of f, $g = f^{-1}$, and g must be unique.
  Almost everything in mathematics is a special case of a category.

1.1 Fields and Field Axioms

A **field** F is a set of elements for which the familiar operations of addition and multiplication are defined and behave in the usual way. Here is a set of axioms for a field. You can use them to prove theorems that are true for any field.

1. Addition is commutative: $a + b = b + a$.
2. Addition is associative: $(a + b) + c = a + (b + c)$.
3. Additive identity: $\exists 0$ such that $\forall a \in F, 0 + a = a + 0 = a$.
4. Additive inverse: $\forall a \in F, \exists -a$ such that $-a + a = a + (-a) = 0$.
5. Multiplication is associative: $(ab)c = a(bc)$.
6. Multiplication is commutative: $ab = ba$.
7. Multiplicative identity: $\exists 1$ such that $\forall a \in F, 1a = a$.
8. Multiplicative inverse: $\forall a \in F - \{0\}, \exists a^{-1}$ such that $a^{-1}a = 1$.
9. Distributive law: $a(b + c) = ab + ac$.

Examples of fields include:
The **rational numbers** $\mathbb{Q}$.
The **real numbers** $\mathbb{R}$.
The **complex numbers** $\mathbb{C}$.
The **finite field** $\mathbb{Z}_p$, constructed for any prime number p as follows:

- Break up the set of integers into p subsets. Each subset is named after the remainder when any of its elements is divided by p.

  $[a]_p = \{m | m = np + a, n \in \mathbb{Z}\}$

  Notice that $[a + kp]_p = [a]_p$ for any k. There are only p sets, but each has many alternate names. These p infinite sets are the elements of the field $\mathbb{Z}_p$.

- Define addition by $[a]_p + [b]_p = [a + b]_p$. Here a and b can be any names for the subsets, because the answer is independent of the choice of name. The rule is "Add a and b, then divide by p and keep the remainder."

- Define multiplication by $[a]_p[b]_p = [ab]_p$. Again a and b can be any names for the subsets, because the answer is independent of the choice of name. The rule is "Multiply a and b, then divide by p and keep the remainder."

1.2 Points and Vectors

F^n denotes the set of ordered lists of n elements from a field F. Usually the field is $\mathbb{R}$, but it could be the field of complex numbers $\mathbb{C}$ or a finite field like $\mathbb{Z}_5$.

A given element of F^n can be regarded either as a point, which represents "position data," or as a vector, which represents "incremental data."

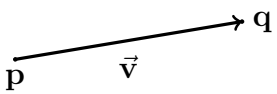
If an element of F^n is a point, we represent it by a bold letter like $\mathbf{p}$ and write it as a column of elements enclosed in parentheses.

$$\mathbf{p} = \begin{pmatrix} 1.1 \\ -3.8 \\ 2.3 \end{pmatrix}$$

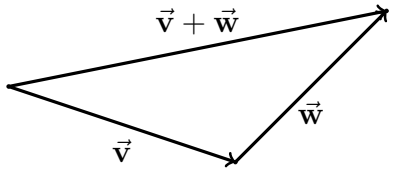
If an element of F^n is a vector, we represent it by a bold letter with an arrow like $\vec{\mathbf{v}}$ and write it as a column of elements enclosed in square brackets.

$$\vec{\mathbf{v}} = \begin{bmatrix} -0.2 \\ 1.3 \\ 2.2 \end{bmatrix}$$

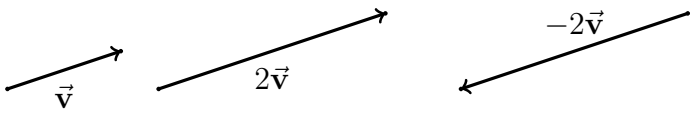
To add a vector to a point, we add the components in identical positions together. The result is a point: $\mathbf{q} = \mathbf{p} + \vec{\mathbf{v}}$. Geometrically we represent this by anchoring the vector at the initial point $\mathbf{p}$. The location of the arrowhead of the vector is the point $\mathbf{q}$ that represents our sum.



To add a vector to a vector, we again add component by component. The result is a vector. Geometrically, the vector created by beginning at the initial point of the first vector and ending at the arrowhead of the second vector is the represents our sum.

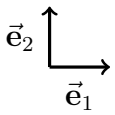


To form a scalar multiple of a vector, we multiply each component by the scalar. In $\mathbb{R}^n$, the geometrical effect is to multiply the length of the vector by the scalar. If the scalar is a negative number, we switch the position of the arrow to the other end of the vector.



1.3 Standard basis vectors

The **standard basis vector** $\vec{\mathbf{e}}_k$ has a 1 as its kth component, and all its other components are 0. Since the additive identity 0 and the multiplicative identity 1 must be present an any field, there will always be n standard basis vectors in F^n. Geometrically, the standard basis vectors in $\mathbb{R}^2$ are usually associated with "one unit east" and "one unit north" respectively.



1.4 Matrices and linear transformations

An $m \times n$ **matrix** over a field F has m rows and n columns.
Matrices represent linear functions, also known as **linear transformations**:
A function $\mathbf{g} : F^n \to F^m$ is called linear if

$$g(a\vec{\mathbf{v}} + b\vec{\mathbf{w}}) = ag(\vec{\mathbf{v}}) + bg(\vec{\mathbf{w}}).$$

For a linear function $\mathbf{g}$, if we know the value of $\mathbf{g}(\vec{\mathbf{e}}_i)$ for each standard basis vector $\vec{\mathbf{e}}_\mathbf{i}$, the value of $\mathbf{g}(\vec{\mathbf{v}})$ for any vector v follows by linearity:

$$\mathbf{g}(v_1\vec{\mathbf{e}}_1 + v_2\vec{\mathbf{e}}_2 + \cdots + v_n\vec{\mathbf{e}}_n) = v_1\mathbf{g}(\vec{\mathbf{e}}_1) + v_2\mathbf{g}(\vec{\mathbf{e}}_2) + \cdots + v_n\mathbf{g}(\vec{\mathbf{e}}_n)$$

The matrix G that represents the linear function $\mathbf{g}$ is formed by using $\mathbf{g}(\vec{\mathbf{e}}_k)$ as the kth column. Then, if $g_{i,j}$ denotes the entry in the ith row and jth column of matrix G, the function value $\vec{\mathbf{w}} = \mathbf{g}(\vec{\mathbf{v}})$ can be computed by the rule

$$w_i = \sum_{j=1}^{n} g_{i,j}v_j$$

1.5 Matrix multiplication

If $m \times n$ matrix G represents linear function $\mathbf{g} : F^n \to F^m$ and $n \times p$ matrix H represents linear function $\mathbf{h} : F^p \to F^n$, then the matrix product GH is defined so that it represents their composition: the linear function $\mathbf{g} \circ \mathbf{h} : F^p \to F^m$.

Start with standard basis vector $\vec{\mathbf{e}}_j$. Function $\mathbf{h}$ converts this to the jth column $\vec{\mathbf{h}}_j$ of matrix H. Then function $\mathbf{g}$ converts this column to $\mathbf{g}(\vec{\mathbf{h}}_j)$, which must therefore be the jth column of matrix GH.

The rule for forming the product GH can be stated in terms of the rule for a matrix acting on a vector: to form GH, just multiply G by each column of H in turn, and put the results side by side to create the matrix GH. If $C = GH$,

$$c_{i,j} = \sum_{k=1}^{n} g_{i,k}h_{k,j}.$$

While matrix multiplication is associative, it is not commutative. Order matters!

1.6 Examples of matrix multiplication

$$B \begin{bmatrix} 0 & 1 \\ 2 & -1 \\ -2 & 0 \end{bmatrix} \qquad A \begin{bmatrix} 2 & 1 & 0 \\ 1 & -1 & -2 \end{bmatrix}$$

$$A \begin{bmatrix} 2 & 1 & 0 \\ 1 & -1 & -2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 2 & 2 \end{bmatrix} AB \qquad B \begin{bmatrix} 0 & 1 \\ 2 & -1 \\ -2 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 & -2 \\ 3 & 3 & 2 \\ -4 & -2 & 0 \end{bmatrix} BA$$

The number of columns in the first factor must equal the number of rows in the second factor.

1.7 Function inverses

A function $f : X \to Y$ is invertible if it has the following two properties:

- It is **injective** (one-to-one): if $f(x_1) = f(x_2)$, then $x_1 = x_2$.
- It is **surjective** (onto): $\forall y \in Y, \exists x \in X$ such that $f(x) = y$.

The inverse function $g = f^{-1}$ has the property that if $f(x) = y$ then $g(y) = x$. So $g(f(x)) = x$ and $f(g(y)) = y$. Both $f \circ g$ and $g \circ f$ are the identity function.

1.8 The determinant of a 2×2 matrix

For matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $\det A = ad - bc$. If you fix one column, it is a linear function of the other column, and it changes sign if you swap the two columns.

1.9 Matrix inverses

A non-square $m \times n$ matrix A can have a "one-sided **inverse**."

If $m > n$, then A takes a vector in $\mathbb{R}^n$ and produces a longer vector in $\mathbb{R}^m$. In general, there will be many matrices B that can recover the original vector in $\mathbb{R}^n$, so that $BA = I_n$. In this case there is no right inverse.

If $m < n$, then A takes a vector in $\mathbb{R}^n$ and produces a shorter vector in $\mathbb{R}^m$. In general, there will be no left inverse matrix B that can recover the original vector in $\mathbb{R}^n$, but there may be many different right inverses for which $AB = I_m$.

For a square matrix, it is possible for both a right inverse B and a left inverse C to exist. In this case, we can prove that B and C are equal and they are unique. We can say that "an inverse" A^{-1} exists, and it represents the inverse of the linear function represented by matrix A.

You can find the inverse of a 2×2 matrix A whose determinant is not zero by using the formula

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

1.10 Matrix transposes

The **transpose** of a given matrix A is written A^T. The two are closely related. The rows of A are the columns of A^T and the columns of A are the rows of A^T.

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, A^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

The transpose of a matrix product is the product of the transposes, but in the opposite order:

$$(AB)^T = B^T A^T$$

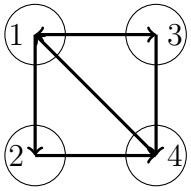
A similar rule holds for matrix inverses:

$$(AB)^{-1} = B^{-1} A^{-1}$$

1.11 Applications of matrix multiplication

In these examples, the "sum of products" rule for matrix multiplication arises naturally, and so it is efficient to use matrix techniques.

- Counting paths: Suppose we have four islands connected by ferry routes:



The entry in row i, column j of the matrix $A = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}$ shows how many ways there are to reach island i by a single ferry ride, starting from island j. The entry in row i, column j of the matrix A^n shows how many ways there are to reach island i by a sequence of n ferry rides, starting from island j.

- Markov processes: A game of beach volleyball has two "states": in state 1, team 1 is serving, in state 2, team 2 is serving. With each point that is played there is a "state transition" governed by probabilities: for example, from state 1, there is a probability of 0.8 of remaining in state 1, a probability of 0.2 of moving to state 2. The transition probabilities can be collected into a matrix like $A = \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix}$. Then the matrix A^n specifies the transition probabilities that result from playing n consecutive points.

2 Lecture Outline

1. Quantifiers and negation

   Especially when you are explaining a proof to someone, it saves some writing to use the symbols $\exists$ (there exists) and $\forall$ (for all).

   Be careful when negating these.

   The negation of "$\forall x, P(x)$ is true" is "$\exists x, P(x)$ is not true."

   The negation of "$\exists x, P(x)$ is true" is "$\forall x, P(x)$ is not true."

   When negating a statement, also bear in mind that

   The negation of "P and Q are true" is "either P or Q is not true."

   The negation of "either P or Q is true" is "both P and Q are not true."

   For practice, let's negate the following statements (which may or may not be true!)

   - There exists an even prime number.
     Negation:

   - All 11-legged alligators are orange with blue spots. (Hubbard, page 5)
     Negation:

   - The function $f(x)$ is continuous on the open interval (0,1), which means that $\forall x \in (0,1), \forall \epsilon > 0, \exists \delta > 0$ such that $\forall y \in (0,1)$,
     $|y - x| < \delta$ implies $|f(y) - f(x)| < \epsilon$.

     Negation: $f(x)$ is discontinuous on the open interval (0,1) means that

2. Set notation
   Here are the standard set-theoretic symbols:

   - $\in$ (is an element of)
   - $\{a|p(a)\}$ (the set of elements a for which $p(a)$ is true)
   - $\subset$ (is a subset of)
   - $\cap$ (intersection)
   - $\cup$ (union)
   - $\times$ (Cartesian product)
   - - or $\setminus$ (set difference)

   Using the integers $\mathbb{Z}$ and the real numbers $\mathbb{R}$, let's construct some sets. In each case there is one way to describe the set using a restriction and another more constructive way to describe the set.

   - The set of real numbers whose cube is greater than 8 in magnitude.
     Restrictive:

     Constructive:

   - The set of coordinate pairs for points on the circle of radius 2 centered at the origin (an example of a "smooth manifold").
     Restrictive:

     Constructive:

3. Function terminology:

Here are some terms that should be familiar from your study of precalculus and calculus:

	Example a	Example b	Example c
domain			
codomain			
image			
one-to-one = injective			
onto = surjective			
invertible = bijective			

Using the sets $X = \{1, 2\}$ and $Y = \{A, B, C\}$, draw diagrams to illustrate the following functions, and fill in the table to show how the terms apply to them:

- $f : X \to Y, f(1) = A, f(2) = B$.

- $g : Y \to X, g(A) = 1, g(B) = 2, g(C) = 1$.

- $h : Y \to Y, h(A) = B, h(B) = C, h(C) = A$. (a permutation)

Here are those function words again, with two additions:

- domain
- natural domain (often deduced from a formula)
- codomain
- image
- one-to-one = injective
- onto = surjective
- invertible = bijective
- inverse image = $\{x | f(x) \in A\}$

Here are functions from $\mathbb{R}$ to $\mathbb{R}$, defined by formulas.

- $f_1(x) = x^2$
- $f_2(x) = x^3$
- $f_3(x) = \log x$(natural logarithm)
- $f_4(x) = e^x$

- Find one that is not injective (not one-to-one)

- For f_1, what is the inverse image of $(1, 4)$?

- Which function is invertible as a function from $\mathbb{R}$ to $\mathbb{R}$?

- What is the natural domain of f_3?

- What is the image of f_4?

- Specify domain and codomain so that f_3 and f_4 are inverses of one another.

- Did your calculus course use "range" as a synonym for "image" or for "codomain?"

4. Composition of functions
   Sometimes people find that a statement is hard to prove because it is so obvious. An example is the associativity of function composition, which will turn out to be crucial for linear algebra.

   Prove that $(f \circ g) \circ h = f \circ (g \circ h)$. Hint: Two functions f_1 and f_2 are equal if they have the same domain X and, $\forall x \in X$, $f_1(x) = f_2(x)$.

   Consider the set of men who have exactly one brother and exactly one son.

   $h(x) =$ "father of x", $g(x) =$ "brother of x", $f(x) =$ "oldest son of x"

   - $f \circ g$ is called
   - $(f \circ g) \circ h$ is
   - $g \circ h$ is called
   - $f \circ (g \circ h)$ is
   - Simpler name for both $(f \circ g) \circ h$ and $f \circ (g \circ h)$

   Consider the real-valued functions

   $g(x) = e^x$, $h(x) = 3 \log x$, $f(x) = x^2$

   - $f \circ g$ has the formula
   - $(f \circ g) \circ h$ has the formula
   - $g \circ h$ has the formula
   - $f \circ (g \circ h)$ has the formula
   - Simpler formula for both $(f \circ g) \circ h$ and $f \circ (g \circ h)$

5. Finite sets and functions form the simplest example of a *category*

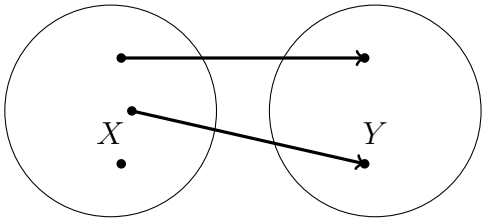
   - The *objects* of the category are finite sets.
   - The *arrows* of the category are functions from one finite set to another. The definition of a function involves quantifiers.

     Requirements for a function $f : X \to Y$

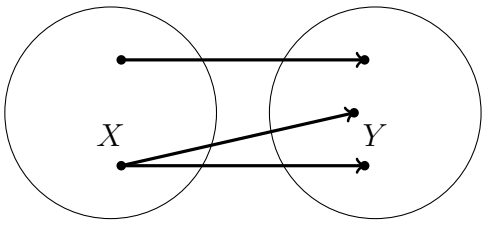
     $\forall x \in X, \exists! y \in Y$ such that $f(x) = y$.

     (The notation $\exists! y$ means "there exists a *unique* y.)"

     What is wrong with the following?

     What is wrong with the following?

   - If arrows $f : X \to Y$ and $g : Y \to Z$ are in the category, then the composition arrow $g \circ f : X \to Z$ is in the category.
   - For any object X there is an identity arrow $I_X : X \to X$
   - Given $f : X \to Y$, $f \circ I_X = f$ and $I_Y \circ f = f$
   - Composition of arrows is associative:

     Given $X \xrightarrow{f} Y \xrightarrow{g} Z \xrightarrow{h} W$, $h \circ (g \circ f) = (h \circ g) \circ f$

   The objects do not have to be sets and the arrows do not have to be functions. For example, the objects could be courses, and an arrow from course X to course Y could mean "if you have taken course X, you will probably do better in course Y as a result." Check that the identity and composition rules are satisfied.

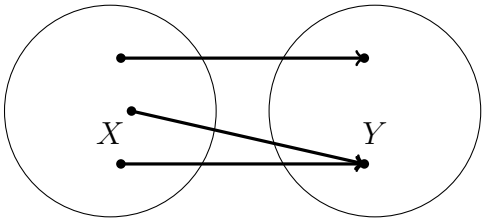
6. Invertible functions - an example of invertible arrows

   First consider the category of finite sets and functions between them.

   The term "inverse" is used only for a "two-sided inverse." Given $f : X \to Y$, an inverse $f^{-1} : Y \to X$ must have the properties
   $f^{-1} \circ f = I_X$ and $f \circ f^{-1} = I_Y$

   Prove that the inverse is unique. This proof uses only things that are true in any category, so it is valid in any category!

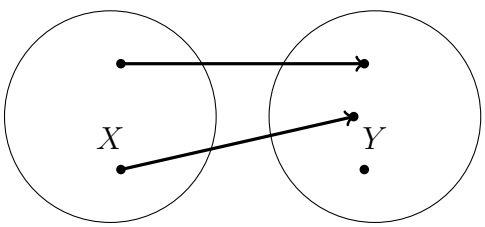
   This function is not invertible because it is not injective, but it is surjective.



   However, it has a "preinverse" (my terminology – the official word is "section.") Starting at an element of Y, choose any element of X from which there is an arrow to that element. Call that function g. Then $f \circ g = I_Y$ but $g \circ f \neq I_X$. Furthermore, g is not unique.

   Prove the cancellation law that if f has a section and $h \circ f = k \circ f$, then $h = k$ (another proof that is valid in any category!)

   This function f is not invertible because it is not surjective, but it is injective.



   It has a "postinverse" g (the official word is "retraction"). First reverse all the arrows to undo its effect, then define g any way you like on the element of Y that is not in the image of f. Then $g \circ f = I_X$ but $f \circ g \neq I_Y$.

7. Fields

Loosely speaking, a field F is a set of elements for which the familiar operations of arithmetic are defined and behave in the usual way. Here is a set of axioms for a field. You can use them to prove theorems that are true for any field.

(a) Addition is commutative: $a + b = b + a$.

(b) Addition is associative: $(a + b) + c = a + (b + c)$.

(c) Additive identity: $\exists 0$ such that $\forall a \in F, 0 + a = a + 0 = a$.

(d) Additive inverse: $\forall a \in F, \exists -a$ such that $-a + a = a + (-a) = 0$.

(e) Multiplication is associative: $(ab)c = a(bc)$.

(f) Multiplication is commutative: $ab = ba$.

(g) Multiplicative identity: $\exists 1$ such that $\forall a \in F, 1a = a$.

(h) Multiplicative inverse: $\forall a \in F - \{0\}, \exists a^{-1}$ such that $a^{-1}a = 1$.

(i) Distributive law: $a(b + c) = ab + ac$.

This set of axioms for a field includes properties (such as the commutativity of addition) that can be proved as theorems by using the other axioms. It therefore does not qualify as an "independent" set, but there is no general requirement that axioms be independent.

Some well-known laws of arithmetic are omitted from the list of axioms because they are easily proved as theorems. The most obvious omission is

$\forall a \in F, 0a = 0$.

Here is the proof. What axiom justifies each step?

- $0 + 0 = 0$ so $(0 + 0)a = 0a$.
- $0a + 0a = 0a$.
- $(0a + 0a) + (-0a) = 0a + (-0a)$.
- $0a + (0a + (-0a)) = 0a + (-0a)$.
- $0a + 0 = 0$.
- $0a = 0$.

8. Finite fields

   Computing with real numbers by hand can be a pain, and most of linear algebra works for an arbitrary field, not just for the real and complex numbers. Alas, the integers do not form a field because in general there is no multiplicative inverse. Here is a simple way to make from the integers a finite field in which messy fractions cannot arise.

   - Choose a prime number p.
   - Break up the set of integers into p subsets. Each subset is named after the remainder when any of its elements is divided by p.

     $[0]_p = \{m | m = np, n \in \mathbb{Z}\}$

     $[1]_p = \{m | m = np + 1, n \in \mathbb{Z}\}$

     $[a]_p = \{m | m = np + a, n \in \mathbb{Z}\}$

     Notice that $[a + kp]_p = [a]_p$ for any k. There are only p sets, but each has many alternate names.

     These p infinite sets are the elements of the field $\mathbb{Z}_p$.
   - Define addition by $[a]_p + [b]_p = [a + b]_p$. Here a and b can be any names for the subsets, because the answer is independent of the choice of name. The rule is "Add a and b, then divide by p and keep the remainder."
   - What is the simplest name for $[5]_7 + [4]_7$?
   - What is the simplest name for the additive inverse of $[3]_7$?
   - Define multiplication by $[a]_p[b]_p = [ab]_p$. Again a and b can be any names for the subsets, because the answer is independent of the choice of name. The rule is "Multiply a and b, then divide by p and keep the remainder."
   - What is the simplest name for $[5]_7[4]_7$?
   - Find the multiplicative inverse for each nonzero element of $\mathbb{Z}_7$

9. Rational numbers
   The rational numbers $\mathbb{Q}$ form a field. You learned how to add and multiply them years ago! The multiplicative inverse of $\frac{a}{b}$ is $\frac{b}{a}$ as long as $a \neq 0$.

   The rational numbers are not a "big enough" field for doing Euclidean geometry or calculus. Here are some irrational quantities:

   - $\sqrt{2}$
   - π.
   - most values of trig functions, exponentials, or logarithms.
   - coordinates of most intersections of two circles.

10. Real numbers
    The real numbers $\mathbb{R}$ constitute a field that is large enough so that any characterization of a number in terms of an infinite sequence of real numbers still leads to a real number.

    A positive real number is an expression like 3.141592... where there is no limit to the number of decimal places that can be provided if requested. To get a negative number, put a minus sign in front. This is Hubbard's definition.

    An equivalent viewpoint is that a positive real number is the sum of an integer and an infinite series of the form

$$\sum_{i=1}^{\infty} a_i (\frac{1}{10})^i$$

    where each a_i is one of the decimal digits 0...9.

    Write the first three terms of an infinite series that converges to π.

    The rational numbers and the real numbers are both "ordered fields." This means that there is a subset of positive elements that is closed under both addition and multiplication. No finite field is ordered.

    In $\mathbb{Z}_5$, you can name the elements $[0], [1], [2], [-2], [-1]$, and try to call the elements [1] and [2] "positive." Why does this attempt to make an ordered field fail?

11. Proof 1.1 - two theorems that are valid in any field

    (a) Using nothing but the field axioms and the theorem that $0a = 0$, prove that if $ab = 0$, then either a or b must be 0.

    (b) Using nothing but the field axioms, prove that the additive inverse of an element a is unique. (Standard strategy for uniqueness proofs: assume that there are two different inverses b and c, and prove that $b = c$.

12. Lists of field elements as points and vectors:

    F^n denotes the set of ordered lists of n elements from a field F. Usually the field is $\mathbb{R}$, but it could be the field of complex numbers $\mathbb{C}$ or a finite field like $\mathbb{Z}_5$.

    An element of F^n can be regarded either as a point, which represents "position data," or as a vector, which represents "incremental data." Beware: many textbooks ignore this distinction!

    If an element of F^n is a point, we represent it by a bold letter like $\mathbf{p}$ and write it as a column of elements enclosed in parentheses.

$$\mathbf{p} = \begin{pmatrix} 1.1 \\ -3.8 \\ 2.3 \end{pmatrix},$$

    If an element of F^n is a vector, we represent it by by a bold letter with an arrow like $\vec{\mathbf{v}}$ and write it as a column of elements enclosed in square brackets.

$$\vec{\mathbf{v}} = \begin{bmatrix} -0.2 \\ 1.3 \\ 2.2 \end{bmatrix}$$

13. Relation between points and vectors, inspired by geometry:

    - Add vector $\vec{\mathbf{v}}$ component by component to point $\mathbf{A}$ to get point $\mathbf{B}$.
    - Subtract point $\mathbf{A}$ component by component from point $\mathbf{B}$ to get vector $\vec{\mathbf{v}}$.
    - Vector addition: if adding $\vec{\mathbf{v}}$ to point $\mathbf{A}$ gives point $\mathbf{B}$ and adding $\vec{\mathbf{w}}$ to point $\mathbf{B}$ gives point $\mathbf{C}$, then adding $\vec{\mathbf{v}} + \vec{\mathbf{w}}$ to point $\mathbf{A}$ gives point $\mathbf{C}$.
    - A vector in F^n can be multiplied by any element of F to get another vector.

    Draw a diagram to illustrate these operations without use of coordinates, as is typically done in a physics course.

14. Examples from coordinate geometry

Here are two points in the plane.

$$\mathbf{p} = \begin{pmatrix} 1.4 \\ -3.8 \end{pmatrix}, \mathbf{q} = \begin{pmatrix} 2.4 \\ -4.8 \end{pmatrix}$$

Here are two vectors.

$$\vec{\mathbf{v}} = \begin{bmatrix} -0.2 \\ 1.3 \end{bmatrix}, \vec{\mathbf{w}} = \begin{bmatrix} 0.6 \\ -0.2 \end{bmatrix}$$

- What is $\mathbf{q} - \mathbf{p}$?

- What is $\mathbf{p} + \vec{\mathbf{v}}$?

- What is $\vec{\mathbf{v}} - 1.5\vec{\mathbf{w}}$?

- What, if anything, is $\mathbf{p} + \mathbf{q}$?

- What is $0.5\mathbf{p} + 0.5\mathbf{q}$? Why is this apparently illegal operation OK?

15. Subspaces of F^n

A subspace is defined only when the elements of F^n are vectors. It must be closed under vector addition and scalar multiplication. The second requirement means that the zero vector must be in the subspace. The empty set $\emptyset$ is not a subspace!

Geometrically, a subspace corresponds to a "flat subset" (line, plane, etc.) *that includes the origin.*

For $\mathbb{R}^3$ there are four types of subspace. What is the geometric interpretation of each?

- 0-dimensional: the set $\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \}$
- 1-dimensional: $\{t\vec{\mathbf{u}} | t \in \mathbb{R}\}$
  Exception: 0-dimensional if

- 2-dimensional: $\{s\vec{\mathbf{u}} + t\vec{\mathbf{v}} | s, t \in \mathbb{R}\}$
  Exception: 1-dimensional if

- 3-dimensional: $\{r\vec{\mathbf{u}} + s\vec{\mathbf{v}} + t\vec{\mathbf{w}} | r, s, t \in \mathbb{R}\}$
  Exceptions: 2-dimensional if

  1-dimensional if

A special type of subset is obtained by adding all the vectors in a subspace to a fixed point. It is in general not a subspace, but it has special properties. Lines and planes that do not contain the origin fall into this category.

We call such a subset an "affine subset." This terminology is not standard: the Math 116 textbook uses "linear variety."

16. Standard basis vectors:

    These are useful when we want to think of F^n more abstractly.

    The standard basis vector $\vec{\mathbf{e_i}}$ has a 1 in position i, a 0 everywhere else. Since 0 and 1 are in every field, these vectors are defined for any F.

    The nice thing about standard basis vectors is that in F^n, any vector can be represented uniquely in the form

    $$\sum_{i=1}^{n} x_i \vec{\mathbf{e_i}}$$

    This will turn out to be true also in an abstract n-dimensional vector space, but in that case there will be no "standard" basis.

17. Matrices

    An $m \times n$ matrix over a field F is a rectangular array of elements of F with m rows and n columns. Watch the convention: the height is specified first!

    As a mathematical object, any matrix can be multiplied by any element of F. This could be meaningless in the context of an application. Suppose you run a small hospital that has two rooms with three patients in each. Then

    $$\begin{bmatrix} 98.6 & 102.4 & 99.7 \\ 103.2 & 98.3 & 99.6 \end{bmatrix}$$

    is a perfectly reasonable way to keep track of the body temperatures of the patients, but multiplying it by 2.7 seems unreasonable. This matrix, viewed as an element of $\mathbb{R}^6$, is a point, not a vector, but we always use braces for matrices.

    Matrices with the same size and shape can be added component by component. What would you get if you add

    $$\begin{bmatrix} 0.2 & -1.4 & 0.0 \\ 0.6 & -0.9 & 2.35 \end{bmatrix}$$

    to the matrix above to update the temperature data by one day?

18. Matrix multiplication

Matrix multiplication is nicely explained on pages 43-46 of Hubbard. To illustrate the rule, we will take

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & -1 & -2 \end{bmatrix}, \; B = \begin{bmatrix} 0 & 1 \\ 2 & -1 \\ -2 & 0 \end{bmatrix}$$

- Compute AB. $\begin{bmatrix} 0 & 1 \\ 2 & -1 \\ -2 & 0 \end{bmatrix}$

  $\begin{bmatrix} 2 & 1 & 0 \\ 1 & -1 & -2 \end{bmatrix}$

- Compute BA. $\begin{bmatrix} 2 & 1 & 0 \\ 1 & -1 & -2 \end{bmatrix}$

  $\begin{bmatrix} 0 & 1 \\ 2 & -1 \\ -2 & 0 \end{bmatrix}$

In a set of $n \times n$ square matrices, addition and multiplication of matrices are always defined. Multiplication is distributive with respect to addition, too. But because matrix multiplication is noncommutative, the $n \times n$ matrices do not form a field if $n > 1$. (They are said to form a ring.) Let

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \; B = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}$$

Find AB. $\begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}$

$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$

Find BA. $\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$

$\begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}$

19. Matrices as functions:
Since a column vector is also an $n \times 1$ matrix, we can multiply an $m \times n$ matrix by a vector in F^n to get a vector in F^m. The product $A\vec{\mathbf{e_i}}$ is the ith column of A. This is usually the best way to think of a matrix A as representing a *linear function* f: the ith column of A is $f(\vec{\mathbf{e_i}})$.

Example: Suppose that f is linear, $f(\begin{bmatrix}1\\0\end{bmatrix}) = \begin{bmatrix}1\\4\end{bmatrix}$, and $f(\begin{bmatrix}0\\1\end{bmatrix}) = \begin{bmatrix}2\\3\end{bmatrix}$.

What matrix A represents f?

By the definition of matrix multiplication, $A(x_i\vec{\mathbf{e_i}} + x_j\vec{\mathbf{e_j}})$ is the sum of x_i times column i and x_j times column j. So we see that

$$f(x_i\vec{\mathbf{e_i}} + x_j\vec{\mathbf{e_j}}) = x_i f(\vec{\mathbf{e_i}}) + x_j f(\vec{\mathbf{e_j}})$$

This is precisely the requirement for f to be a linear function.

Use matrix multiplication to calculate $f(\begin{bmatrix}2\\-1\end{bmatrix})$.

The rule for forming the product AB can be stated in terms of the rule for a matrix acting on a vector: to form AB, just let A act on each column of B in turn, and put the results side by side to create the matrix AB.

What function does the matrix product AB represent? Consider $(AB)\vec{\mathbf{e_i}}$. This is the ith column of the matrix AB, and it is also the result of letting B act on $\vec{\mathbf{e_i}}$, then letting A act on the result. So for any standard basis vector, the matrix AB represents the composition $A \circ B$ of the functions represented by B and by A.

What about the matrices $(AB)C$ and $A(BC)$? These represent the composition of three functions: say $(f \circ g) \circ h$ and $f \circ (g \circ h)$. But we already know that composition of functions is associative. So we have proved, without any messy algebra, that multiplication of matrices is associative also.

20. Proving associativity by brute force (proof 1.2)

A is an $n \times m$ matrix.

B is an $m \times p$ matrix.

C is an $p \times q$ matrix.

What is the shape of the matrix ABC?

Show how you would lay out the calculation of $(AB)C$.

If $a_{i,j}$ represents the entry in the ith row, jth column of A, then

$$(AB)_{i,k} = \sum_{j=1}^{m} a_{i,j} b_{j,k}$$

$$((AB)C)_{i,q} = \sum_{k=1}^{p} (AB)_{i,k} c_{k,q} = \sum_{j=1}^{m} \sum_{k=1}^{p} (a_{i,j} b_{j,k}) c_{k,q}$$

Show how you would lay out the calculation of $A(BC)$.

$(BC)_{j,q} =$

$(A(BC))_{i,q} =$

On what basis can you now conclude that matrix multiplication is associative for matrices over any field F?

21. Identity matrix:
    It must be square, and the ith column is the ith basis vector. For example,

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

22. Matrices as the arrows for a category $\mathscr{C}$

    Choose a field F, perhaps the real numbers $\mathbb{R}$.

    - An object of $\mathscr{C}$ is a vector space F^n.
    - An arrow of $\mathscr{C}$ is an $n \times m$ matrix A, with domain F^m and codomain F^n.
    - Given $F^p \xrightarrow{B} F^m \xrightarrow{A} F^n$ the composition of arrows A and B is the matrix product AB. Show that the "shape" of the matrices is right for multiplication.

    - The identity arrow for object F^n is the $n \times n$ identity matrix.

    Now we just have to check the two rules that must hold in any category:

    - The associative law for composition of arrows holds because, as we just proved, matrix multiplication is associative.
    - Verify the two identity rules for the case where $A = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 2 & 3 \end{bmatrix}$.

23. Matrix inverses:
Consider first the case of a non-square $m \times n$ matrix A.

If $m > n$, then A takes a vector in $\mathbb{R}^n$ and produces a longer vector in $\mathbb{R}^m$. In general, there will be many matrices B that can recover the original vector in $\mathbb{R}^n$. In the lingo of categories, such a matrix B is a *retraction*.

Here is a matrix that converts a 2-component vector (price of silver and price of gold) into a three-component vector that specifies the price of alloys containing 25%, 50%, and 75% gold respectively. Calculate $\vec{\mathbf{v}} = A \begin{bmatrix} 4 \\ 8 \end{bmatrix}$.

$$A = \begin{bmatrix} .75 & .25 \\ .5 & .5 \\ .25 & .75 \end{bmatrix}, \vec{\mathbf{v}} = A \begin{bmatrix} 4 \\ 8 \end{bmatrix} =$$

By elementary algebra you can reconstruct the price of silver and of gold from the price of any two of the alloys, so it is no surprise to find two different left inverses. Apply each of the following to $\vec{\mathbf{v}}$.

$$B_1 = \begin{bmatrix} 2 & -1 & 0 \\ -2 & 3 & 0 \end{bmatrix}, B_1\vec{\mathbf{v}} =$$

$$B_2 = \begin{bmatrix} 0 & 3 & -2 \\ 0 & -1 & 2 \end{bmatrix}, B_2\vec{\mathbf{v}} =$$

However, in this case there is no right inverse.

If $m < n$, then A takes a vector in $\mathbb{R}^n$ and produces a shorter vector in $\mathbb{R}^m$. In general, there will be no left inverse matrix B that can recover the original vector in $\mathbb{R}^n$, but there may be many different right inverses. Let $A = \begin{bmatrix} 1 & -1 \end{bmatrix}$ and find two different right inverses. In the lingo of categories, such a matrix A is a *section*.

24. Inverting square matrices

For a square matrix, the interesting case is where both a right inverse B and a left inverse C exist. In this case, B and C are equal and they are unique. We can say that "an inverse" A^{-1} exists.

Proof of both uniqueness and equality:
To prove uniqueness of the left inverse matrix, assume that matrix A has two different left inverses C and C' and a right inverse B:

$$C'A = CA = I$$
$$C'(AB) = C(AB) = IB$$
$$C'I = CI = B$$
$$C' = C = B$$

In general, inversion of matrices is best done by "row reduction," discussed in Chapter 2 of Hubbard. For 2×2 matrices there is a simple formula that is worth memorizing:

If

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

then

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

If $ad - bc = 0$ then no inverse exists.

Write down the inverse of $\begin{bmatrix} 3 & 1 \\ 4 & 2 \end{bmatrix}$, where the elements are in $\mathbb{R}$.

The matrix inversion recipe works in any field: try inverting $A = \begin{bmatrix} 3 & 1 \\ 4 & 2 \end{bmatrix}$ where the elements are in $\mathbb{Z}_5$.

25. Other matrix terminology:
All these terms are nicely explained on pp 49-50 of Hubbard.

- transpose
- symmetric matrix
- antisymmetric matrix
- diagonal matrix
- upper or lower triangular matrix

Try applying them to some 3×3 matrices:

$$A = \begin{bmatrix} 3 & 1 & 2 \\ 1 & 2 & 3 \\ 2 & 3 & 4 \end{bmatrix}$$

$$B = \begin{bmatrix} 3 & 0 & 0 \\ 1 & 2 & 0 \\ 2 & 3 & 4 \end{bmatrix}$$

$$C = \begin{bmatrix} 3 & 1 & 2 \\ 0 & 2 & 3 \\ 0 & 0 & 4 \end{bmatrix}$$

$$D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

$$E = \begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & -3 \\ 2 & 3 & 0 \end{bmatrix}$$

26. Linear transformations:

A function $T : F^n \to F^m$ is called *linear* if, for any vectors $\vec{\mathbf{v}}, \vec{\mathbf{w}} \in F^n$ and any scalars $a, b \in F$

$$T(a\vec{\mathbf{v}} + b\vec{\mathbf{w}}) = aT(\vec{\mathbf{v}}) + bT(\vec{\mathbf{w}})$$

Example:

The components of $\vec{\mathbf{v}}$ are the quantities of sugar, flour, and chocolate required to produce a batch of brownies. The components of $\vec{\mathbf{w}}$ are the quantities of these ingredients required to produce a batch of fudge. T is the function that converts such a vector into the total cost of ingredients. T is represented by a matrix $[T]$ (row vector) of prices for the various ingredients.

Write these vectors for the following data:

- A batch of brownies takes 3 pounds of sugar, 6 of flour, 1 of chocolate, while a batch of fudge takes 4 pounds of sugar, 0 of flour, 2 of chocolate.

- Sugar costs \$2 per pound, flour costs \$1 per pound, chocolate costs \$6 per pound.

Then $a\vec{\mathbf{v}} + b\vec{\mathbf{w}}$ is the vector of ingredients required to produce a batches of brownies and b batches of fudge, while $T(\vec{\mathbf{v}})$ is the cost of parts for a single batch of brownies. The statement

$T(a\vec{\mathbf{v}} + b\vec{\mathbf{w}}) = aT(\vec{\mathbf{v}}) + bT(\vec{\mathbf{w}})$ is sound economics.

Two ways to find the cost of 3 batches of brownies plus 2 batches of fudge.

$T(3\vec{\mathbf{v}} + 2\vec{\mathbf{w}}) =$

$3T(\vec{\mathbf{v}}) + 2T(\vec{\mathbf{w}}) =$

Suppose that T produces a 2-component vector of costs from two competing grocers. In that case $[T]$ is a 2×3 matrix.

27. Matrices and linear transformations

Use * to denote the mechanical operation of matrix multiplication.

Any vector can be written as $\vec{\mathbf{v}} = x_1\vec{\mathbf{e_1}} + ... + x_n\vec{\mathbf{e_n}}$.

The rule for multiplying a matrix $[T]$ by a vector $\vec{\mathbf{v}}$ is equivalent to

$$[T] * \vec{\mathbf{v}} = x_1[T] * \vec{\mathbf{e_1}} + ... + x_n[T] * \vec{\mathbf{e_n}} = [T] * (x_1\vec{\mathbf{e_1}} + ... + x_n\vec{\mathbf{e_n}})$$

.

So multiplication by $[T]$ specifies a linear transformation of F^n.

The matrix $[T]$ has columns $[T] * (\vec{\mathbf{e_1}}), ...[T] * (\vec{\mathbf{e_n}})$.

The distinction is subtle. T is a function, a rule. $[T]$ is just a collection of numbers, but the general rule for matrix multiplication turns it into a function.

28. Composition and multiplication:

Suppose $S : F^n \to F^m$ and $T : F^m \to F^p$ are both linear transformations. Then the codomain of S equals the domain of T and we can define the composition $U = T \circ S$.

Prove that U is linear.

To find the matrix of U, we need only determine its action on each standard basis vector.

$$U(\vec{\mathbf{e_i}}) = T(S(\vec{\mathbf{e_i}})) = T([S] * \vec{\mathbf{e_i}}) = [T] * ([S] * \vec{\mathbf{e_i}}) = ([T] * [S]) * \vec{\mathbf{e_i}}$$

So the matrix of $T \circ S$ is $[T] * [S]$.

29. Inversion

A function f is invertible if it is 1-to-1 (injective) and onto (surjective). If g is the inverse of f, then both $g \circ f$ and $f \circ g$ are the identity function. How do we reconcile this observation with the existence of matrices that have one-sided inverses?

Here are two simple examples that identify the problem.

(a) Define f by the formula $f(x) = 2x$. Then

$f : \mathbb{R} \to \mathbb{R}$ is invertible.

$f : \mathbb{Z}_3 \to \mathbb{Z}_3$ is invertible.

$f : \mathbb{Z} \to \mathbb{Z}$ is not invertible.

$f : \mathbb{Z} \to 2\mathbb{Z}$ is invertible. ($2\mathbb{Z}$ is the set of even integers)

In the last case, we have made f invertible by redefining its codomain to equal its image.

(b) If we want to say that the inverse of $f(x) = x^2$ is $g(x) = \sqrt{x}$, we have to redefine $f(x)$ so that its codomain is the nonnegative reals (makes it onto) and its domain is the nonnegative reals (makes it one-to-one).

The codomain of the function that an $m \times n$ matrix represents is all of $\mathbb{R}^m$.

Hubbard p. 64 talks about the invertibility of a linear transformation $T : F^n \to F^m$ and ends up commenting that m and n must be equal. Here is the problem, whose proof will have to wait:

If $m > n$, T cannot be onto, because its image is just a subspace of F^m.

Show how the case where $[T] = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ illustrates the problem.

If $m < n$, T cannot be one-to-one, because there is always a subspace of F^n that gets mapped to the zero vector.

Show how the case where $[T] = \begin{bmatrix} 1 & -1 \end{bmatrix}$ illustrates the problem.

30. Example - constructing the matrix of a linear transformation

    Here is what we know about function f:

    - Its domain and codomain are both $\mathbb{R}^2$.
    - It is linear.
    - $f(\begin{bmatrix}1\\2\end{bmatrix}) = \begin{bmatrix}7\\5\end{bmatrix}$.
    - $f(\begin{bmatrix}1\\4\end{bmatrix}) = \begin{bmatrix}11\\9\end{bmatrix}$.

    Find the matrix T that represents f by using linearity to determine what f does to the standard basis vectors.

    Then automate the calculation by writing down a matrix equation and solving it for T.

31. Invertibility of linear functions and of matrices
(proof 1.3, Hubbard, proposition 1.3.14)

Since the key issue in this proof is the subtle distinction between a linear function T and the matrix $[T]$ that represents it, it is a good idea to use * to denote matrix multiplication and $\circ$ to denote composition of linear transformations.

It is also a good idea to use $\vec{\mathbf{x}}$ for a vector in the domain of T and $\vec{\mathbf{y}}$ for a vector in the codomain of T.

Suppose that linear transformation $T : F^n \to F^m$ is represented by the $m \times n$ matrix $[T]$.

(a) Suppose that the matrix $[T]$ is invertible. Prove that the linear transformation T is one-to-one and onto (injective and surjective), hence invertible.

(b) Suppose that linear transformation T is invertible. Prove that its inverse S is linear and that the matrix of S is $[S] = [T]^{-1}$

The shortest version of this proof starts by exploiting the linearity of T when it is applied to a cleverly-chosen sum of vectors.

$$T(aS(\vec{\mathbf{y_1}}) + bS(\vec{\mathbf{y_2}})) = aT \circ S(\vec{\mathbf{y_1}}) + bT \circ S(\vec{\mathbf{y_2}}).$$

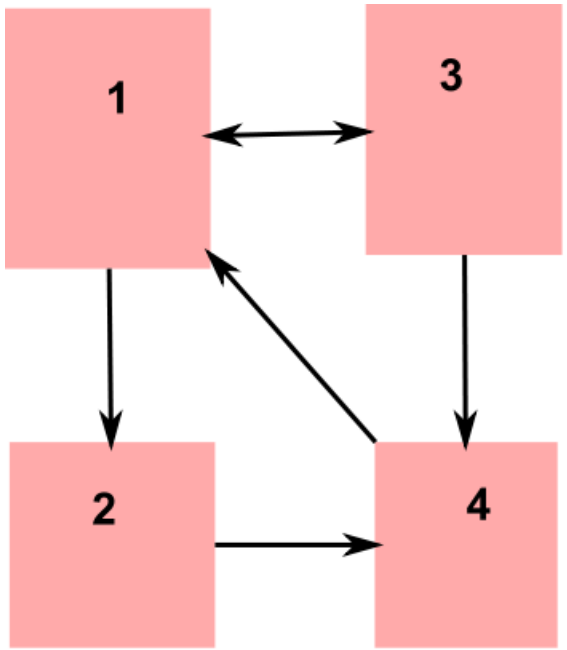
32. Application: graph theory

This is inspired by example 1.2.22 in Hubbard (page 51), but I have extended it by allowing one-way edges and multiple edges.

A graph has n vertices: think of them as islands. Given two vertices V_i and V_j, there may be $A_{i,j}$ edges (ferryboats) that lead from V_j to V_i and $A_{j,i}$ edges that lead from V_i to V_j. If a ferryboat travels in both directions between two islands, it counts twice. In the interest of simplicity, we allow at most one ferryboat between a given pair of islands. The matrix

$$A = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

corresponds to the following directed graph:



The matrix A describes the graph completely.

33. The category of islands and itineraries

The *objects* of the category are the islands.

The *arrows* of the category are itineraries: e.g 4-1 (one step), 1-3-4 (two steps) or 4-1-3-1-2 (four steps). A zero-step itinerary like 1 means "just stay on island 1."

Checking the requirements for a category:

- What are the domain and codomain of the itinerary 4-1-3-1-2?

- If f = 4-1, g = 1-3-4 and h = 4-1-3-1-2 , what is the itinerary $g \circ f$?

  What is $h \circ g$?

- Check the associative law $h \circ (g \circ f) = (h \circ g) \circ f$.

- Check the identity rules for the itinerary 1-3-4.

Now consider a vector $\vec{\mathbf{v}} = \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix}$ whose entries are arbitrary non-negative integers. x_j represents the number of itineraries that end at island j After one more ferryboat ride, the number of itineraries that end at island i is

$$\sum_{j=1}^{n} A_{i,j} x_j.$$

So we see that the vector $A\vec{\mathbf{v}}$ represents the number of itineraries ending at each island after extending the existing list of itineraries by taking one extra ferry ride wherever possible.

If you start on island j and take n ferryboat rides, then the number of itineraries leading to each island is specified by the components of the vector $A^n \vec{\mathbf{e_j}}$.

Hubbard does the example of a cube, where all edges are two-way.

For the four-island graph, with $A = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}$,

use matrix multiplication to find

(a) the number of two-step itineraries from island 1 to island 4.

(b) the number of three-step itineraries from island 1 to island 2.

(c) the number of four-step itineraries from island 3 to island 1.

$$\begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

34. Application: Markov processes

This is inspired by example 1.2.21 in Hubbard, but in my opinion he breaks his own excellent rule by using a "line matrix" to represent probabilities. The formulation below uses a column vector.

Think of a graph where the vertices represent "states" of a random process. A state could, for example, be

(a) A traveler is on a specific island.

(b) Player 1 is serving in a game of badminton.

(c) Hubbard's reference books are on the shelf in the order (2,1,3).

(d) A roulette player has two chips.

(e) During an inning of baseball, there is one man out and runners on first base and third base.

All edges are one way, and attached to each edge is a number in [0,1], the "transition probability" of following that edge in one step of the process. The sum of the probabilities on all the edges leading out of a state cannot exceed 1, and if it is less than 1 there is some probability of remaining in that state.

Examples: write at least one column of the matrix for each case.

(a) If you are on Oahu, the probability of taking a ferry to Maui is 0.2, and the probability of taking a ferry to Lanai is 0.1. Otherwise you stay put.

(b) Badminton: if player 1 serves, the probability of losing the point and the serve is 0.2. If player 2 serves, the probability of losing the point and the serve is 0.3.

(c) If John Hubbard's reference books are on the shelf in the order (2,1,3), the probability that he consults book 3 and places it at the left to make the order (3,2,1) is P_3.

(d) Roulette: after starting with 2 chips and betting a chip on red, the probability of having 3 chips is $\frac{9}{19}$ and the probability of having 1 chip is $\frac{10}{19}$. (in a fair casino, each probability would be $\frac{1}{2}$).

For the badminton example, the transition matrix is

$$A = \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix}.$$

What matrix represents the transition resulting from two successive points?

$$\begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix}$$

$$\begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix}$$

What matrix represents the transition resulting from four successive points?

$$\begin{bmatrix} 0.7 & 0.45 \\ 0.3 & 0.55 \end{bmatrix}$$

$$\begin{bmatrix} 0.7 & 0.45 \\ 0.3 & 0.55 \end{bmatrix}$$

If you raise the transition matrix A to a high power, you might conjecture that after a long time the probability that player 1 is serving is 0.6, no matter who served first.

In support of this conjecture, show that the matrix $A^\infty = \begin{bmatrix} 0.6 & 0.6 \\ 0.4 & 0.4 \end{bmatrix}$ has the property that $AA^\infty = A^\infty$.

3 Seminar Topics

On the course Web site is a link for "Seminar Topic Signup." Usually there will be five topics, but this year in Fortnight 1 there will be four topics for the first week, five for the second. That will leave time for introductions.

Usually, up to three students may sign up for a topic, and one will be chosen at random. This week we are allowing four signups per topic in the first week, in case attendance is large because of shoppers. However, no one may sign up as the third or fourth presenter on a topic until we have at least one presenter for each topic!

You only need to sign up ten times in thirteen weeks; so there is no problem if you join the course late.

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 your own handwritten notes, but be discreet about it. Copying a printout of Paul's scanned lecture notes makes a bad impression. Students who do a good job without any notes quickly get spotted as prospective course assitants for next year.

Topics for Part 1 (Thursday. Sept. 6 - Sunday, Sept. 9)

1. (Proof 1.1)
   Suppose that a and b are two elements of a field F. Using only the axioms for a field, prove the following:

   - $\forall a \in F$, $0a = 0$.
   - If $ab = 0$, then either a or b must be 0.
   - The additive inverse of a is unique.

2. Consider a category whose one and only object is $\mathbb{R}^2$ and whose arrows are vectors in $\mathbb{R}^2$ (more precisely, the arrow is the function "add this vector to the source point.") Show that the following requirements for a category are met (sometimes rather trivially):

   - Every arrow has a source(domain) object and a target(codomain) object.
   - Composition of arrows is associative.
   - For every object, there is an identity arrow that has the required properties.

   This sort of category, with only a single object, is called a "monoid." If every arrow has an inverse, then a monoid is called a "group." Show that this category is a group.

3. (Proof 1.2) A is an $n \times m$ matrix. The entry in row i, column j is $a_{i,j}$.
   B is an $m \times p$ matrix.
   C is an $p \times q$ matrix.
   The entries in these matrices are all from the same field F. Using summation notation, prove that matrix multiplication is associative:
   that $(AB)C = A(BC)$. Include a diagram showing how you would lay out the calculation in each case so the intermediate results do not have to be recopied.

4. Here are the "data" for a category:

   - The objects are vector spaces $\mathbb{R}^n$, one for each $n \geq 1$.
   - The arrows are $m \times n$ matrices with real entries (that means m rows, n columns).
   - Composition of arrows is accomplished by matrix multiplication, which was just proved to be associative.

   Your job:

   - Show that if the codomain for arrow g is the same as the domain for arrow f, then the composition $f \circ g$ is defined.
   - Show that the identity and composition requirements for a category are satisfied.

Topics for Part 2 (Thursday. Sept. 13 - Sunday, Sept. 16)

1. Invent a 2×2 matrix A whose entries are elements of the finite field $\mathbb{Z}_5$ and whose determinant is not 0 or 1. To save time, write 2 instead of $[2]_5$ – everyone will know what you mean. Using the recipe for inverting a 2×2 matrix, construct the matrix A^{-1}, and show that $AA^{-1} = A^{-1}A = I$. If you are clever, you can lay out this calculation so that you only have to copy A^{-1} once.

2. (straight from the textbook)
   Suppose that matrix $F : V \to W$ has a left inverse G and a right inverse H. Prove that G is unique and that $G = H$.
   Let $V = \mathbb{R}^n$ and $W = \mathbb{R}^m$ so that F is an $m \times n$ matrix.
   Show an example where $m = 2, n = 1$, no right inverse exists, and a left inverse is not unique. Then show an example where $m = 1, n = 2$, no left inverse exists and a right inverse is not unique.

3. (a more general statement from category theory, for which the preceding example is a special case.)
   In category $\mathcal{C}$, consider arrow $f : A \to B$. Suppose that f has a "retraction" $g : B \to A$ that undoes its effect so that $g \circ f = I_A$ and also has a "section" (preinverse) h whose effect is undone by f so that $f \circ h = I_B$. Prove that g is unique and that $g = h$. For the case where A and B are finite sets and the arrows are functions, let B have three elements. With the aid of a diagram, show that if A has two elements, no section exists and a retraction is not unique, while if A has four elements, no retraction exists and a section is not unique.

4. (Proof 1.3) Suppose that linear transformation $T : F^n \to F^m$
   is represented by the $m \times n$ matrix $[T]$.

   - Suppose that the matrix $[T]$ is invertible. Prove that the linear transformation T is one-to-one and onto (injective and surjective), hence invertible.
   - Suppose that linear transformation T is invertible. Prove that its inverse S is linear and that the matrix of S is $[S] = [T]^{-1}$

   Note: Use * to denote matrix multiplication and $\circ$ to denote composition of linear transformations. You may take it as already proved that matrix multiplication represents composition of linear transformations. Do not assume that $m = n$.

5. Draw a directed graph like (but not identical to) the one in the notes that shows four islands linked by ferry routes. Write down the matrix A that represents this graph (column specifies origin, row specifies destination). Make a category in which the objects are islands and the arrows are itineraries like 1-3-4-2-3. Show how to compose two arrows, and explain informally how the associativity and identity requirements for a category are met.

   Show how, by calculating one entry in the matrix A^2, you can determine the number of two-step itineraries of the form a-x-b (with a good choice of a and b, you can make the answer be 2). Then explain how, with only four matrix multiplications, you could determine the number of 16-step itineraries that start at island a and end at island b.

4 Workshop Problems

Usually there will be three pairs of problems that can be done without a computer, plus an extra optional pair of problems that require editing R scripts. This last pair will be easy if you have watched the R script videos and downloaded the R scripts onto a laptop that you bring to section. Students enrolled for graduate credit and anyone who wants to use the graduate credit grading option should do one of these. Undergraduates are free to ignore R completely.

Organize the section into groups of three or four students. Preferably either all members of a group should be interested in R, or none should.

Work your problems on the whiteboards. That makes it easy for the section leaders to see how things are going.

Section leaders will assign a number to each group. In the first week
Group 1 does 1a and 2a.
Group 2 does 1b and 2b.
Group 3 does 1a and 2b.
Group 4 does 1b and 2a.
Groups interested in R also do 3a or 3b. Others work whatever remaining problems look interesting, as time permits.

In the second week
Group 1 does 4a and 5a.
Group 2 does 4b and 5b.
Group 3 does 4a and 5b.
Group 4 does 4b and 5a.
Groups interested in R also do 3a or 3b. Others work whatever remaining problems look interesting, as time permits.

Once your group has solved its problems, use a cell phone to take pictures of your solutions, and upload them to the topic box for your section on the Week 1 page of the Web site.

1. Some very short proofs

   (a) Starting with $-1 + 1 = 0$, prove that $(-1)a = -a$ for any $a \in F$. Justify each step of your proof by reference to one or more of the field axioms or to a theorem that was proved in class.

   (b) Prove that $\mathbb{Z}_6$ is not a field. (Either show that it fails to satisfy one of the field axioms, or show that it violates a theorem that is true any any field.)

2. Finite fields

   (a) Make a table of all the powers of [2] in the finite field $\mathbb{Z}_{13}$. Each nonzero field element will appear once, until finally $[2]^{12} = [1]$. Using this table along with the fact that $[2]^m] \times [2]^n = [2]^{m+n}$, find the multiplicative inverse of every element of $\mathbb{Z}_{13}$. Also show how to compute $[11] \times [5]$ by doing addition of exponents. You have constructed a finite table of logarithms!

   (b) Let $A = \begin{bmatrix} [2] & [2] \\ [1] & [0] \end{bmatrix}$, where all entries are in the finite field $\mathbb{Z}_3$. There are only 3^4 different 2×2 matrices with entries, so when you compute powers of A, some power must eventually repeat. If $A^m = A^n$, then $A^{n-m} = I$. By matrix multiplication, find the smallest positive integer p for which $A^p = I$. Then invent a matrix B for which p has a different value.

3. Problems to be solved by writing or editing R scripts

   (a) (similar to script 1.1A, topics 2 and 3)
   Use the `outer()` function of R to make a table of the multiplication facts for $\mathbb{Z}_{17}$ and use it to find the multiplicative inverse of each nonzero element. Then use these inverses to find the result of dividing 11 by 5 and the result of dividing 5 by 11 in this field.

   (b) Let $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$. Use R to calculate A^2, A^4, A^8, and A^{16}. Then use these results to find A^{29}. This is an efficient way to calculate large Fibonacci numbers.

4. Matrices and linear functions

   (a) Here is what we know about the function f:

   - The space it maps from and the space it maps to (the domain and codomain, respectively) are both $\mathbb{R}^2$.
   - It is linear.
   - $f(\begin{bmatrix}1\\1\end{bmatrix}) = \begin{bmatrix}4\\2\end{bmatrix}$
   - $f(\begin{bmatrix}1\\3\end{bmatrix}) = \begin{bmatrix}6\\4\end{bmatrix}$

   i. Find the matrix T that represents f by using linearity to determine what f does to the standard basis vectors.

   ii. Automate the calculation of T by writing down a matrix equation and solving it for T.

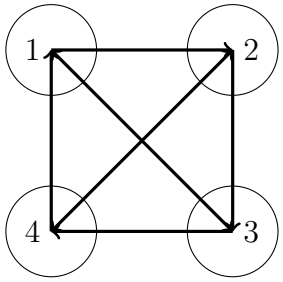
   (b) Suppose that $T : (Z_5)^2 \to (Z_5)^2$ is a linear transformation for which $T\begin{bmatrix}1\\-1\end{bmatrix} = \begin{bmatrix}1\\2\end{bmatrix}, T\begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}3\\0\end{bmatrix}$.

   Construct the matrix $[T]$ that represents T and the matrix $[S]$ that represents T^{-1}.

   Since you are working in a finite field, there are no fractions. Dividing by 2 is the same as multiplying by 3.

5. Applications of matrix multiplication

   (a) Suppose we have four islands connected by ferry routes. The ferries run clockwise around the four islands, and there are also ferries in both directions between 3 and 1 and between 4 and 2.



   Write down the matrix A that represents this graph, and use matrix multiplication to find how many six-step itineraries start at island 1 and end at island 3. If you are efficient, you will need to do only three multiplications.

   (b) The final exam in Humanities 10 always includes either a question about the Iliad or a question about the Odyssey. The choice is made by rolling a die, according to the following rules:

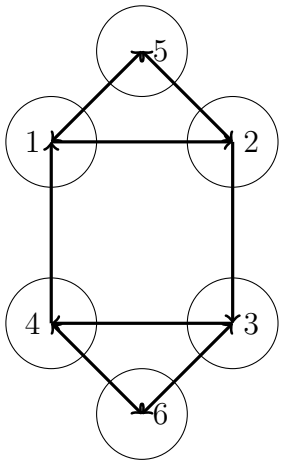
   - If the exam in year n had an Iliad question, the exam in year $n+1$ will have an Iliad question if the die roll is even, an Odyssey question if the die roll is odd.
   - If the exam in year n had an Odyssey question, the exam in year $n+1$ will have an Odyssey question if the die roll is 1, an Iliad question otherwise.

   i. You know that there was an Iliad question in 2014. Use matrix multiplication to calculate the probability of an Iliad question in 2017.

   ii. You are planning that your first child will take Humanities 10 some time in the 2040s. Estimate the probability that there will be an Iliad question on the final exam. (Hint: find a vector of probabilities that does not change from year n to year $n+1$.)

6. Problems to be solved by writing or editing R scripts

   (a) Suppose we have six islands connected by ferry routes. The ferries run clockwise around islands 1-4, and there are also ferries from 2 to 5, 5 to 1, 4 to 6, and 6 to 3



   Write down the matrix A that represents this graph, and use matrix multiplication to find how many nine-step itineraries start at island 5 and end at island 3. If you are efficient, you will need to do only four multiplications.

   (b) (similar to script 1.1D, topic 1)
   You are playing roulette in an American casino, and for any play you may have 0, 1, 2, or 3 chips. When you bet a chip on "odd" you have only an 18/38 chance of winning, because the wheel has 18 odd numbers, 18 even numbers, plus 0 and 00 which count as neither even nor odd.

   - If you have 0 chips you cannot bet and continue to have 0 chips.
   - If you have 1 chip you have probability 9/19 of moving up to 2 chips, probability 10/19 of moving down to 0 chips.
   - If you have 2 chip you have probability 9/19 of moving up to 3 chips, probability 10/19 of moving down to 1 chip.
   - If you have 3 chips you declare victory, do not bet, and continue to have 3 chips.

   Create the 4×4 matrix that represents the effect of one play. Assume that before the first play you are certain to have 2 chips. Use matrix multiplication to determine the probability of your having 0, 1, 2, or 3 chips after 1, 2, 4 and 8 plays. Make a conjecture about the situation after a very large number of plays.

5 Homework

(PROBLEM SET 1 - due on Tuesday, September 18 at 11:59 PM Eastern time)

Problems 1-7 should be done in a single .pdf file and uploaded to the Assignments page of the Web site. The easiest way is to solve the problems on paper and scan them. If you are expert with LaTeX, MS Word, or the Canvas editor, that is a fine alternative. If you take pictures with a smart phone, you must combine the images into a single .pdf file.

Problems 8 and 9 are only for students who are doing the graduate credit work. Both should be done in a single R script and uploaded to the Assignments page of the course Web site.

You will be prepared to do problems 1-3 and 8 after the first week of classes on Sept. 6-9

1. Prove the following, using only the field axioms and the results of workshop problem 1(a).

   (a) The multiplicative inverse a^{-1} of a nonzero element a of a field is unique.

   (b) $(-a)(-b) = ab$.

2. Function composition (Hubbard, exercise 0.4.10.)
   Prove the following:

   (a) Let the functions $f : B \to C$ and $g : A \to B$ be onto. Then the composition $(f \circ g)$ is onto.

   (b) Let the functions $f : B \to C$ and $g : A \to B$ be one-to-one. Then the composition $(f \circ g)$ is one-to-one.

   This problem asks you to prove two results that we will use again and again. All you need to do is to use the definitions of "one-to-one" and "onto."

   Here are some strategies that may be helpful:

   - Exploit the definition:
     If you are told that $f(x)$ is onto, then, for any y in the codomain Y, you can assert the existence of an x such that $f(x) = y$.
     If you are told that $f(x)$ is one-to-one, then, for any a and b such that $f(a) = f(b)$, you can assert that $a = b$.
   - Construct what the definition requires by a procedure that cannot fail:
     To prove that $h(x)$ is onto, describe a procedure for constructing an x such that $h(x) = y$. The proof consists in showing that this procedure works for all y in the codomain Y.

- Prove uniqueness by introducing two names for the same thing:

  To prove that $h(x)$ is one-to-one, give two different names to the same thing: assume that $h(a) = h(b)$, and prove that $a = b$.

3. Hubbard, exercise 1.2.2, parts (a) and (e) only. Do part (a) in the field $\mathbb{R}$, and do part (e) in the field $\mathbb{Z}_7$, where -1 is the same as 6. Check your answer in (e) by doing the calculation in two different orders: according to the associative law these should give the same answer. See Hubbard, figure 1.2.5, for a nice way to organize the calculation.

4. (a) Prove theorem 1.2.17 in Hubbard: that the transpose of a matrix product is the product of the matrices in the opposite order: $(AB)^T = B^T A^T$.

   (b) Let $A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}, B = \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix}$. Calculate AB. Then, using the theorem you just proved, write down the matrix BA without doing any matrix multiplication. (Notice that A and B are symmetric matrices.)

   (c) Prove that if A is any matrix, then $A^T A$ is symmetric.

5. (a) Here is a matrix whose entries are in the finite field $\mathbb{Z}_5$.

$$A = \begin{bmatrix} [1]_5 & [2]_5 \\ [3]_5 & [3]_5 \end{bmatrix}$$

   Write down the inverse of A, using the names $[0]_5 \cdots [4]_5$ for the entries in the matrix. Check your answer by matrix multiplication.

   (b) Count the number of different 2×2 matrices with entries in the finite field $\mathbb{Z}_5$. Of these, how many are invertible? Hint: for invertibility, the left column cannot be zero, and the right column cannot be a multiple of the left column.

6. (a) Hubbard, Exercise 1.3.19, which reads:

   "If A and B are $n \times n$ matrices, their Jordan product is $\frac{AB+BA}{2}$. Show that this product is commutative but not associative."

   Since this problem has an odd number, it is solved in the solutions manual for the textbook. If you cannot resist the temptation to consult the manual, you must cite it as a source!

   (b) Denote the Jordan product of A and B by $A*B$. Prove that it satisfies the distributive law $A * (B + C) = A * B + A * C$.

   (c) Prove that the Jordan product satisfies the special associative law $A * (B * A^2) = (A * B) * A^2$.

7. (a) Suppose that T is linear and that $T\begin{bmatrix}3\\2\end{bmatrix} = \begin{bmatrix}6\\8\end{bmatrix}, T\begin{bmatrix}2\\1\end{bmatrix} = \begin{bmatrix}5\\5\end{bmatrix}$.

   Use the linearity of T to determine $T\begin{bmatrix}1\\0\end{bmatrix}$ and $T\begin{bmatrix}0\\1\end{bmatrix}$, and thereby determine the matrix $[T]$ that represents T. (This brute-force approach works fine in the 2×2 case but not in the $n \times n$ case.)

   (b) Express the given information about T from part (a) in the form $[T][A] = [B]$, and determine the matrix $[T]$ that represents T by using the matrix $[A]^{-1}$. (This approach will work in the general case once you know how to invert an $n \times n$ matrix .)

   The last two problems (graduate credit only) require R scripts. It is fine to copy and edit similar scripts from the course Web site, but it is unacceptable to copy and edit your classmates' scripts!

8. (similar to script 1.1C, topic 5)
   Let $\vec{\mathbf{v1}}$ and $\vec{\mathbf{v2}}$ denote the columns of a 2×2 matrix M. Write an R script that draws a diagram to illustrate the rule for the sign of det M, namely

   - If you have to rotate $\vec{\mathbf{v1}}$ counterclockwise (through less than 180°) to make it line up with $\vec{\mathbf{v2}}$, then det $M > 0$.
   - If you have to rotate $\vec{\mathbf{v1}}$ clockwise (through less than 180°) to make it line up with $\vec{\mathbf{v2}}$, then det $M < 0$.
   - If $\vec{\mathbf{v1}}$ and $\vec{\mathbf{v2}}$ lie on the same line through the origin, then det $M = 0$.

9. (similar to script 1.1D, topic 2)
   Busch Gardens proposes to open a theme park in Beijing, with four regions connected by monorail. From region 1 (the Middle Kingdom), a guest can ride on a two-way monorail to region 2(Tibet), region 3(Shanghai) or region 4(Hunan) or back. Regions 2, 3, and 4 are connected by a one-way monorail that goes from 2 to 3 to 4 and back to 2.

   (a) Draw a diagram to show the four regions and their monorail connections.

   (b) Construct the 4×4 transition matrix A for this graph of four vertices.

   (c) Using matrix multiplication in R, determine how many different sequences of four monorail rides start in Tibet and end in the Middle Kingdom.

1 Major Concepts

1. Matrices and one-sided inverses

   (a) Determine which of these matrices is/are invertible and, where possible, find their inverses. For the invertible matrix/matrices, confirm that the inverse you calculate is both a left and a right inverse.

   $$A = \begin{bmatrix} 4 & 2 \\ 6 & 3 \end{bmatrix}, B = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, C = \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 2 \end{bmatrix}$$

   (b) Matrix D is not invertible, but it does have a one-sided inverse. Which does it have—a left inverse or a right inverse? Find a left/right inverse, and explain why an other-side inverse cannot exist.

   $$D = \begin{bmatrix} 4 & 3 \end{bmatrix}$$

   (c) Similarly, matrix E is not invertible, but it does have a one-sided inverse. Which does it have—a left inverse or a right inverse? Find a left/right inverse, and explain why an other-side inverse cannot exist.

   $$E = \begin{bmatrix} 7 \\ 2 \\ 4 \end{bmatrix}$$

2. Categories, retractions, and sections: Consider a category in which our objects are finite sets A and B and we have an arrow $f : A \to B$.

   (a) Draw an example A, B, and f where A has more elements than B and f has neither a section or a retraction.

   (b) Draw an example A, B, and f where A has more elements than B and f has a section. Can f have a retraction?

   (c) Draw an example A, B, and f where A has fewer elements than B and f has neither a section or a retraction.

   (d) Draw an example A, B, and f where A has fewer elements than B and f has a retraction. Can f have a section?

3. Summation notation

   (a) If $f(i)$ is some expression involving i, then

   $$\sum_{i=1}^{n} f(i) = f(1) + f(2) + \cdots + f(n)$$

   (b) Calculate $\sum_{i=0}^{3} \frac{i^2+1}{i+1}$ and $\sum_{i=1}^{3} \left(\sum_{j=0}^{i} j \right)$.

   (c) Given that matrix A is 2×5, matrix B is 5×3, and matrix C is 3×4, give an expression in summation notation for $(ABC)_{2,3}$. Do this by considering $(AB)C$.

4. Markov processes

   You are playing a board game where you can move your piece to a red, yellow, or blue square, at each round. Here are the rules for the game:

   - If you are currently on a red square, you toss a coin, if you end up with head you move to a yellow square, if you end up with tail, you move to another red square.
   - If you are on a yellow square, you roll a die. If you get 1, you move to a red square. If you get 2,3,or,4 you move to another yellow square. If you get 5 or 6, you move to a blue square.
   - If you are currently on a blue square, you roll a die. If the number you get is even, you stay on blue.If it is odd, you move to a red square.

   (a) Write down the Markov matrix that represents the probabilities for moving from one square to another.

   (b) If you start on a yellow square, after 4 rounds of playing, what is the probability that your piece moves to a red square?

5. More field axioms: Prove $(-1)(-1) = 1$.

6. Matrices as functions: Here's an example for the derivative function as a matrix: The domain and co-domain have bases :x^0, x^1, x^2, x^3
   With this convention, the elements of the matrix are the coefficients of each corresponding basis element. With the 4*4 matrix below, we can differentiate all polynomials of degree four or less.

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

   For instance, the polynomial $1 + 2x + 3x^2 + 4x^3$can be represented by a column vector, with the entries being the coefficients for the powers of x:

$$v = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

   Now,we need to differentiate our polynomial:
   • In the language of functions: the derivative function, takes in our polynomial, and acts on it.
   • In the language of matrices: matrix A acts on vector b. (We should multiply matrix A by vector b).

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 & 6 & 12 & 0 \end{bmatrix}$$

   We can read the resulting vector as the polynomial:

$$2 + 6x + 12x^2$$

   Our matrix has indeed accomplished differentiation!

7. Hints for question 2 of homework:
   - For part a), pay attention to the definition of surjectivity:
   $\forall c \in C, \ \exists b \in B \ \ s.t. \ \ f(b) = c$
   -Similarly, for part b), pay attention to the definition of injectivity:
   if $h(a_i) = h(a_j)$ then, $a_i = a_j$
   Here, it is easier to use the contrapositive of the statement, so we want to show that if $a_i \neq a_j$ then, $h(a_i) \neq h(a_j)$.

MATHEMATICS 23a/E-23a, Fall 2018
Linear Algebra and Real Analysis I
Week 2 (Dot and Cross Products, Euclidean Geometry of $\mathbb{R}^n$)

Authors: Paul Bamberg and Kate Penner
R scripts by Paul Bamberg
Last modified: August 13, 2018 by Paul Bamberg

Reading

- Hubbard, section 1.4

Recorded Lectures

- Lecture 4 (Week 2, Class 1) (watch on September 18 or 19)
- Lecture 5 (Week 2, Class 2) (watch on September 20 or 21)

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

- 2.1 Given vectors $\vec{\mathbf{v}}$ and $\vec{\mathbf{w}}$ in Euclidean $\mathbb{R}^n$, prove that $|\vec{\mathbf{v}} \cdot \vec{\mathbf{w}}| \leq |\vec{\mathbf{v}}||\vec{\mathbf{w}}|$ (Cauchy-Schwarz) and that $|\vec{\mathbf{v}} + \vec{\mathbf{w}}| \leq |\vec{\mathbf{v}}| + |\vec{\mathbf{w}}|$ (triangle inequality). Use the distributive law for the scalar product and the fact that no vector has negative length.

  (The standard version of this proof is in the textbook. An alternative is in sections 1.3 and 1.4 of the Executive Summary.)

- 2.2 For a 3×3 matrix A, define $\det(A)$ in terms of the cross and dot products of the columns of the matrix. Then, using the definition of matrix multiplication and the linearity of the dot and cross products, prove that $\det(AB) = \det(A)\det(B)$.

R Scripts Scripts labeled A, B, ... are closely tied to the Executive Summary. Scripts labeled X, Y, ... are interesting examples. There is a narrated version on the Web site. Scripts labeled L are library scripts that you may wish to include in your own scripts.

- Script 1.2A-LengthDotAngle.R
  Topic 1 - Length, Dot Product, Angles
  Topic 2 - Components of a vector
  Topic 3 - Angles in Pythagorean triangles
  Topic 4 - Vector calculation using components

- Script 1.2B-RotateReflect.R
  Topic 1 - Rotation matrices
  Topic 2 - Reflection matrices

- Script 1.2C-ComplexConformal.R
  Topic 1 - Complex numbers in R
  Topic 2 - Representing complex numbers by 2x2 matrices

- Script 1.2D-CrossProduct.R
  Topic 1 - Algebraic properties of the cross product
  Topic 2 - Geometric properties of the cross product
  Topic 3 - Using cross products to invert a 3x3 matrix

- Script 1.2E-DeterminantProduct.R
  Topic 1 - Product of 2x2 matrices
  Topic 2 - Product of 3x3 matrices

- Script 1.2L-VectorLibrary.R
  Topic 1 - Some useful angles and basis vectors
  Topic 2 - Functions for working with angles in degrees

- Script 1.2X-Triangle.R
  Topic 1 - Generating and displaying a randomly generated triangle
  Topic 2 - Checking some formulas of trigonometry

- Script 1.2Y-Angles3D.R
  Topic 1 - Angles between vectors in $\mathbb{R}^3$
  Topic 2 - Angles and distances in a cube
  Topic 3 - Calculating the airline mileage between cities

1 Executive Summary

1.1 The dot product

The dot product of two vectors in $\mathbb{R}^n$ is $\vec{\mathbf{x}} \cdot \vec{\mathbf{y}} = x_1y_1 + x_2y_2 + \ldots + x_ny_n = \sum_{i=1}^{n} x_iy_i$

- It requires two vectors and returns a scalar.
- It is commutative and it is distributive with respect to addition.
- In $\mathbb{R}^2$ or $\mathbb{R}^3$, the dot product of a vector with itself (a concept of algebra) is equal to the square of its length (a concept of geometry):

$$\vec{\mathbf{x}} \cdot \vec{\mathbf{x}} = |\vec{\mathbf{x}}|^2$$

- Taking the dot product with any standard basis vector $\vec{\mathbf{e_i}}$ extracts the corresponding component:

$$\vec{\mathbf{x}} \cdot \vec{\mathbf{e_i}} = x_i$$

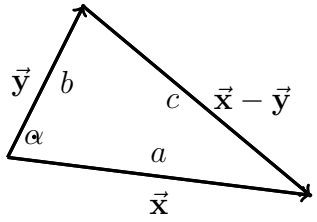
- Taking the dot product with any unit vector $\vec{\mathbf{a}}$ (not necessarily a basis vector) extracts the component of $\vec{\mathbf{x}}$ along $\vec{\mathbf{a}}$:

$$\vec{\mathbf{x}} \cdot \vec{\mathbf{a}} = x_a$$

  This means that the difference $\vec{\mathbf{x}} - x_a\vec{\mathbf{a}}$ is orthogonal to $\vec{\mathbf{a}}$.

1.2 Dot products and angles

We have the law of cosines, usually written $c^2 = a^2 + b^2 - 2ab\cos\alpha$.



Consider the triangle whose sides lie along the vectors $\vec{\mathbf{x}}$(length a), $\vec{\mathbf{y}}$ (length b), and $\vec{\mathbf{x}} - \vec{\mathbf{y}}$ (length c). Let α denote the angle between the vectors $\vec{\mathbf{x}}$ and $\vec{\mathbf{y}}$.

By the distributive law,

$$(\vec{\mathbf{x}} - \vec{\mathbf{y}}) \cdot (\vec{\mathbf{x}} - \vec{\mathbf{y}}) = \vec{\mathbf{x}} \cdot \vec{\mathbf{x}} + \vec{\mathbf{y}} \cdot \vec{\mathbf{y}} - 2\vec{\mathbf{x}} \cdot \vec{\mathbf{y}} \implies c^2 = a^2 + b^2 - 2\vec{\mathbf{x}} \cdot \vec{\mathbf{y}}$$

Comparing with the law of cosines, we find that angles and dot products are related by:

$$\vec{\mathbf{x}} \cdot \vec{\mathbf{y}} = ab\cos\alpha = |\vec{\mathbf{x}}||\vec{\mathbf{y}}|\cos\alpha$$

1.3 Cauchy-Schwarz inequality

The dot product provides a way to extend the definition of length and angle for vectors to $\mathbb{R}^n$, but now we can no longer invoke Euclidean plane geometry to guarantee that $|\cos\alpha| \leq 1$.

We need to show that for any vectors $\vec{\mathbf{v}}$ and $\vec{\mathbf{w}}$ in $\mathbb{R}^n$

$$|\vec{\mathbf{v}} \cdot \vec{\mathbf{w}}| \leq |\vec{\mathbf{v}}||\vec{\mathbf{w}}|$$

This is generally known as the "Cauchy-Schwarz inequality."

For a short proof of the Cauchy-Schwarz inequality, make $\vec{\mathbf{v}}$ and $\vec{\mathbf{w}}$ into unit vectors and form their sum and difference.

$$(\frac{\vec{\mathbf{v}}}{|\vec{\mathbf{v}}|} \pm \frac{\vec{\mathbf{w}}}{|\vec{\mathbf{w}}|}) \cdot (\frac{\vec{\mathbf{v}}}{|\vec{\mathbf{v}}|} \pm \frac{\vec{\mathbf{w}}}{|\vec{\mathbf{w}}|}) \geq 0$$

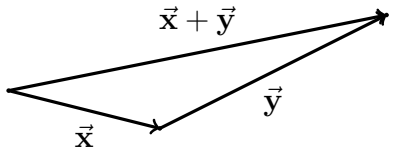
$$1 + 1 \pm 2\frac{\vec{\mathbf{v}} \cdot \vec{\mathbf{w}}}{|\vec{\mathbf{v}}||\vec{\mathbf{w}}|} \geq 0, \text{ and by algebra } |\frac{\vec{\mathbf{v}} \cdot \vec{\mathbf{w}}}{|\vec{\mathbf{v}}||\vec{\mathbf{w}}|}| \leq 1$$

We now have a useful definition of angle for vectors in $\mathbb{R}^n$ in general:

$$\alpha = \arccos \frac{\vec{\mathbf{v}} \cdot \vec{\mathbf{w}}}{|\vec{\mathbf{v}}||\vec{\mathbf{w}}|}$$

1.4 The triangle inequality

If $\vec{\mathbf{x}}$ and $\vec{\mathbf{y}}$, placed head-to-tail, determine two sides of a triangle, the third side coincides with the vector $\vec{\mathbf{x}} + \vec{\mathbf{y}}$.



We need to show that its length cannot exceed the sum of the lengths of the other two sides:

$$|\vec{\mathbf{x}} + \vec{\mathbf{y}}| \leq |\vec{\mathbf{x}}| + |\vec{\mathbf{y}}|$$

The proof uses the distributive law for the dot product.

$$|\vec{\mathbf{x}} + \vec{\mathbf{y}}|^2 = (\vec{\mathbf{x}} + \vec{\mathbf{y}}) \cdot (\vec{\mathbf{x}} + \vec{\mathbf{y}}) = (\vec{\mathbf{x}} + \vec{\mathbf{y}}) \cdot \vec{\mathbf{x}} + (\vec{\mathbf{x}} + \vec{\mathbf{y}}) \cdot \vec{\mathbf{y}}$$

Applying Cauchy-Schwarz to each term on the right-hand side, we have:

$$|\vec{\mathbf{x}} + \vec{\mathbf{y}}|^2 \leq |\vec{\mathbf{x}} + \vec{\mathbf{y}}||\vec{\mathbf{x}}| + |\vec{\mathbf{x}} + \vec{\mathbf{y}}||\vec{\mathbf{y}}|$$

In the special case where $|\vec{\mathbf{x}} + \vec{\mathbf{y}}| = 0$ the inequality is clearly true. Otherwise we can divide by the common factor of $|\vec{\mathbf{x}} + \vec{\mathbf{y}}|$ to complete the proof.

1.5 Isometries of $\mathbb{R}^2$

A linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ is completely specified by its effect on the basis vectors $\vec{\mathbf{e}}_1$ and $\vec{\mathbf{e}}_2$. These vectors are the two columns of the matrix that represents T. If you know what a transformation is supposed to do to each basis vector, you can simply use this information to fill out the necessary columns of its matrix representation.

Of special interest are **isometries**: transformations that preserve the distance between any pair of points, and hence the length of any vector.

Since dot products can be expressed in terms of lengths, it follows that any isometry also preserves dot products.

So the transformation T is an isometry if and only if for any pair of vectors:

$$T\vec{\mathbf{a}} \cdot T\vec{\mathbf{b}} = \vec{\mathbf{a}} \cdot \vec{\mathbf{b}}$$

For the matrix associated with an isometry, both columns must be unit vectors and their dot product is zero.

Two isometries:

- A rotation, $R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$, with $\det R = +1$.
- A reflection, $F(\theta) = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}$, with $\det F = -1$.

Matrix R represents a counterclockwise rotation through angle θ about the origin. Matrix F represents reflection in a line through the origin that makes an angle θ with the first standard basis vector.

There are many other isometries of Euclidean geometry; translations, or rotations about points other than the origin. However, these do not hold the origin fixed, and so they are not linear transformations and cannot represented by 2×2 matrices.

Since the composition of isometries is an isometry, the product of any number of matrices of this type is another rotation or reflection. Remember that composition is a series of transformations acting on a vector in a specific order that must be preserved during multiplication.

1.6 Matrices and algebra: complex numbers

The same field axioms we reviewed on the first day apply here to the complex numbers, notated $\mathbb{C}$.

The real and imaginary parts of a complex number can be used as the two components of a vector in $\mathbb{R}^2$. The rule for addition of complex numbers is the same as the rule for addition of vectors in $\mathbb{R}^2$ (in that they are to be kept separate from each other), and the modulus of a complex number is the same as the length of the vector that represents it. So the triangle inequality applies for complex numbers: $|z_1 + z_2| \leq |z_1| + |z_2|$.

This property extends to vector spaces over complex numbers.

1.7 What about complex multiplication?

The geometrical interpretation of multiplication by a complex number $z = a + ib = re^{i\theta}$ is multiplication of the modulus by r combined with addition of θ to the angle with the x-axis.
This is precisely the geometrical effect of the linear transformation represented by the matrix

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} r\cos\theta & -r\sin\theta \\ r\sin\theta & r\cos\theta \end{bmatrix}$$

Such a matrix is the product of the constant matrix $\begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}$ and the rotation matrix $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$.
It is called a **conformal matrix** and it preserves angles even though it does not preserve lengths.

1.8 Complex numbers as a field of matrices

In general, matrices do not form a field because multiplication is not commutative. There are two notable exceptions: $n \times n$ matrices that are multiples of the identity matrix and 2×2 conformal matrices. Since multiples of the identity matrix and rotations all commute, the product of two conformal matrices $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ and $\begin{bmatrix} c & -d \\ d & c \end{bmatrix}$ is the same in either order.

1.9 The cross product

$$\vec{\mathbf{a}} \times \vec{\mathbf{b}} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \times \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{bmatrix}$$

Properties

1. $\vec{\mathbf{a}} \times \vec{\mathbf{b}} = -\vec{\mathbf{b}} \times \vec{\mathbf{a}}$.
2. $\vec{\mathbf{a}} \times \vec{\mathbf{a}} = 0$.
3. For fixed $\vec{\mathbf{a}}$, $\vec{\mathbf{a}} \times \vec{\mathbf{b}}$ is a linear function of $\vec{\mathbf{b}}$, and vice versa.
4. For the standard basis vectors, $\vec{\mathbf{e_i}} \times \vec{\mathbf{e_j}} = \vec{\mathbf{e_k}}$ if i, j and k are in cyclic increasing order (123, 231, or 312). Otherwise $\vec{\mathbf{e_i}} \times \vec{\mathbf{e_j}} = -\vec{\mathbf{e_k}}$.
5. $\vec{\mathbf{a}} \times \vec{\mathbf{b}} \cdot \vec{\mathbf{c}} = \vec{\mathbf{a}} \cdot \vec{\mathbf{b}} \times \vec{\mathbf{c}}$. This quantity is also the determinant of the matrix whose columns are $\vec{\mathbf{a}}$, $\vec{\mathbf{b}}$, and $\vec{\mathbf{c}}$.
6. $(\vec{\mathbf{a}} \times \vec{\mathbf{b}}) \times \vec{\mathbf{c}} = (\vec{\mathbf{a}} \cdot \vec{\mathbf{c}})\vec{\mathbf{b}} - (\vec{\mathbf{b}} \cdot \vec{\mathbf{c}})\vec{\mathbf{a}}$
7. $\vec{\mathbf{a}} \times \vec{\mathbf{b}}$ is orthogonal to the plane spanned by $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$.
8. $|\vec{\mathbf{a}} \times \vec{\mathbf{b}}|^2 = |\vec{\mathbf{a}}|^2 |\vec{\mathbf{b}}|^2 - (\vec{\mathbf{a}} \cdot \vec{\mathbf{b}})^2$
9. The length of $\vec{\mathbf{a}} \times \vec{\mathbf{b}}$ is $|\vec{\mathbf{a}}||\vec{\mathbf{b}}| \sin \alpha$.
10. The length of $\vec{\mathbf{a}} \times \vec{\mathbf{b}}$ is equal to the area of the parallelogram spanned by $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$.

1.10 Cross product and determinants

If a 3×3 matrix A has columns $\vec{\mathbf{a}}_1$, $\vec{\mathbf{a}}_2$, and $\vec{\mathbf{a}}_3$, then its determinant $\det(A) = \vec{\mathbf{a}}_1 \times \vec{\mathbf{a}}_2 \cdot \vec{\mathbf{a}}_3$.

1. $\det(A)$ changes sign if you interchange any two columns. (easiest to prove for columns 1 and 2, but true for any pair)
2. $\det(A)$ is a linear function of each column (easiest to prove for column 3, but true for any column)
3. For the identity matrix I, $\det(I) = 1$.

The magnitude of $\vec{\mathbf{a}} \times \vec{\mathbf{b}} \cdot \vec{\mathbf{c}}$ is equal to the volume of the parallelepiped spanned by $\vec{\mathbf{a}}$, $\vec{\mathbf{b}}$ and $\vec{\mathbf{c}}$.
If $C = AB$, then $\det(C) = \det(A) \det(B)$

2 Lecture Outline

1. The dot product:

   This is defined for vectors in $\mathbb{R}^n$ as

   $$\vec{\mathbf{x}} \cdot \vec{\mathbf{y}} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

   It has the following properties. The proof of the first four (omitted) is brute-force computation.

   - Commutative law:
     $$\vec{\mathbf{x}} \cdot \vec{\mathbf{y}} = \vec{\mathbf{y}} \cdot \vec{\mathbf{x}}$$
   - Distributive law:
     $$\vec{\mathbf{x}} \cdot (\vec{\mathbf{y_1}} + \vec{\mathbf{y_2}}) = \vec{\mathbf{x}} \cdot \vec{\mathbf{y_1}} + \vec{\mathbf{x}} \cdot \vec{\mathbf{y_2}}$$
   - For Euclidean geometry, in $\mathbb{R}^2$ or $\mathbb{R}^3$, the dot product of a vector with itself (defined by algebra) is equal to the square of its length (a physically meaningful quantity).
   - Taking the dot product with any standard basis vector $\vec{\mathbf{e_i}}$ extracts the corresponding component:
     $$\vec{\mathbf{x}} \cdot \vec{\mathbf{e_i}} = x_i$$
   - Taking the dot product with any unit vector $\vec{\mathbf{a}}$ (not necessarily a standard basis vector) extracts the component of $\vec{\mathbf{x}}$ along $\vec{\mathbf{a}}$:
     $$\vec{\mathbf{x}} \cdot \vec{\mathbf{a}} = x_a$$
     This means that the difference $\vec{\mathbf{x}} - x_a \vec{\mathbf{a}}$ is orthogonal to $\vec{\mathbf{a}}$.

     Proof: Orthogonality of two vectors means that their dot product is zero. So to show orthogonality, evaluate
     $$(\vec{\mathbf{x}} - (\vec{\mathbf{x}} \cdot \vec{\mathbf{a}})\vec{\mathbf{a}}) \cdot \vec{\mathbf{a}}.$$

2. Dot products and angles

From elementary trigonometry we have the law of cosines, usually written $c^2 = a^2 + b^2 - 2ab\cos\alpha$.

In this formula, c denotes the length of the side opposite angle α. Just in case you forgot the proof, let's review it.

Angles and dot products are related by the formula

$\vec{\mathbf{x}} \cdot \vec{\mathbf{y}} = |\vec{\mathbf{x}}||\vec{\mathbf{y}}| \cos\alpha$

Proof (Hubbard, page 69):

Consider the triangle whose sides lie along the vectors $\vec{\mathbf{x}}, \vec{\mathbf{y}}$, and $\vec{\mathbf{x}} - \vec{\mathbf{y}}$, and let α denote the angle between the vectors $\vec{\mathbf{x}}$ and $\vec{\mathbf{y}}$.

$$c^2 = (\vec{\mathbf{x}} - \vec{\mathbf{y}}) \cdot (\vec{\mathbf{x}} - \vec{\mathbf{y}}).$$

Expand the dot product using the distributive law, and you can identify one of the terms as $2ab\cos\alpha$.

3. Cauchy-Schwarz inequality

The dot product provides a way to extend the definition of length and angle to vectors in $\mathbb{R}^n$, but now we can no longer invoke Euclidean plane geometry to guarantee that $|\cos\alpha| \leq 1$.

We need to show that for any vectors $\vec{\mathbf{v}}$ and $\vec{\mathbf{w}}$ in $\mathbb{R}^n$,

$$|\vec{\mathbf{v}} \cdot \vec{\mathbf{w}}| \leq |\vec{\mathbf{v}}||\vec{\mathbf{w}}|$$

This is generally known as the "Cauchy-Schwarz inequality." Hubbard points out that it was first published by Bunyakovsky. This fact illustrates Stigler's Law of Eponymy:
"No law, theorem, or discovery is named after its originator."
The law applies to itself, since long before Stigler formulated it, A. N. Whitehead noted that,

"Everything of importance has been said before, by someone who did not discover it."

The best-known proof of the Cauchy-Schwarz inequality incorporates two useful strategies.

- No vector has negative length.
- Discriminant of quadratic equation.

Define a quadratic function of the real variable t by

$$f(t) = |t\vec{\mathbf{v}} - \vec{\mathbf{w}}|^2 = (t\vec{\mathbf{v}} - \vec{\mathbf{w}}) \cdot (t\vec{\mathbf{v}} - \vec{\mathbf{w}})$$

Since $f(t)$ is the square of a length of a vector, it cannot be negative, so the quadratic equation $f(t) = 0$ does not have two real roots.

But by the quadratic formula, if the equation $at^2 + bt + c = 0$ does not have two real roots, its discriminant $b^2 - 4ac$ is not positive.

Complete the proof by writing out $b^2 - 4ac \leq 0$ for quadratic function $f(t)$.

So we have a useful definition of angle for vectors in $\mathbb{R}^n$ in general:

$$\alpha = \arccos \frac{\vec{\mathbf{v}} \cdot \vec{\mathbf{w}}}{|\vec{\mathbf{v}}||\vec{\mathbf{w}}|}$$

The function $\arccos(x)$ can be computed on your electronic calculator by summing an infinite series. It is guaranteed to return a value between 0 and π.

Example: In $\mathbb{R}^4$, what is the angle between vectors $\begin{bmatrix} 1 \\ 2 \\ 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \\ 1 \\ 2 \end{bmatrix}$?

4. The triangle inequality (second part of proof 2.1)

   If $\vec{\mathbf{x}}$ and $\vec{\mathbf{y}}$, placed head-to-tail, determine two sides of a triangle, the third side coincides with the vector $\vec{\mathbf{x}}+\vec{\mathbf{y}}$. We need to show that its length cannot exceed the sum of the lengths of the other two sides:

   $|\vec{\mathbf{x}} + \vec{\mathbf{y}}| \leq |\vec{\mathbf{x}}| + |\vec{\mathbf{y}}|$

   The proof uses the distributive law for the dot product and the Cauchy-Schwarz inequality.

   Express $|\vec{\mathbf{x}} + \vec{\mathbf{y}}|^2$ as a dot product:

   Apply the distributive law:

   Use Cauchy-Schwarz to get an inequality for lengths:

   Take the square root of both sides:

5. Proof 2.1 – start to finish, done in a slightly different way

   Given vectors $\vec{\mathbf{v}}$ and $\vec{\mathbf{w}}$ in Euclidean $\mathbb{R}^n$, prove that $|\vec{\mathbf{v}}\cdot\vec{\mathbf{w}}| \leq |\vec{\mathbf{v}}||\vec{\mathbf{w}}|$ (Cauchy-Schwarz) and that $|\vec{\mathbf{v}} + \vec{\mathbf{w}}| \leq |\vec{\mathbf{v}}| + |\vec{\mathbf{w}}|$ (triangle inequality). Use the distributive law for the scalar product and the fact that no vector has negative length.

6. Isometries of $\mathbb{R}^2$.

   A linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ is completely specified by its effect on the basis vectors $\vec{\mathbf{e}}_1$ and $\vec{\mathbf{e}}_2$. These vectors are the two columns of the matrix that represents T.

   Of special interest are "isometries:" transformations that preserve the distance between any pair of points, and hence the length of any vector.

   Since

$$4\vec{\mathbf{a}} \cdot \vec{\mathbf{b}} = |\vec{\mathbf{a}} + \vec{\mathbf{b}}|^2 - |\vec{\mathbf{a}} - \vec{\mathbf{b}}|^2,$$

   dot products can be expressed in terms of lengths, and any isometry also preserves dot products.

   Prove this useful identity.

   So T is an isometry if and only if

   $T\vec{\mathbf{a}} \cdot T\vec{\mathbf{b}} = \vec{\mathbf{a}} \cdot \vec{\mathbf{b}}$ for any pair of vectors.

   This means that the first column of T must be a unit vector, which can be written without any loss of generality as

$$\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}.$$

   The second column must also be a unit vector, and its dot product with the first column must be zero. So there are only two possibilities:

   - A rotation,

$$R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

     which has $\det R = 1$.

   - A reflection,

$$F(\theta) = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix},$$

     which has $\det F = -1$.

     This represents reflection in a line through the origin that makes an angle θ with the first basis vector.

   Since the composition of isometries is an isometry, the product of any number of matrices of this type is another rotation or reflection.

7. Using matrices to represent rotations and reflections

   (a) Use matrix multiplication to show that if a counterclockwise rotation though angle β is followed by a counterclockwise rotation though angle α, the net effect is a counterclockwise rotation though angle $\alpha + \beta$. (The proof requires some trig identities that you can rederive, if you ever forget them, by doing this calculation.)

   (b) Confirm, both by geometry and by matrix multiplication, that if you reflect a point P first in the line $y = 0$, then in the line $y = x$, the net effect is to rotate the point counterclockwise through $90°$.

8. Complex numbers as vectors and as matrices

The field axioms that you learned last week apply also to the complex numbers, notated $\mathbb{C}$.

The real and imaginary parts of a complex number can be used as the two components of a vector in $\mathbb{R}^2$. The rule for addition of complex numbers is the same as the rule for addition of vectors in $\mathbb{R}^2$, and the modulus of a complex number is the same as the length of the vector that represents it. So the triangle inequality applies for complex numbers: $|z_1+z_2| \leq |z_1|+|z_2|$.

This property extends to vector spaces over complex numbers.

The geometrical interpretation of multiplication by a complex number $z = a + ib = re^{i\theta}$ is multiplication of the modulus by r combined with addition of θ to the angle with the x-axis.
This is precisely the geometrical effect of the linear transformation represented by the matrix

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} r\cos\theta & -r\sin\theta \\ r\sin\theta & r\cos\theta \end{bmatrix}$$

Such a matrix is the product of the constant matrix $\begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}$ and the rotation matrix $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$.
It is called a **conformal matrix** and it preserves angles even though it does not preserve lengths.

Example: Compute the product of the complex numbers $2+i$ and $3+i$ by useing matrix multiplication.

9. Complex numbers as a field of matrices

In general, matrices do not form a field because multiplication is not commutative. There are two notable exceptions: $n \times n$ matrices that are multiples of the identity matrix and 2×2 conformal matrices. Since multiples of the identity matrix and rotations all commute, the product of two conformal matrices $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ and $\begin{bmatrix} c & -d \\ d & c \end{bmatrix}$ is the same in either order.

10. Cross products:

At this point it is inappropriate to try to define the determinant of an $n \times n$ matrix. For $n = 3$, however, anything that can be done with determinants can also be done with cross products, which are peculiar to $\mathbb{R}^3$. So we will start with cross products:

Definition:

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \times \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{bmatrix}$$

Since this is a computational definition, the way to prove the following properties is by brute-force computation.

(a) $\vec{\mathbf{a}} \times \vec{\mathbf{b}} = -\vec{\mathbf{b}} \times \vec{\mathbf{a}}$.

(b) $\vec{\mathbf{a}} \times \vec{\mathbf{a}} = \vec{\mathbf{0}}$.

(c) For fixed $\vec{\mathbf{b}}$, $\vec{\mathbf{a}} \times \vec{\mathbf{b}}$ is a linear function of $\vec{\mathbf{b}}$, and vice versa.

(d) For the standard basis vectors, $\vec{\mathbf{e_i}} \times \vec{\mathbf{e_j}} = \vec{\mathbf{e_k}}$ if i, j and k are in cyclic increasing order (123, 231, or 312). Otherwise $\vec{\mathbf{e_i}} \times \vec{\mathbf{e_j}} = -\vec{\mathbf{e_k}}$.
You may find it easiest to calculate cross products in general as
$(a_1\vec{\mathbf{e_1}} + a_2\vec{\mathbf{e_2}} + a_3\vec{\mathbf{e_3}}) \times (b_1\vec{\mathbf{e_1}} + b_2\vec{\mathbf{e_2}} + b_3\vec{\mathbf{e_3}})$,
using the formula for the cross products of basis vectors. Try this approach for
$\vec{\mathbf{a}} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$, $\vec{\mathbf{b}} = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$.

(e) $\vec{\mathbf{a}} \times \vec{\mathbf{b}} \cdot \vec{\mathbf{c}} = \vec{\mathbf{a}} \cdot \vec{\mathbf{b}} \times \vec{\mathbf{c}}$. No parentheses are necessary, because the operations only make sense if the cross product is done first. This quantity is also the determinant of the matrix whose columns are $\vec{\mathbf{a}}$, $\vec{\mathbf{b}}$, and $\vec{\mathbf{c}}$.

(f) $(\vec{\mathbf{a}} \times \vec{\mathbf{b}}) \times \vec{\mathbf{c}} = (\vec{\mathbf{a}} \cdot \vec{\mathbf{c}})\vec{\mathbf{b}} - (\vec{\mathbf{b}} \cdot \vec{\mathbf{c}})\vec{\mathbf{a}}$
Physicists, memorize this formula ! The vector in the middle gets the plus sign.

11. Geometric properties of the cross product:

    We can now prove these without messy calculations involving components. Justify each step, using properties of the dot product and properties (a) through (f) from the preceding page.

    - $\vec{\mathbf{a}} \times \vec{\mathbf{b}}$ is orthogonal to the plane spanned by $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$.
      Proof: Let $\vec{\mathbf{v}} = s\vec{\mathbf{a}} + t\vec{\mathbf{b}}$ be a vector in this plane. Then

      $$\vec{\mathbf{v}} \cdot \vec{\mathbf{a}} \times \vec{\mathbf{b}} = s\vec{\mathbf{a}} \cdot \vec{\mathbf{a}} \times \vec{\mathbf{b}} + t\vec{\mathbf{b}} \cdot \vec{\mathbf{a}} \times \vec{\mathbf{b}}$$

      $$\vec{\mathbf{v}} \cdot \vec{\mathbf{a}} \times \vec{\mathbf{b}} = s\vec{\mathbf{a}} \cdot \vec{\mathbf{a}} \times \vec{\mathbf{b}} - t\vec{\mathbf{b}} \cdot \vec{\mathbf{b}} \times \vec{\mathbf{a}}$$

      $$\vec{\mathbf{v}} \cdot \vec{\mathbf{a}} \times \vec{\mathbf{b}} = s\vec{\mathbf{a}} \times \vec{\mathbf{a}} \cdot \vec{\mathbf{b}} - t\vec{\mathbf{b}} \times \vec{\mathbf{b}} \cdot \vec{\mathbf{a}}$$

      $$\vec{\mathbf{v}} \cdot \vec{\mathbf{a}} \times \vec{\mathbf{b}} = 0 - 0 = 0.$$

    - $|\vec{\mathbf{a}} \times \vec{\mathbf{b}}|^2 = |\vec{\mathbf{a}}|^2|\vec{\mathbf{b}}|^2 - (\vec{\mathbf{a}} \cdot \vec{\mathbf{b}})^2$
      Proof:

      $$|\vec{\mathbf{a}} \times \vec{\mathbf{b}}|^2 = (\vec{\mathbf{a}} \times \vec{\mathbf{b}}) \cdot (\vec{\mathbf{a}} \times \vec{\mathbf{b}})$$

      $$|\vec{\mathbf{a}} \times \vec{\mathbf{b}}|^2 = ((\vec{\mathbf{a}} \times \vec{\mathbf{b}}) \times \vec{\mathbf{a}}) \cdot \vec{\mathbf{b}}$$

      $$|\vec{\mathbf{a}} \times \vec{\mathbf{b}}|^2 = ((\vec{\mathbf{a}} \cdot \vec{\mathbf{a}})\vec{\mathbf{b}} - (\vec{\mathbf{a}} \cdot \vec{\mathbf{b}})\vec{\mathbf{a}}) \cdot \vec{\mathbf{b}}$$

      $$|\vec{\mathbf{a}} \times \vec{\mathbf{b}}|^2 = (\vec{\mathbf{a}} \cdot \vec{\mathbf{a}})(\vec{\mathbf{b}} \cdot \vec{\mathbf{b}}) - (\vec{\mathbf{a}} \cdot \vec{\mathbf{b}})(\vec{\mathbf{a}} \cdot \vec{\mathbf{b}})$$

      $$|\vec{\mathbf{a}} \times \vec{\mathbf{b}}|^2 = |\vec{\mathbf{a}}|^2|\vec{\mathbf{b}}|^2 - (\vec{\mathbf{a}} \cdot \vec{\mathbf{b}})^2$$

    - The length of $\vec{\mathbf{a}} \times \vec{\mathbf{b}}$ is $|\vec{\mathbf{a}}||\vec{\mathbf{b}}| \sin \alpha$.
      Proof:

      $$|\vec{\mathbf{a}} \times \vec{\mathbf{b}}|^2 = |\vec{\mathbf{a}}|^2|\vec{\mathbf{b}}|^2(1 - \cos^2 \alpha) = |\vec{\mathbf{a}}|^2|\vec{\mathbf{b}}|^2(\sin^2 \alpha)$$

    - The length of $\vec{\mathbf{a}} \times \vec{\mathbf{b}}$ is equal to the area of the parallelogram spanned by $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$.
      Proof: $|\vec{\mathbf{a}}|$ is the base of the parallelogram and $|\vec{\mathbf{b}}| \sin \alpha$ is its height. Draw a diagram to illustrate this property.

12. Cross products and determinants.

You should be familiar with 2×2 and 3×3 determinants from high-school algebra. The general definition of the determinant, to be introduced in the spring term, underlies the general technique for calculating volumes in $\mathbb{R}^n$ and will be used to define differential forms.

If a 2×2 matrix A has columns $\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ and $\begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$, then its determinant $\det(A) = a_1 b_2 - a_2 b_1$.

Equivalently,

$$\begin{bmatrix} a_1 \\ a_2 \\ 0 \end{bmatrix} \times \begin{bmatrix} b_1 \\ b_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \det \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \end{bmatrix}$$

You can think of the determinant as a function of the entire matrix A or as a function of its two columns.

Matrix A maps the unit square, spanned by the two standard basis vectors, into a parallelogram whose area is $|\det(A)|$.

Let's prove this for the case where all the entries of A are positive and $\det(A) > 0$. The area of the parallelogram formed by the columns of A is twice the area of the triangle that has these columns as two of its sides. The area of this triangle can be calculated in terms of elementary formulas for areas of rectangles and right triangles.

13. Determinants in $\mathbb{R}^3$

Here is our definition:

If a 3×3 matrix A has columns $\vec{\mathbf{a}}_1$, $\vec{\mathbf{a}}_2$, and $\vec{\mathbf{a}}_3$, then its determinant $\det(A) = \vec{\mathbf{a}}_1 \times \vec{\mathbf{a}}_2 \cdot \vec{\mathbf{a}}_3$.

Apply this definition to the matrix

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 2 \\ 0 & 1 & 0 \end{bmatrix}.$$

Check the following properties of the definition.

(a) $\det(A)$ changes sign if you interchange any two columns. (easiest to prove for columns 1 and 2, but true for any pair)

(b) $\det(A)$ is a linear function of each column (easiest to prove for column 3, but true for any column)

(c) For the identity matrix I, $\det(I) = 1$.

14. Determinants, triple products, and geometry

The magnitude of $\vec{\mathbf{a}} \times \vec{\mathbf{b}} \cdot \vec{\mathbf{c}}$ is equal to the volume of the parallelepiped spanned by $\vec{\mathbf{a}}$, $\vec{\mathbf{b}}$ and $\vec{\mathbf{c}}$.

Proof: $\vec{\mathbf{a}} \times \vec{\mathbf{b}}$ is the area of the base of the parallelepiped, and $|\vec{\mathbf{c}}| \cos \alpha$, where α is the angle between $\vec{\mathbf{c}}$ and the direction orthogonal to the base, is its height.

Matrix A maps the unit cube, spanned by the three basis vectors, into a parallelepiped whose volume is $|\det(A)|$. You can think of $|\det(A)|$ as a "volume stretching factor." This interpretation will underly much of the theory for change of variables in multiple integrals, a major topic in the spring term.

If three vectors in $\mathbb{R}^3$ all lie in the same plane, the cross product of any two of them, which is orthogonal to that plane, is orthogonal to the third vector, so $\vec{\mathbf{v}}_1 \times \vec{\mathbf{v}}_2 \cdot \vec{\mathbf{v}}_3 = 0$.

Apply this test to $\vec{\mathbf{v}}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \vec{\mathbf{v}}_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \vec{\mathbf{v}}_3 = \begin{bmatrix} 3 \\ 2 \\ 2 \end{bmatrix}$.

If four points in $\mathbb{R}^3$ all lie in the same plane, the vectors that join any one of the points to each of the other three points all lie in that plane. Apply this test to $\mathbf{p} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \mathbf{q} = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}, \mathbf{r} = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}, \mathbf{s} = \begin{pmatrix} 4 \\ 3 \\ 3 \end{pmatrix}$.

15. Calculating angles and areas

Let $\vec{\mathbf{v}}_1 = \begin{bmatrix} -2 \\ 2 \\ -1 \end{bmatrix}, \vec{\mathbf{v}}_2 = \begin{bmatrix} -4 \\ 1 \\ 1 \end{bmatrix}$.

Both these vectors happen to be perpendicular to the vector $\vec{\mathbf{v}}_3 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$.

(a) Determine the angle between $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$.

(b) Determine the volume of the parallelepiped spanned by $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2,$ and $\vec{\mathbf{v}}_3$, and thereby determine the area of the parallelogram spanned by $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$.

16. Determinants and matrix multiplication

If $C = AB$, then $\det(C) = \det(A)\det(B)$

This useful result is easily proved by brute force for 2×2 matrices, and a brute-force proof in Mathematica would be valid for 3×3 matrices. Here is a proof that relies on properties of the cross product.

Recall that each column of a matrix is the image of a standard basis vector.

Consider the first column of the matrix $C = AB$, and exploit the fact that A is linear.

$$\vec{\mathbf{c}}_1 = A\vec{\mathbf{b}}_1 = A(\sum_{i=1}^{3} b_{i,1}\vec{\mathbf{e}}_i) = \sum_{i=1}^{3} b_{i,1}A(\vec{\mathbf{e}}_i)) = \sum_{i=1}^{3} b_{i,1}\vec{\mathbf{a}}_i.$$

The same is true of the second and third columns.

Now consider $\det C = \vec{\mathbf{c_1}} \times \vec{\mathbf{c_2}} \cdot \vec{\mathbf{c_3}}$.

$$\det C = (\sum_{i=1}^{3} b_{i,1}\vec{\mathbf{a_i}}) \times (\sum_{j=1}^{3} b_{j,2}\vec{\mathbf{a_j}}) \cdot (\sum_{k=1}^{3} b_{k,3}\vec{\mathbf{a_k}})$$

Now use the distributive law for dot and cross products.

$$\det C = \sum_{i=1}^{3} b_{i,1} \sum_{j=1}^{3} b_{j,2} \sum_{k=1}^{3} b_{k,3}(\vec{\mathbf{a_i}} \times \vec{\mathbf{a_j}} \cdot \vec{\mathbf{a_k}})$$

There are 27 terms in this sum, but all but six of them involve two subscripts that are equal, and these are zero because a triple product with two equal vectors is zero.

The six that are not zero all involve $\vec{\mathbf{a}}_1 \times \vec{\mathbf{a}}_2 \cdot \vec{\mathbf{a}}_3$, three with a plus sign and three with a minus sign. So

$\det C = f(B)(\vec{\mathbf{a}}_1 \times \vec{\mathbf{a}}_2 \cdot \vec{\mathbf{a}}_3) = f(B)\det(A)$, where $f(B)$ is some messy function of products of all the entries of B.

This formula is valid for any A. In particular, it is valid when A is the identity matrix, $C = B$, and $\det(A) = 1$.

So $\det B = f(B)\det(I) = f(B)$

and the messy function is the determinant!

17. Proof 2.2 – start to finish

    For a 3×3 matrix A, define $\det(A)$ in terms of the cross and dot products of the columns of the matrix. Then, using the definition of matrix multiplication and the linearity of the dot and cross products, prove that $\det(AB) = \det(A)\det(B)$.

18. Isometries of $\mathbb{R}^2$.

A linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ is completely specified by its effect on the basis vectors $\vec{\mathbf{e}}_1$ and $\vec{\mathbf{e}}_2$. These vectors are the two columns of the matrix that represents T.

Of special interest are "isometries:" transformations that preserve the distance between any pair of points, and hence the length of any vector.

Since

$$4\vec{\mathbf{a}} \cdot \vec{\mathbf{b}} = |\vec{\mathbf{a}} + \vec{\mathbf{b}}|^2 - |\vec{\mathbf{a}} - \vec{\mathbf{b}}|^2,$$

dot products can be expressed in terms of lengths, and any isometry also preserves dot products.

Prove this useful identity.

So T is an isometry if and only if

$T\vec{\mathbf{a}} \cdot T\vec{\mathbf{b}} = \vec{\mathbf{a}} \cdot \vec{\mathbf{b}}$ for any pair of vectors.

This means that the first column of T must be a unit vector, which can be written without any loss of generality as

$$\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}.$$

The second column must also be a unit vector, and its dot product with the first column must be zero. So there are only two possibilities:

- A rotation,

$$R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

  which has $\det R = 1$.

- A reflection,

$$F(\theta) = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix},$$

  which has $\det F = -1$.

  This represents reflection in a line through the origin that makes an angle θ with the first basis vector.

Since the composition of isometries is an isometry, the product of any number of matrices of this type is another rotation or reflection.

19. Transposes and dot products

Start by proving in general that $(AB)^T = B^T A^T$. This is a statement about matrices, and you have to prove it by brute force.

The dot product of vectors $\vec{\mathbf{v}}$ and $\vec{\mathbf{w}}$ can also be written in terms of matrix multiplication as

$$\vec{\mathbf{v}} \cdot \vec{\mathbf{w}} = \vec{\mathbf{v}}^T \vec{\mathbf{w}}$$

where we think of $\vec{\mathbf{v}}^T$ as a $1 \times m$ matrix and think of $\vec{\mathbf{w}}$ as an $m \times 1$ matrix. The product is a 1×1 matrix, so it equals its own transpose.

Prove that $\vec{\mathbf{v}} \cdot A\vec{\mathbf{w}} = A^T\vec{\mathbf{v}} \cdot \vec{\mathbf{w}}$. This theorem lets you move a matrix from one factor in a dot product to the other, as long as you replace it by its transpose.

20. Orthogonal matrices

If a matrix R represents an isometry, then each column is a unit vector and the columns are orthogonal. Since the columns of R are the rows of R^T we can express this property as

$R^T R = I$

Perhaps a nicer way to express this condition for a matrix to represent an isometry is $R^T = R^{-1}$. Check that this is true for the 2×2 matrices that represent rotations and reflections.

For a rotation matrix

$$R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

For a reflection matrix

$$F(\theta) = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}.$$

21. Isometries and cross products

Many vectors of physical importance (torque, angular momentum, magnetic field) are defined as cross products, so it is useful to know what happens to a cross product when an isometry is applied to each vector in the product.

Consider the matrix whose columns are $R\vec{\mathbf{u}}, R\vec{\mathbf{v}}$, and $\vec{\mathbf{w}}$.

Multiply this matrix by R^T to get a matrix whose columns are

$R^T R\vec{\mathbf{u}}, R^T R\vec{\mathbf{v}}$, and $R^T\vec{\mathbf{w}}$. In the process you multiply the determinant by $\det(R^T) = \det(R)$.

Now, since $R^T R = I$ for an isometry, $\vec{\mathbf{u}} \times \vec{\mathbf{v}} \cdot R^T\vec{\mathbf{w}} = \det(R)R\vec{\mathbf{u}} \times R\vec{\mathbf{v}} \cdot \vec{\mathbf{w}}$

Equivalently, $R(\vec{\mathbf{u}} \times \vec{\mathbf{v}}) \cdot \vec{\mathbf{w}} = \det(R)R\vec{\mathbf{u}} \times R\vec{\mathbf{v}} \cdot \vec{\mathbf{w}}$.

Since this is true for any $\vec{\mathbf{w}}$, in particular for any basis vector, it follows that

$R(\vec{\mathbf{u}} \times \vec{\mathbf{v}}) = \det(R)R\vec{\mathbf{u}} \times R\vec{\mathbf{v}}$

If R is a rotation, then $\det(R) = 1$ and $R(\vec{\mathbf{u}} \times \vec{\mathbf{v}}) = R\vec{\mathbf{u}} \times R\vec{\mathbf{v}}$

If R is a reflection, then $\det(R) = -1$ and $R(\vec{\mathbf{u}} \times \vec{\mathbf{v}}) = -R\vec{\mathbf{u}} \times R\vec{\mathbf{v}}$

This is reasonable. Suppose you are watching a physicist in a mirror as she calculates the cross product of two vectors. You see her apparently using a left-hand rule and think that she has got the sign of the cross-product wrong.

22. Using cross products to invert a 3×3 matrix

Thinking about transposes also leads to a formula for the inverse of a 3×3 matrix in terms of cross products. Suppose that matrix A has columns $\vec{\mathbf{a}}_1, \vec{\mathbf{a}}_2$, and $\vec{\mathbf{a}}_3$. Form the vector $\vec{\mathbf{s}}_1 = \vec{\mathbf{a}}_2 \times \vec{\mathbf{a}}_3$.

This is orthogonal to $\vec{\mathbf{a}}_2$ and $\vec{\mathbf{a}}_3$, and its dot product with $\vec{\mathbf{a}}_1$ is $\det(A)$.

Similarly, the vector $\vec{\mathbf{s}}_2 = \vec{\mathbf{a}}_3 \times \vec{\mathbf{a}}_1$

is orthogonal to $\vec{\mathbf{a}}_3$ and $\vec{\mathbf{a}}_1$, and its dot product with $\vec{\mathbf{a}}_2$ is $\det(A)$.

Finally, the vector $\vec{\mathbf{s}}_3 = \vec{\mathbf{a}}_1 \times \vec{\mathbf{a}}_2$

is orthogonal to $\vec{\mathbf{a}}_1$ and $\vec{\mathbf{a}}_2$, and its dot product with $\vec{\mathbf{a}}_3$ is $\det(A)$.

So if you form these vectors into a matrix S and take its transpose,

$S^T A = \det(A) I$.

If $\det A = 0$, A has no inverse. Otherwise

$$A^{-1} = \frac{S^T}{\det(A)}.$$

You may have learned this rule in high-school algebra in terms of 2×2 determinants.

Summarize the proof that this recipe is correct.

.

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 2.1) Given vectors $\vec{\mathbf{v}}$ and $\vec{\mathbf{w}}$ in Euclidean $\mathbb{R}^n$, prove that $|\vec{\mathbf{v}} \cdot \vec{\mathbf{w}}| \leq |\vec{\mathbf{v}}||\vec{\mathbf{w}}|$ (Cauchy-Schwarz) and that $|\vec{\mathbf{v}} + \vec{\mathbf{w}}| \leq |\vec{\mathbf{v}}| + |\vec{\mathbf{w}}|$ (triangle inequality). Use the distributive law for the scalar product and the fact that no vector has negative length.

   (The standard version of this proof is in the textbook. An alternative is in sections 1.3 and 1.4 of the Executive Summary.)

2. Let $F(\alpha)$ be the 2×2 matrix that represents reflection in an upward-sloping line through the origin that makes an angle α with the positive horizontal axis.
   Let $R(\theta)$ be the 2×2 matrix that represents rotation counterclockwise about the origin through an angle θ.

   Prove by matrix multiplication that $F(\beta)F(\alpha) = R(2(\beta - \alpha))$. You will need to use the trigonometric identities

   $$\sin(x + y) = \sin x \cos y + \cos x \sin y;\ \cos(x + y) = \cos x \cos y - \sin x \sin y,$$

   which every scientist should memorize!

   Optional: As you apply $F(\alpha)$ and $F(\beta)$ successively to a vector $\vec{\mathbf{v}}$, the tip of the vector moves along a circular arc. If you draw a diagram with $\beta > \alpha > 0$, it is not hard also to prove your result geometrically.

3. Using the formulas
$$\vec{\mathbf{a}} \times \vec{\mathbf{b}} \cdot \vec{\mathbf{c}} = \vec{\mathbf{a}} \cdot \vec{\mathbf{b}} \times \vec{\mathbf{c}}$$
and
$$(\vec{\mathbf{a}} \times \vec{\mathbf{b}}) \times \vec{\mathbf{c}} = (\vec{\mathbf{a}} \cdot \vec{\mathbf{c}})\vec{\mathbf{b}} - (\vec{\mathbf{b}} \cdot \vec{\mathbf{c}})\vec{\mathbf{a}},$$
which every physicist should memorize,
prove that $|\vec{\mathbf{a}} \times \vec{\mathbf{b}}|$ is equal to the area of a parallelogram in $\mathbb{R}^3$ spanned by vectors $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$, with angle α between them, namely $|\vec{\mathbf{a}}||\vec{\mathbf{b}}||\sin\alpha|$.

4. (Proof 2.2) For a 3×3 matrix A, define $\det(A)$ in terms of the cross and dot products of the columns of the matrix. Then, using the definition of matrix multiplication and the linearity of the dot and cross products, prove that $\det(AB) = \det(A)\det(B)$.

5. Given the rule that the transpose of a matrix product is the product of the transposes in reverse order, $(AB)^T = B^T A^T$ (if there is lots of time left, you may want to prove this rule), prove that
$$\vec{\mathbf{v}} \cdot A\vec{\mathbf{w}} = A^T\vec{\mathbf{v}} \cdot \vec{\mathbf{w}}.$$

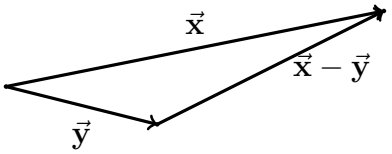
6. (Extra topic) A field isomorphism:
Let f be the function that converts the complex number $c = a + bi$ into the conformal matrix $C = f(c) = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ and the complex number $z = x + yi$ into the conformal matrix $Z = f(z) = \begin{bmatrix} x & -y \\ y & x \end{bmatrix}$.

   - Prove that $f(c + z) = f(c) + f(z)$.
   - Prove that $f(cz) = f(c)f(z)$, where the multiplication on the left is standard multiplication of complex numbers and the multiplication on the right is matrix multiplication.

4 Workshop Problems

1. Proofs that use dot products

   (a) A triangle is formed by using vectors $\vec{\mathbf{x}}$ and $\vec{\mathbf{y}}$, both anchored at one vertex. The vectors are labeled so that the longer one is called $\vec{\mathbf{x}}$: i.e. $|\vec{\mathbf{x}}| > |\vec{\mathbf{y}}|$. The vector $\vec{\mathbf{x}} - \vec{\mathbf{y}}$ then lies along the third side of the triangle. Prove that
   $|\vec{\mathbf{x}} - \vec{\mathbf{y}}| \geq |\vec{\mathbf{x}}| - |\vec{\mathbf{y}}|$.



   (b) A parallelogram has sides with lengths a and b. Its diagonals have lengths c and d, Prove the "parallelogram law," which states that

$$c^2 + d^2 = 2(a^2 + b^2).$$

2. Applying the dot product to parallelograms

   (a) A parallelogram is spanned by two vectors that meet at a 60 degree angle, one of which is twice as long as the other. Find the ratio of the lengths of the diagonals and the cosine of the acute angle between the diagonals. Confirm that the parallelogram law holds in this case.

   (b) Consider a parallelogram spanned by vectors $\vec{\mathbf{v}}$ and $\vec{\mathbf{w}}$. Using the dot product, prove that it is a rhombus if and only if the diagonals are perpendicular and that it is a rectangle if and only if the diagonals are equal in length.

3. Proofs and applications that involve cross products

   (a) Prove that the cross product is not associative but that it satisfies the "Jacobi identity"

$$(\vec{\mathbf{a}} \times \vec{\mathbf{b}}) \times \vec{\mathbf{c}} + (\vec{\mathbf{b}} \times \vec{\mathbf{c}}) \times \vec{\mathbf{a}} + (\vec{\mathbf{c}} \times \vec{\mathbf{a}}) \times \vec{\mathbf{b}} = 0.$$

   (b) Consider a parallelepiped whose base is a parallelogram spanned by two unit vectors, anchored at the origin, with a 60 degree angle between them. The third side leaving the origin, also a unit vector, makes a 60 degree angle with each of the other two sides, so that each face is made of of a pair of equilateral triangles. Using dot and cross products, show that the angle α between the third side and a line that bisects the angle between the other two sides satisfies $\cos\alpha = 1/\sqrt{3}$ and that the volume of this parallelepiped is $\frac{1}{\sqrt{2}}$.

4. Problems that involve writing or editing R scripts
   Script 1.2A will be helpful for both of these. The library script 2L has functions for dealing with angles in degrees.

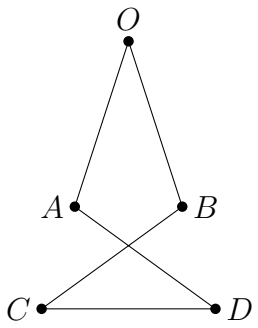
   (a) Construct a triangle where vector AB has length 5 and is directed east, while vector AC has length 10 and is directed 53 degrees north of east. On side BC, construct point D that is 1/3 of the way from B to C. Using dot products, confirm that the vector AD bisects the angle at A. Draw a diagram that illustrates this result.
   This is a special case of Euclid's Elements, Book VI, Proposition 3.

   (b) You are playing golf, and the hole is located 350 yards from the tee in a direction 18 degrees south of east. You hit a tee shot that travels 220 yards 14 degrees south of east, followed by an iron shot that travels 150 yards 23 degrees south of east. How far from the hole is your golf ball now located? Draw a diagram that illustrates this result.

5 Homework, due at 11:59 pm on Sept. 25

In working on these problems, you may collaborate with classmates and consult books and general online references. If, however, you encounter a posted solution to one of the problems, do not look at it, and email Paul, who will try to get it removed.

1. One way to construct a regular pentagon



Take five ball-point pens or other objects of equal length(call it 1) and arrange them symmetrically, as shown in the diagram above, so that O, A, C and O, B, D are collinear and $|OC| = |OD|$. Let $AO = \vec{\mathbf{v}}, |BO| = |\vec{\mathbf{v}}|$, $CD = \vec{\mathbf{w}}$, $CA = x\vec{\mathbf{v}}, |DB| = x|\vec{\mathbf{v}}|$.

   (a) Express vectors AD and OB in terms of x, $\vec{\mathbf{v}}$, and $\vec{\mathbf{w}}$. By using the fact that these vectors have the same length 1 as $\vec{\mathbf{v}}$ and $\vec{\mathbf{w}}$, get two equations relating x and $\vec{\mathbf{v}} \cdot \vec{\mathbf{w}}$. (Use the distributive law for the dot product).

   (b) Eliminate x to find a quadratic equation satisfied by $\vec{\mathbf{v}} \cdot \vec{\mathbf{w}}$. Show that the angle α between $\vec{\mathbf{v}}$ and $\vec{\mathbf{w}}$ satisfies the equation $\sin 3\alpha = -\sin 2\alpha$ and that therefore $\alpha = \frac{2\pi}{5}$. (In case you have forgotten, $\sin 3\alpha = \sin\alpha(4\cos^2\alpha - 1)$).

   (c) Explain how, given five identical ball-point pens, you can construct a regular pentagon. (Amazingly, the obvious generalization with seven pens lets you construct a regular heptagon. Crockett Johnson claims to have discovered this fact while dining with friends in a restaurant in Italy in 1975, using a menu, a wine list, and seven toothpicks)

2. One vertex of a quadrilateral in $\mathbb{R}^3$ is located at point $\mathbf{p}$. The other three vertices, going around in order, are located at $\mathbf{q} = \mathbf{p} + \vec{\mathbf{a}}$, $\mathbf{r} = \mathbf{p} + \vec{\mathbf{b}}$, and $\mathbf{s} = \mathbf{p} + \vec{\mathbf{c}}$.

   (a) Invent an expression involving cross products that is equal to zero if and only if the four vertices of the quadrilateral lie in a plane.

   (b) Prove that the midpoints of the four sides $\mathbf{pq}, \mathbf{qr}, \mathbf{rs}$, and $\mathbf{sp}$ are the vertices of a parallelogram.

3. Isometries and dot products

   The transpose of a (column) vector $\vec{\mathbf{v}}$ is a "row vector" $\vec{\mathbf{v}}^T$, which is also a $1 \times n$ matrix.

   Suppose that $\vec{\mathbf{v}}$ and $\vec{\mathbf{w}}$ are vectors in $\mathbb{R}^n$ and A is an $n \times n$ matrix.

   (a) Prove that $\vec{\mathbf{v}} \cdot A\vec{\mathbf{w}} = \vec{\mathbf{v}}^T A\vec{\mathbf{w}}$. (You can think of the right-hand side as the product of three matrices.)

   (b) Prove that $\vec{\mathbf{v}} \cdot A\vec{\mathbf{w}} = A^T\vec{\mathbf{v}} \cdot \vec{\mathbf{w}}$. You can do this by brute force using summation notation, or you can do it by using part (a) and the rule for the transpose of a matrix product (Therem 1.2.17 in Hubbard).

   (c) Now suppose that $\vec{\mathbf{v}}$ and $\vec{\mathbf{w}}$ are vectors in $\mathbb{R}^3$ and R is an 3×3 isometry matrix. Prove that $R\vec{\mathbf{v}} \cdot R\vec{\mathbf{w}} = \vec{\mathbf{v}} \cdot \vec{\mathbf{w}}$. If you believe that physical laws should remain valid when you rotate your epcerimental apparatus, this result shows that dot products are appropriate to use in expressing physical laws.

4. Using vectors to prove theorems of trigonometry.

   (a) For vectors $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$,
   $|\vec{\mathbf{a}} \times \vec{\mathbf{b}}| = |\vec{\mathbf{a}}||\vec{\mathbf{b}}| \sin\alpha$, where α is the angle between the vectors.
   By applying this formula to a triangle whose sides are $\vec{\mathbf{v}}, \vec{\mathbf{w}}$, and $\vec{\mathbf{v}} - \vec{\mathbf{w}}$, prove the Law of Sines.

   (b) Consider a parallelogram spanned by vectors $\vec{\mathbf{v}}$ and $\vec{\mathbf{w}}$.
   Its diagonal is $\vec{\mathbf{v}} + \vec{\mathbf{w}}$.

   Let α denote the angle between $\vec{\mathbf{v}}$ and the diagonal ; let β denote the angle between $\vec{\mathbf{w}}$ and the diagonal. By expressing sines and cosines in terms of cross products, dot products, and lengths of vectors, prove the addition formula
   $\sin(\alpha + \beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta$.

5. Let $R(\theta)$ denote the 2×2 matrix that represents a counterclockwise rotation about the origin through angle θ. Let $F(\alpha)$ denote the 2×2 matrix that represents a reflection in the line through the origin that makes angle α with the x axis. Using matrix multiplication and the trigonometric identities

   $\sin(\alpha + \beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta$
   $\cos(\alpha + \beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta$, prove the following:

   (a) $F(\beta)F(\alpha) = R(2(\beta - \alpha))$.

   (b) $F(\gamma)F(\beta)F(\alpha) = F(\gamma + \alpha - \beta)$. (If you are doing R, you might want to work problem 7 first.)

   (c) The product of any even number of reflections in lines through the origin is a rotation about the origin and the product of any odd number of reflections in lines through the origin is a reflection in a line through the origin. (Hint: use induction. First establish the base cases $n = 1$ and $n = 2$. Then do the "inductive step:" show that if the result is true for the product of n reflections, it is true for $n + 2$ reflections.)

6. Matrices that represent complex numbers

   (a) Confirm that $i^2 = -1$ using conformal matrices.

   (b) Represent $4 + 2i$ as a matrix. Square it and interpret its result as a complex number. Confirm your answer by checking what you get when expanding algebraically.

   (c) Show that using matrices to represent complex numbers still preserves addition as we would expect.

   That is, write two complex numbers as matrices. Then add the matrices, and interpret the sum as a complex number. Confirm your answer is correct algebraically.

The last two problems require R scripts. Feel free to copy and edit existing scripts and to use the library script 2L, which has functions for dealing with angles in degrees.

7. Vectors in two dimensions

   (a) You are playing golf and have made a good tee shot. Now the hole is located only 30 yards from your ball, in a direction 32 degrees north of east. You hit a chip shot that travels 25 yards 22 degrees north of east, followed by a putt that travels 8 yards 60 degrees north of east. How far from the hole is your golf ball now located? For full credit, include a diagram showing the relevant vectors.

   (b) The three-reflections theorem, whose proof was problem 5b, states that if you reflect successively in lines that make angle α, β, and γ with the $x-$axis, the effect is simply to reflect in a line that makes angle $\alpha + \gamma - \beta$ with the x-axis. Confirm this, using R, for the case where $\alpha = 40°$, $\beta = 30°$, and $\gamma = 80°$. Make a plot in R to show where the point $P = (1, 0)$ ends up after each of the three successive reflections.

8. Vectors in three dimensions (see script 2Y, topic 3)

   The least expensive way to fly from Boston (latitude 42.36° N, longitude 71.06° W) to Naples (latitude 40.84° N, longitude 14.26° E) is to buy a ticket on Aer Lingus and change planes in Dublin (latitude 53.35° N, longitude 6.26° W). Since Dublin is more than 10 degrees further north than either Boston or Naples, it is possible that the stop in Dublin might lengthen the journey substantially.

   (a) Construct unit vectors in $\mathbb{R}^3$ that represent the positions of the three cities.

   (b) By computing angles between these vectors, compare the length in kilometers of a nonstop flight with the length of a trip that stops in Dublin. Remember that, by the original definition of the meter, the distance from the North Pole to the Equator along the meridian through Paris is 10,000 kilometers. (You may treat the Earth as a sphere of unit radius.)

   (c) Any city that is on the great-circle route from Boston to Naples has a vector that lies in the same plane as the vectors for Boston and Naples. Invent a test for such a vector (you may use either cross products or determinants), and apply it to Dublin.

1 Major Concepts

1. You will learn next week that inversion of a matrix A can be accomplished by multiplication on the left by some sequence of "elementary matrices" $E_k * ... * E_1$. For now, all that matters is that the determinant of an elementary matrix is nonzero. Given this, prove that a matrix A is invertible iff (if and only if) its determinant is nonzero. (You may take it as given that the result of Proof 2.2 generalizes to any number of matrices[1], not just 2, and generalizes beyond 3×3 matrices.)

2. $T : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear transformation. For each of the following sets of three vectors, can we determine the matrix that represents T by T's actions on each vector in the set? (The intended solution involves calculating a **determinant**, so that's why we've put it here. This question anticipates next week's material and will likely seem more interesting after you watch those lectures.)

   (a) $\begin{bmatrix}1\\2\\1\end{bmatrix}, \begin{bmatrix}-2\\-4\\3\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}$

   (b) $\begin{bmatrix}1\\2\\1\end{bmatrix}, \begin{bmatrix}-2\\-4\\3\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}$

3. Classic Inequality Results and interesting applications

   (a) Demonstrate that: $|x + y| \geq ||x| - |y||$

   (b) Demonstrate the Triangle Inequality for in (n) dimensions: $|x_1 + ... + x_n| \leq |x_1| + |x_2| + ...|x_n|$

   (c) Prove that $(x\cos(\theta) + y\sin(\theta))^2 \leq x^2 + y^2$

   (d) Prove that $\sum_{i=1}^{n} |x_i - [i/2]| \geq \sum_{i=1}^{n} |x_{i+1} - x_i|$

4. Making sense of rotation and reflection matrices

   (a) Let's think about rotating a vector through an angle θ. This is a linear transformation (let's call it f), so we can represent it by a matrix F. One way to think about building F is to think about how f would act on vectors in our *standard basis*.

   (b) f sends a vector extending straight along the x-axis, $\begin{bmatrix}1\\0\end{bmatrix}$, to the vector $\begin{bmatrix}\cos\theta\\\sin\theta\end{bmatrix}$. (Check this by geometry/trig.) f also sends a vector extending straight up along the y-axis to $\begin{bmatrix}-\sin\theta\\\cos\theta\end{bmatrix}$. Overall, we can then obtain a rotation matrix

   $$F = \begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}$$

   (c) Now let's think about the linear transformation g that accomplishes reflection along a line that makes an angle θ with the positive x-axis. g sends $\begin{bmatrix}1\\0\end{bmatrix}$ to $\begin{bmatrix}\cos 2\theta\\\sin 2\theta\end{bmatrix}$. What g does to $\begin{bmatrix}0\\1\end{bmatrix}$ is less straightforward to figure out using trig identities and stuff, but looking at the special cases $\theta = \frac{\pi}{4}$ and $\theta = \frac{\pi}{2}$, it seems reasonable that g takes $\begin{bmatrix}1\\0\end{bmatrix}$ to $\begin{bmatrix}\sin 2\theta\\-\cos 2\theta\end{bmatrix}$. So we can obtain a reflection matrix

   $$G = \begin{bmatrix}\cos 2\theta & \sin 2\theta\\ \sin 2\theta & -\cos 2\theta\end{bmatrix}$$

[1] To prove this, we would use a technique called *mathematical induction*, which we'll really get into in a couple weeks.

5. Parallelepiped, vectors, and volume
   A parallelepiped is spanned by the following vectors:
   $\vec{v_1} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ $\vec{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ $\vec{v}_3 = \begin{bmatrix} 2 \\ -2 \\ 2 \end{bmatrix}$

   (a) Calculate the volume of the parallelepiped.

   (b) Find the cosine of the angle between $\vec{v}_1 and\ \vec{v}_2$.

   (c) Find the sine of the angle between $\vec{v}_1\ and\ \vec{v}_2$ (don't use the result from previous part).

   (d) check that the cosine and sine you have found are correct, by checking that they satisfy the identity $\sin^2\theta + \cos^2\theta = 1$.

   (e) Starting from the cross and dot product definitions of sine and cosine, prove the identity that you used above.

MATHEMATICS 23a/E-23a, Fall 2018
Linear Algebra and Real Analysis I
Week 3 (Row Reduction, Independence, Basis)

Authors: Paul Bamberg and Kate Penner
R scripts by Paul Bamberg
Last modified: October 1, 2018 by Paul Bamberg (improved HW problem 2)

Reading

- Hubbard, Sections 2.1 through 2.5

Recorded Lectures

- Lecture 6 (Week 3, Class 1) (watch on September 25 or 26)
- Lecture 7 (Week 3, Class 2) (watch on September 27 or 28)

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

- 3.1. Equivalent descriptions of a basis:
  Prove that a maximal set of linearly independent vectors for a subspace of $\mathbb{R}^n$ is also a minimal spanning set for that subspace.
- 3.2 Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Prove that Ker T and Img T are subspaces of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively and that

  $\dim(\text{Ker } T) + \dim(\text{Img } T) = n.$

  This is Hubbard, Theorem 2.5.8. You may use the results of Theorems 2.5.4 and 2.5.6, which show that, after row reducing T, you can easily construct a basis for Ker T and for Img T.

R Scripts .
Note: if you are ever going to dabble in R, this is the week to do it. Row reduction and Gram-Schmidt are a pain to do by hand, but you can modify these scripts to solve many linear algebra problems.

- Script 1.3A-RowReduction.R
  Topic 1 - Row reduction to solve two equations, two unknowns
  Topic 2 - Row reduction to solve three equations, three unknowns
  Topic 3 - Row reduction by elementary matrices
  Topic 4 - Automating row reduction in R
  Topic 5 - Row reduction to solve equations in a finite field

- Script 1.3B-RowReductionApplications.R
  Topic 1 - Testing for linear independence or dependence
  Topic 2 - Inverting a matrix by row reduction
  Topic 3 - Showing that a given set of vectors fails to span $\mathbb{R}^n$
  Topic 4 - Constructing a basis for the image and kernel

- Script 1.3C-OrthonormalBasis.R
  Topic 1 - Using Gram-Schmidt to construct an orthonormal basis
  Topic 2 - Making a new orthonormal basis for $\mathbb{R}^3$
  Topic 3 - Testing the cross-product rule for isometries

- Script 1.3P-RowReductionProofs.R
  Topic 1 - In $\mathbb{R}^n$, $n+1$ vectors cannot be independent
  Topic 2 - In $\mathbb{R}^n$, $n-1$ vectors cannot span
  Topic 3 - An invertible matrix must be square

1 Executive Summary

1.1 Row reduction for solving systems of equations

When you solve the equation $A\vec{\mathbf{v}} = \vec{\mathbf{b}}$ you combine the matrix A and the vector $\vec{\mathbf{b}}$ into a single matrix. Here is a simple example.

$$x + 2y = 7, 2x + 5y = 16$$

Then $A = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}$, $\vec{\mathbf{v}} = \begin{bmatrix} x \\ y \end{bmatrix}$, $\vec{\mathbf{b}} = \begin{bmatrix} 7 \\ 16 \end{bmatrix}$, so that $A\vec{\mathbf{v}} = \vec{\mathbf{b}}$ exactly corresponds to our system of equations. Our matrix of interest is therefore $\begin{bmatrix} 1 & 2 & 7 \\ 2 & 5 & 16 \end{bmatrix}$

First, subtract twice row 1 from row 2, then subtract twice row 2 from row 1 to get $\begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \end{bmatrix}$

Interpret the result as a pair of equations (remember what each column corresponded to when we first appended A and $\vec{\mathbf{b}}$ together: $x = 3, y = 2$.
The final form we are striving for is **row-reduced echelon form**, in which

- The leftmost nonzero entry in every row is a "pivotal 1."
- Pivotal 1's move to the right as you move down the matrix.
- A column with a pivotal 1 has 0 for all its other entries.
- Any rows with all 0's are at the bottom.

The row-reduction algorithm converts a matrix to echelon form. Briefly,

1. SWAP rows, if necessary, so that the leftmost column that is not all zeroes has a nonzero entry in the first row.
2. DIVIDE by this entry to get a pivotal 1.
3. SUBTRACT multiples of the first row from the others to clear out the rest of the column under the pivotal 1.
4. Repeat these steps to get a pivotal 1 in the next row, with nothing but zeroes elsewhere in the column (including in the first row). Continue until the matrix is in echelon form.

A pivotal 1 in the final column indicates no solutions. A bottom row full of zeroes means that there are infinitely many solutions.

Row reduction can be used to find the inverse of a matrix. By appending the appropriately sized identity matrix, row reducing will give the inverse of the matrix.

1.2 Row reduction by elementary matrices

Each basic operation in the row-reduction algorithm for a matrix A can be achieved by multiplication on the left by an appropriate invertible elementary matrix.

- Type 1: Multiplying the kth row by a scalar m is accomplished by an elementary matrix formed by starting with the identity matrix and replacing the kth element of the diagonal by the scalar m.

  Example: $E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ multiplies the second row of matrix A by 3.

- Type 2: Adding b times the jth row to the kth row is accomplished by an elementary matrix formed by starting with the identity matrix and changing the jth element in the kth row for 0 to the scalar b.

  Example: $E_2 = \begin{bmatrix} 1 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ adds three times the second row of matrix A to the first row.

  You want to multiply the second row of A by 3, so the 3 must be in the second column of E_2. Since the 3 is in the first row of E_2, it will affect the first row of E_2A.

- Type 3: Swapping row j with row k is accomplished by an elementary matrix formed by starting with the identity matrix, changing the jth and kth elements on the diagonal to 0, and changing the entries in row j, column k and in row k, column j from 0 to 1.

  Example: $E_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ swaps the first and third rows of matrix A.

Suppose that $A|I$ row-reduces to $A'|B$. Then $EA = A'$ and $EI = B$, where $E = E_k \cdots E_2 E_1$ is a product of elementary matrices. Since each elementary matrix is invertible, so is E. Clearly $E = B$, which means that we can construct E during the row-reduction process by appending the identity matrix I to the matrix A that we are row reducing.

If matrix A is invertible, then $A' = I$ and $E = A^{-1}$. However, the matrix E is invertible even when the matrix A is not invertible. Remarkably, E is also unique: it comes out the same even if you carry out the steps of the row-reduction algorithm in a non-standard order.

1.3 Row reduction for determining linear independence

Given a set of elements such as $\{a_1, a_2, a_3, a_4\}$, a **linear combination** is the name given to any arbitrary sum of scalar multiples of those elements. For instance: $a_1 - 2a_2 + 4a_3 - 5a_4$ is a linear combination of the above set.

Given some set of vectors, we describe the set as **linearly independent** if none of the vectors can be written as a linear combination of the others. Similarly, we describe the set as **linearly dependent** if one or more of the vectors can be written as a linear combination of the others.

A **subspace** is a set of vectors (usually an infinite number of them) that is **closed** under addition and scalar multiplication. "Closed" means that the sum of any two vectors in the set is also in the set and any scalar multiple of a vector in the set is also in the set. A subspace of F^n is the set of all possible linear combinations of some set of vectors. This set is said to **span** or to **generate** the subspace

A subspace $W \in F^n$ has the following properties:

1. The element $\vec{\mathbf{0}}$ is in W.
2. For any two elements $\vec{\mathbf{u}}$, $\vec{\mathbf{v}}$ in W, the sum $\vec{\mathbf{u}} + \vec{\mathbf{w}}$ is also in W.
3. For any element $\vec{\mathbf{v}}$ in W and any scalar c in F, the element $c\vec{\mathbf{v}}$ is also in W.

A **basis** of a vector space or subspace is a linearly independent set that spans that space.

The definition of a basis can be stated in three equivalent ways, each of which implies the other two:

a) It is a maximal set of linearly independent vectors in V: if you add any other vector in V to this set, it will no longer be linearly independent.

b) It is a minimal spanning set: it spans V, but if you remove any vector from this set, it will no longer span V.

c) It is a set of linearly independent vectors that spans V.

The number of elements in a basis for a given vector space is called the **dimension** of the vector space. A subspace has at most the same dimension as the space of which it is a subspace.

By creating a matrix whose columns are the vectors in a set and row reducing, we can find a maximal linearly independent subset, namely the columns that become columns with pivotal 1's. Any column that becomes a column without a pivotal 1 is a linear combination of the columns to its left.

1.4 Finding a vector outside the span

To show that a set of vectors $\{\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \cdots, \vec{\mathbf{v}}_k\}$ does not span F^n, we must exhibit a vector $\vec{\mathbf{w}}$ that is not a linear combination of the vectors in the given set.

- Create an $n \times k$ matrix A whose columns are the given vectors.
- Row-reduce this matrix, forming the product E of the elementary matrices that accomplish the row reduction.
- If the original set of vectors spans F^n, the row-reduced matrix EA will have n pivotal columns. Otherwise it will have fewer than n pivotal 1s, and there will be a row of zeroes at the bottom. If that is the case, construct the vector $\vec{\mathbf{w}} = E^{-1}\vec{\mathbf{e}}_n$.
- Now consider what happens when you row reduce the matrix $A|\vec{\mathbf{w}}$. The last column will contain a pivotal 1. Therefore the vector $\vec{\mathbf{w}}$ is independent of the columns to its left: it is not in the span of the set $\{\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \cdots, \vec{\mathbf{v}}_k\}$.

If $k < n$, then matrix A has fewer than n columns, so the matrix EA has fewer than n pivotal columns and must have a row of zeroes at the bottom. It follows that the vector $\vec{\mathbf{w}} = E^{-1}\vec{\mathbf{e}}_n$ can be constructed and that a set of fewer than n vectors cannot span F^n.

1.5 Image, kernel, and the dimension formula

Consider linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$, represented by matrix $[T]$.

- The image of T, Img T, is the set of vectors that lie in the subspace spanned by the columns of $[T]$.
- Img T is a subspace of $\mathbb{R}^m$. Its dimension is r, the **rank** of matrix $[T]$.
- A solution to the system of equations $T(\vec{\mathbf{x}}) = \vec{\mathbf{b}}$ is guaranteed to exist (though it may not be unique) if and only if Img T is m-dimensional.
- To find a basis for Img T, use the columns of the matrix $[T]$ that become pivotal columns as a result of row reduction.
- The kernel of T, Ker T, is the set of vectors $\vec{\mathbf{x}}$ for which $T(\vec{\mathbf{x}}) = \mathbf{0}$.
- Ker T is a subspace of $\mathbb{R}^n$.
- A system of equations $T(\vec{\mathbf{x}}) = \vec{\mathbf{b}}$ has a unique solution (though perhaps no solution exists) if and only if Ker T is zero-dimensional.
- There is an algorithm (Hubbard pp 196-197) for constructing an independent vector in Ker T from each of the $n - r$ nonpivotal columns of $[T]$.
- Since dim Img $T = r$ and dim Ker $T = n - r$,
  dim Img T + dim Ker $T = n$ (the "rank-nullity theorem.")

1.6 Linearly independent rows

Hubbard (page 200) gives two arguments that the number of linearly independent rows of a matrix equals its rank. Here is yet another.

Swap rows to put a nonzero row as the top row. Then swap a row that is linearly independent of the top row into the second position. Swap a row that is linearly independent of the top two rows into the third position. Continue until the top r rows are a linearly independent set, while each of the bottom $m - r$ rows is a linear combination of the top r rows.

Continuing with elementary row operations, subtract appropriate multiples of the top r rows from each of the bottom rows in succession, reducing it to zero. (Easy in principle but hard in practice!). The top rows, still untouched, are linearly independent, so there is no way for row reduction to convert any of them to a zero row. In echelon form, the matrix will have r pivotal 1s: rank r.

It follows that r is both the number of linearly independent columns and the number of linearly independent rows: the rank of A is equal to the rank of its transpose A^T.

1.7 Orthonormal basis

A basis is called **orthogonal** if any two distinct vectors in the basis have a dot product of zero. If, in addition, each basis vector is a unit vector, then the basis is called **orthonormal**.

Given any basis $\{\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \cdots, \vec{\mathbf{v}}_k\}$ of a subspace W and any vector $\vec{\mathbf{x}} \in W$, we can express $\vec{\mathbf{x}}$ as a linear combination of the basis vectors:

$$\vec{\mathbf{x}} = c_1\vec{\mathbf{v}}_1 + c_2\vec{\mathbf{v}}_2 + \cdots + c_k\vec{\mathbf{v}}_k,$$

but determining the coefficients requires row reducing a matrix.

If the basis $\{\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \cdots, \vec{\mathbf{v}}_k\}$ is orthonormal, just take the dot product with $\vec{\mathbf{v}}_i$ to determine that $\vec{\mathbf{x}} \cdot \vec{\mathbf{v}}_i = c_i$.

We can convert any spanning set of vectors into a basis. Here is the algorithm, sometimes called the "Gram-Schmidt process."
Choose any vector $\vec{\mathbf{w}}_1$: divide it by its length to make the first basis vector $\vec{\mathbf{v}}_1$.
Choose any vector $\vec{\mathbf{w}}_2$ that is linearly independent of $\vec{\mathbf{v}}_1$ and subtract off a multiple of $\vec{\mathbf{v}}_1$ to make a vector $\vec{\mathbf{x}}$ that is orthogonal to $\vec{\mathbf{v}}_1$.

$\vec{\mathbf{x}} = \vec{\mathbf{w}}_2 - (\vec{\mathbf{w}}_2 \cdot \vec{\mathbf{v}}_1)\vec{\mathbf{v}}_1$

Divide this vector by its length to make the second basis vector $\vec{\mathbf{v}}_2$.
Choose any vector $\vec{\mathbf{w}}_3$ that is linearly independent of $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$, and subtract off multiples of $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$ to make a vector $\vec{\mathbf{x}}$ that is orthogonal to both $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$.

$\vec{\mathbf{x}} = \vec{\mathbf{w}}_3 - (\vec{\mathbf{w}}_3 \cdot \vec{\mathbf{v}}_1)\vec{\mathbf{v}}_1 - (\vec{\mathbf{w}}_3 \cdot \vec{\mathbf{v}}_2)\vec{\mathbf{v}}_2$

Divide this vector by its length to make the third basis vector $\vec{\mathbf{v}}_3$.
Continue until you can no longer find any vector that is linearly independent of your basis vectors.

2 Lecture Outline

1. Row reduction
   This is just an organized version of the techniques for solving simultaneous equations that you learned in high school.

   When you solve the equation $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ you combine the matrix A and the vector $\vec{\mathbf{b}}$ into a single matrix. Here is a simple example.

   The equations are
   $x + 2y = 7$
   $2x + 5y = 16.$

   Then $A = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}$, $\vec{\mathbf{b}} = \begin{bmatrix} 7 \\ 16 \end{bmatrix}$,

   and we must row-reduce the 2×3 matrix $\begin{bmatrix} 1 & 2 & 7 \\ 2 & 5 & 16 \end{bmatrix}$.

   First, subtract twice row 1 from row 2 to get

   Then subtract twice row 2 from row 1 to get

   Interpret the result as a pair of equations:

   Solve these equations (by inspection) for x and y

   You see the general strategy. First eliminate x from all but the first equation, then eliminate y from all but the second, and keep going until, with luck, you have converted each row into an equation that involves only a single variable with a coefficient of 1.

2. Echelon Form

The result of row reduction is a matrix in echelon form, whose properties are carefully described on p. 165 of Hubbard (definition 2.1.5). Here is Hubbard's messiest example:

$$\begin{bmatrix} 0 & 1 & 3 & 0 & 0 & 3 & 0 & 4 \\ 0 & 0 & 0 & 1 & -2 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}.$$

Key properties:

- The leftmost nonzero entry in every row is a "pivotal 1."
- Pivotal 1's move to the right as you move down the matrix.
- A column with a pivotal 1 has 0 for all its other entries.
- Any rows with all 0's are at the bottom.

If a matrix is not is echelon form, you can convert it to echelon form by applying one or more of the following row operations.

(a) Multiply a row by a nonzero number.

(b) Add (or subtract) a multiple of one row from another row.

(c) Swap two rows.

Here are the "what's wrong?" examples from Hubbard. Find row operations that fix them.

$$\begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & 1 \end{bmatrix}.$$

$$\begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

$$\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

$$\begin{bmatrix} 0 & 1 & 0 & 3 & 0 & -3 \\ 0 & 0 & -1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}.$$

3. Row reduction algorithm

   The row-reduction algorithm (Hubbard, p. 166) converts a matrix to echelon form. Briefly,

   (a) SWAP rows so that the leftmost column that is not all zeroes has a nonzero entry in the first row.

   (b) DIVIDE by this entry to get a pivotal 1.

   (c) SUBTRACT multiples of the first row from the others to clear out the rest of the column under the pivotal 1.

   (d) Repeat these steps to get a pivotal 1 in the second row, with nothing but zeroes elsewhere in the column (including in the first row).

   (e) Repeat until the matrix is in echelon form.

   Carry out this procedure to row-reduce the matrix $\begin{bmatrix} 0 & 3 & 3 & 6 \\ 2 & 4 & 2 & 4 \\ 3 & 8 & 4 & 7 \end{bmatrix}$.

4. Solving equations

   Once you have row-reduced the matrix, you can interpret it as representing the equation $\tilde{A}\vec{\mathbf{x}} = \tilde{\vec{\mathbf{b}}}$,

   which has the same solutions as the equation with which you started, except that now they can be solved by inspection.

   A pivotal 1 in the last column $\tilde{\vec{\mathbf{b}}}$ is the kiss of death, since it is an equation like $0x + 0y = 1$. There is no solution. This happens, for example,

   when row reduction converts $\begin{bmatrix} 1 & 2 & 3 & 2 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 2 & 0 \end{bmatrix}$ to $\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

   The top and bottom rows say that $x + 2y + 3z = 2, x + y + 2z = 0$, so that $y + z = 2$. The middle row says that $y + z = 1$, inconsistent with $y + z = 2$. The row-reduced matrix expresses the inconsistency as $0x + 0y = 1$.

   Otherwise, choose freely the values of the "active" unknowns in the non-pivotal columns(excluding the last one). Then each row gives the value of the "passive" unknown in the column that has the pivotal 1 for that row. This happens, for example,

   when row reduction converts $\begin{bmatrix} 2 & 1 & 3 & 1 \\ 1 & -1 & 0 & 1 \\ 1 & 1 & 2 & \frac{1}{3} \end{bmatrix}$ to $\begin{bmatrix} 1 & 0 & 1 & \frac{2}{3} \\ 0 & 1 & 1 & -\frac{1}{3} \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

   The only nonpivotal column(except the last one) is the third. So we can choose the value of the active unknown z freely.

   Then the first row gives x in terms of z: $x = \frac{2}{3} - z$.

   The second row gives y in terms of z: $y = -\frac{1}{3} - z$.

   If there are as many equations as unknowns, this situation is exceptional. If there are fewer equations than unknowns, it is the usual state of affairs. Expressing the passive variables in terms of the active ones will be the subject of the important implicit function theorem in week 11.

   A column that is all zeroes is nonpivotal. Such a column must have been there from the start; it cannot come about as a result of row reduction. It corresponds to an unknown that was never mentioned. This sounds unlikely, but it can happen when you represent a system of equations by an arbitrary matrix.

   Example: In $\mathbb{R}^3$, solve the equations $x = 0, y = 0$ (z not mentioned)

5. Many for the price of one
   If you have several equations with the same matrix A on the left and different vectors on the right, you can solve them all in the process of row-reducing A. See Example 2.2.9 in Hubbard for some gory details.

   Row reduction is more efficient than computing A^{-1}, and it works even when A is not invertible. Here is simple example with a non-invertible A:

   $x + 2y = 3$
   $2x + 4y = 6$

   $x + 2y = 3$
   $2x + 4y = 7$

   The first pair has infinitely many solutions: choose any y and take $x = 3 - 2y$. The second set has none.

   We must row-reduce the 2×4 matrix

$$\begin{bmatrix} 1 & 2 & 3 & 3 \\ 2 & 4 & 6 & 7 \end{bmatrix}.$$

   This quickly gives

$$\begin{bmatrix} 1 & 2 & 3 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

   and then

$$\begin{bmatrix} 1 & 2 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

   The last column has a pivotal 1 – no solution for the second set.

   The third column has no pivotal 1, and the second column is also nonpivotal, so there are multiple solutions for the first set of equations. Make a free choice of the active variable y that goes with nonpivotal column 2.

   How does the first row now determine the passive unknown x?

6. Matrix inversion by row reduction

   If A is square and you choose each standard basis vector in turn for the right-hand side, then row reduction constructs the inverse of A if it exists.

   As a simple example, we invert $A = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}$.

   Begin by appending the standard basis vectors as third and fourth columns to get

$$\begin{bmatrix} 1 & 2 & 1 & 0 \\ 2 & 5 & 0 & 1 \end{bmatrix}.$$

   Now row-reduce this in two easy steps:

   The right two columns of the row-reduced matrix are the desired inverse: check it!

   For matrices larger than 2×2, row reduction is a more efficient way of constructing a matrix inverse than any techniques involving determinants that you may have learned!

   Hubbard, Example 2.3.4, states that the matrix $\begin{bmatrix} 2 & 1 & 3 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 2 & 0 & 0 & 1 \end{bmatrix}$

   row reduces to $\begin{bmatrix} 1 & 0 & 0 & 3 & -1 & -4 \\ 0 & 1 & 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & -2 & 1 & 3 \end{bmatrix}$.

   Identify A and A^{-1}.

   If A is not square, it cannot row-reduce to the identity; so it is not invertible. We have finally proved that *a matrix can be invertible only if it is square.*

7. Elementary matrices:
Each basic operation in the row-reduction algorithm can be achieved by multiplication on the left by an appropriate invertible elementary matrix.

Here are examples of the three types of elementary matrix. For each, figure out what row operation is achieved by converting $A = \begin{bmatrix} 2 & 4 \\ -1 & 1 \\ 1 & 0 \end{bmatrix}$ to EA.

- Type 1: $E_1 = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
- Type 2: $E_2 = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
- Type 3: $E_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

In practice, use of elementary matrices does not speed up computation, but it provides a nice way to think about row reduction for purposes of doing proofs.

For example, as on page 180 of Hubbard, suppose that $A|I$ row-reduces to $I|B$.

Then $EA = I$ and $EI = B$, where

$E = E_k \cdots E_2 E_1$ is a product of elementary matrices. Since each elementary matrix is invertible, so is E. Clearly $E = B$, which means that we can construct E during the row-reduction process.

In the case where A row-reduces to the identity there is an easy proof that E is unique.

Start with $EA = I$.

Multiply by E^{-1} on the left, E on the right, to get

$E^{-1}EAE = E^{-1}E$,

from which it follows that $AE = I$. So E is also a right inverse of A. But we earlier proved that if a matrix A has a right inverse and a left inverse, both are unique.

8. Row reduction and elementary matrices

We want to solve the equations
$3x + 6y = 21$
$2x + 5y = 16.$

Then $A = \begin{bmatrix} 3 & 6 \\ 2 & 5 \end{bmatrix}$, $\vec{\mathbf{b}} = \begin{bmatrix} 21 \\ 16 \end{bmatrix}$,

and we must row-reduce the 2×3 matrix $\begin{bmatrix} 3 & 6 & 21 \\ 2 & 5 & 16 \end{bmatrix}$.

Use an elementary matrix to accomplish each of the three steps needed to accomplish row reduction.

Matrix E_1 divides the top row by 3.

Matrix E_2 subtracts twice row 1 from row 2.

Matrix E_3 subtracts twice row 2 from row 1.

Interpret the result as a pair of equations and solve them (by inspection) for x and y.

Show that the product $E_3E_2E_1$ is the inverse of A.

9. Linear combinations and span

The defining property of a linear function T: for any collection of k vectors in F^n, $\vec{\mathbf{v}}_1, \cdots \vec{\mathbf{v}}_k$, and any collection of coefficients $a_1 \cdots a_k$ in field F,

$$T(\sum_{i=1}^{k} a_i \vec{\mathbf{v}}_i) = \sum_{i=1}^{k} a_i T(\vec{\mathbf{v}}_i).$$

The sum $\sum_{i=1}^{k} a_i \vec{\mathbf{v}}_i$ is called a linear combination of the vectors $\vec{\mathbf{v}}_1, \cdots \vec{\mathbf{v}_k}$.

The set of all the linear combinations of $\vec{\mathbf{v}}_1, \cdots \vec{\mathbf{v}_k}$ is called the span of the set $\vec{\mathbf{v}}_1, \cdots \vec{\mathbf{v}}_k$.

Prove that it is a subspace of F^n.

Suppose $\vec{\mathbf{v}}_1 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$, $\vec{\mathbf{v}}_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$, $\vec{\mathbf{v}}_3 = \begin{bmatrix} 3 \\ -1 \\ -2 \end{bmatrix}$,$\vec{\mathbf{w}}_1 = \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}$,$\vec{\mathbf{w}}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

- Show that $\vec{\mathbf{w}}_1$ is a linear combination of $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$.

- Invent an easy way to describe the span of $\vec{\mathbf{v}}_1$, $\vec{\mathbf{v}}_2$, and $\vec{\mathbf{v}}_3$. (Hint: consider the sum of the components.)

- Thereby show that $\vec{\mathbf{w}}_1$ is in the span of $\vec{\mathbf{v}}_1$, $\vec{\mathbf{v}}_2$, and $\vec{\mathbf{v}}_3$ but $\vec{\mathbf{w}}_2$ is not.

- The matrix $\begin{bmatrix} 1 & 0 & 3 & 2 & 1 \\ -2 & 1 & -1 & -3 & 0 \\ 1 & -1 & -2 & 1 & 0 \end{bmatrix}$ row reduces to $\begin{bmatrix} 1 & 0 & 3 & 2 & 0 \\ 0 & 1 & 5 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$.
  How does this result answer the question of whether or not $\vec{\mathbf{w}}_1$ or $\vec{\mathbf{w}}_2$ is in the span of $\vec{\mathbf{v}}_1$, $\vec{\mathbf{v}}_2$, and $\vec{\mathbf{v}}_3$?

10. Linear independence

$\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \cdots \vec{\mathbf{v}}_k$ are linearly independent if the system of equations

$x_1\vec{\mathbf{v}}_1 + x_2\vec{\mathbf{v}}_2 + \cdots + x_k\vec{\mathbf{v}}_k = \vec{\mathbf{w}}$ has at most one solution.

To test for linear independence, make the vectors $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \cdots \vec{\mathbf{v}}_k$ into a matrix and row-reduce it. If any column is nonpivotal, then the vectors are linearly dependent. Here is an example.

The vectors to test for independence are $\vec{\mathbf{v}}_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 1 \end{bmatrix}$, $\vec{\mathbf{v}}_2 = \begin{bmatrix} 2 \\ 0 \\ 1 \\ 1 \end{bmatrix}$, $\vec{\mathbf{v}}_3 = \begin{bmatrix} 0 \\ 2 \\ 3 \\ 1 \end{bmatrix}$.

The vector $\vec{\mathbf{w}}$ is irrelevant and might as well be zero, so we just make a matrix from the three given vectors:

$\begin{bmatrix} 1 & 2 & 0 \\ 1 & 0 & 2 \\ 2 & 1 & 3 \\ 1 & 1 & 1 \end{bmatrix}$ reduces to $\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

The third column is nonpivotal; so the given vectors are linearly dependent. How can you write the third one as a linear combination of the first two?

Change $\vec{\mathbf{v}}_3$ to $\begin{bmatrix} 0 \\ 2 \\ 1 \\ 1 \end{bmatrix}$ and test again.

Now $\begin{bmatrix} 1 & 2 & 0 \\ 1 & 0 & 2 \\ 2 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ reduces to $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

There is no nonpivotal column. The three vectors are linearly independent.

Setting $\vec{\mathbf{w}} = \vec{\mathbf{0}}$, as we have already done, leads to the standard definition of linear independence: if

$a_1\vec{\mathbf{v}}_1 + a_2\vec{\mathbf{v}}_2 + \cdots a_k\vec{\mathbf{v}}_k = \vec{\mathbf{0}}$

then $a_1 = a_2 = \cdots = a_k = 0$.

11. Constructing a vector outside the span

The vectors are

$\vec{\mathbf{v}}_1 = \begin{bmatrix} 4 \\ 2 \\ 3 \end{bmatrix}$, $\vec{\mathbf{v}}_2 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$

$A = \begin{bmatrix} 4 & 2 \\ 2 & 1 \\ 3 & 2 \end{bmatrix}$ reduces to $EA = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$, and the matrix that does the job is

$E = \begin{bmatrix} 1 & 0 & -1 \\ -\frac{3}{2} & 0 & 2 \\ \frac{1}{2} & 1 & 0 \end{bmatrix}$.

We want to append a third column $\vec{\mathbf{b}}$ such that when we row reduce the square matrix $A|\vec{\mathbf{b}}$, the resulting matrix $EA|E\vec{\mathbf{b}}$ will have a pivotal 1 in the third column. In this case it will be in the bottom row. Since E, being a product of elementary matrices, must be invertible, we compute

$$E^{-1} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

We have found a vector, $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, that is not in the span of $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$.

Key point: the proof relies on the fact that this procedure will always work, because the matrix E that accomplishes row reduction is guaranteed to be invertible!

12. Two key theorems

- In $\mathbb{R}^n$, a set of $n+1$ vectors cannot be linearly independent.

  If we start with $n+1$ vectors in $\mathbb{R}^n$, make a matrix that has these vectors as its columns, and row-reduce, the best we can hope for is to get a pivotal 1 in each of n columns. There must be at least one non-pivotal column (not necessarily the last column), and the $n+1$ vectors must be linearly dependent: they cannot be linearly independent.

  Show what the row-reduced matrix looks like and how it is possible for the non-pivotal column not to be the last column.

- In $\mathbb{R}^n$, a set of $n-1$ vectors cannot span.

  Remember that "span" means

  $\forall \vec{\mathbf{w}}, x_1\vec{\mathbf{v}}_1 + x_2\vec{\mathbf{v}}_2 + \cdots x_k\vec{\mathbf{v}}_k = \vec{\mathbf{w}}$ has at least one solution.

  Since "exists" is easier to work with than "for all", convert this into a definition of "does not span." A set of k vectors does not span if

  $\exists \vec{\mathbf{w}}$ such that $x_1\vec{\mathbf{v}}_1 + x_2\vec{\mathbf{v}}_2 + \cdots x_k\vec{\mathbf{v}}_k = \vec{\mathbf{w}}$ has no solution.

  We invent a method for constructing $\vec{\mathbf{w}}$, using elementary matrices.

  Make a matrix A whose columns are $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \cdots \vec{\mathbf{v}}_k$, and row-reduce it by elementary matrices whose product can be called E. Then EA is in echelon form.

  If A has only $n-1$ columns, it cannot have more than $n-1$ pivotal 1's, and there cannot be a pivotal 1 in the bottom row. That means that if we had chosen a $\vec{\mathbf{w}}$ that row-reduced to a pivotal 1 in the last row, the set of equations

  $x_1\vec{\mathbf{v}}_1 + x_2\vec{\mathbf{v}}_2 + \cdots x_k\vec{\mathbf{v}}_k = \vec{\mathbf{w}}$

  would have had no solution.

  Now E is the product of invertible elementary matrices, hence invertible. Just construct $\vec{\mathbf{w}} = E^{-1}\vec{\mathbf{e_n}}$ as an example of a vector that is not in the span of the given $n-1$ vectors.

13. Definition of basis (Hubbard, Definition 2.4.12)

A basis for a subspace $V \subset \mathbb{R}^n$ has the following equivalent properties:

(a) It is a maximal set of linearly independent vectors in V: if you add any other vector in V to the set, it will no longer be linearly independent.

(b) It is a minimal spanning set: it spans V, but if you remove any vector from the set, it will no longer span.

(c) It is a set of linearly independent vectors that spans V.

To show that any of these three properties implies the other two would require six proofs. Let's do just one. Call the basis vectors $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \cdots \vec{\mathbf{v}}_k$.

Prove that (a) implies (b) (this is your proof 3.1).

When we add any other vector $\vec{\mathbf{w}}$ to the basis set, the resulting set is linearly dependent. Express this statement as an equation that includes the term $b\vec{\mathbf{w}}$.

Show that if $b \neq 0$, we can express $\vec{\mathbf{w}}$ as a linear combination of the basis set. This will prove "spanning set".

To prove that $b \neq 0$, assume the contrary, and show that the vectors $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \cdots \vec{\mathbf{v}}_k$ would be linearly dependent.

To prove "minimal spanning set," just exhibit a vector that is not in the span of $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \cdots \vec{\mathbf{v}}_{k-1}$.

Now we combine this definition of basis with what we already know about sets of vectors in $\mathbb{R}^n$.

Our conclusions:

In $\mathbb{R}^n$, a basis cannot have $< n$ elements, since they would not span.

In $\mathbb{R}^n$, a basis cannot have $> n$ elements, since they would not be linearly independent.

So any basis for $\mathbb{R}^n$ must, like the standard basis, have exactly n elements.

14. Basis for a subspace

Consider any subspace $E \subset \mathbb{R}^n$. We need to prove the following:

- E has a basis.
- Any two bases for E have the same number of elements, called the dimension of E.

Before the proof, consider an example.

$E \subset \mathbb{R}^3$ is the set of vectors for which $x_1 + x_2 + x_3 = 0$.

One basis is $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$.

Another basis is $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$.

It's obvious that either basis is linearly independent, since neither basis vector is zero, and one is not a multiple of the other.

How could we establish linear independence by using row reduction?

To show that each basis spans is less trivial. Fortunately, in this simple case we can write an expression for the general element of E as $\begin{bmatrix} a \\ b \\ -a-b \end{bmatrix}$

How would we express this general element as a linear combination of basis vectors?

Now we proceed to the proof. First we must prove the existence of a basis by explaining how to construct one.

How to make a basis for a non-empty subspace E in general:

Choose any $\vec{\mathbf{v}}_1$ to get started. Notice that we need not specify a method for doing this! The justification for this step is the so-called "axiom of choice."

If $\vec{\mathbf{v}}_1$ does not span E, choose $\vec{\mathbf{v}}_2$ that is not in the span of $\vec{\mathbf{v}}_1$ (not a multiple of it). Again, we do not say how to do this, but it must be possible since $\vec{\mathbf{v}}_1$ does not span E.

If $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$ do not span E, choose $\vec{\mathbf{v}}_3$ that is not in the span of $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$ (not a linear combination).

Keep going until you have spanned the space. By construction, the set is linearly independent. So it is a basis.

Second, we must prove that every basis has the same number of vectors.

Imagine that two people have done this and come up with bases of possibly different sizes.

One is $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \cdots \vec{\mathbf{v}}_m$.

The other is $\vec{\mathbf{w}}_1, \vec{\mathbf{w}}_2, \cdots \vec{\mathbf{w}}_p$.

Since each basis spans E, we can write each $\vec{\mathbf{w}}_j$ as a linear combination of the $\vec{\mathbf{v}}$. It takes m coefficients to do this for each of the p vectors, so we end up with an $m \times p$ matrix A, each of whose columns is one of the $\vec{\mathbf{w}}_j$.

We can also write each $\vec{\mathbf{v}}_i$ as a linear combination of the $\vec{\mathbf{w}}_j$. It takes p coefficients to do this for each of the m vectors, so we end up with a $p \times m$ matrix B, each of whose columns is one of the $\vec{\mathbf{v}}_i$.

Clearly $AB = I$ and $BA = I$. So A is invertible, hence square, and $m = p$.

15. Kernels and Images(first part of proof 3.2)

Consider linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$. This can be represented by a matrix, but we want to stay abstract for the moment.

- The kernel of T, Ker T, is the set of vectors $\vec{\mathbf{x}}$ for which $T(\vec{\mathbf{x}}) = \mathbf{0}$.
- A system of equations $T(\vec{\mathbf{x}}) = \vec{\mathbf{b}}$ has a unique solution if and only if Ker T is zero-dimensional.
  Assume that $T(\vec{\mathbf{x}}_1) = \vec{\mathbf{b}}$ and $T(\vec{\mathbf{x}}_2) = \vec{\mathbf{b}}$.
  Since T is linear,
  $T(\vec{\mathbf{x}}_1 - \vec{\mathbf{x}}_2) = \vec{\mathbf{b}} - \vec{\mathbf{b}} = \mathbf{0}$.
  If the kernel is zero-dimensional, it contains only the zero vector, and $\vec{\mathbf{x}}_1 = \vec{\mathbf{x}}_2$.
  Conversely, if the solution is unique: the only way that $\vec{\mathbf{x}}_1$ and $\vec{\mathbf{x}}_2$ can both be solutions is $\vec{\mathbf{x}}_1 = \vec{\mathbf{x}}_2$, the kernel is zero-dimensional.
- Ker T is a subspace of $\mathbb{R}^n$.
  Proof:
  If $\vec{\mathbf{x}}$ and $\vec{\mathbf{y}}$ are elements of Ker T, then, because T is linear,
  $T(a\vec{\mathbf{x}} + b\vec{\mathbf{y}}) = aT(\vec{\mathbf{x}}) + bT(\vec{\mathbf{y}}) = \mathbf{0}$.
- The image of T, Img T, is the set of vectors $\vec{\mathbf{w}}$ for which $\exists\vec{\mathbf{v}}$ such that $\vec{\mathbf{w}} = T(\vec{\mathbf{v}})$.
- Img T is a subspace of $\mathbb{R}^m$.
  Proof:
  If $\vec{\mathbf{w}}_1$ and $\vec{\mathbf{w}}_2$ are elements of Img T, then
  $\exists\vec{\mathbf{v}}_1$ such that $\vec{\mathbf{w}}_1 = T(\vec{\mathbf{v}}_1)$ and
  $\exists\vec{\mathbf{v}}_2$ such that $\vec{\mathbf{w}}_2 = T(\vec{\mathbf{v}}_2)$
  $T(a\vec{\mathbf{v}}_1 + b\vec{\mathbf{v}}_2) = aT(\vec{\mathbf{v}}_1) + bT(\vec{\mathbf{v_2}}) = a\vec{\mathbf{w_1}} + b\vec{\mathbf{w}}_2$.
  We have shown that any linear combination of elements of Img T is also an element of Img T.

16. Basis for the image

To find a basis for the image of T, we must find a linearly independent set of vectors that span the image. Spanning the image is not a problem: the columns of the matrix for T do that. The hard problem is to choose a linearly independent set. The secret is to use row reduction.

Each nonpivotal column is a linear combination of the columns to its left, hence inappropriate to include in a basis. It follows that the pivotal columns of T form a basis for the image. Of course, you can permute the columns and come up with a different basis: no one said that a basis is unique.

Here is an example.

The matrix $T = \begin{bmatrix} 1 & 2 & 1 & 1 \\ 0 & 0 & 1 & -1 \\ 2 & 4 & 1 & 3 \end{bmatrix}$ row reduces to $\begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

By inspecting these two matrices, find a basis for Img T. Notice that the dimension of Img T is 2, which is less than the number of rows, and that the two leftmost columns do not form a basis.

17. Basis for the kernel

The matrix $T = \begin{bmatrix} 1 & 2 & 1 & 1 \\ 0 & 0 & 1 & -1 \\ 2 & 4 & 1 & 3 \end{bmatrix}$ row reduces to $\begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

To find a basis for Ker T, look at the row-reduced matrix and identify the nonpivotal columns. For each nonpivotal column i in turn, put a 1 in the position of that column, a 0 in the position of all other nonpivotal columns, and leave blanks in the other positions. The resulting vectors must be linearly independent, since for each of them, there is a position where it has a 1 and where all the others have a zero. What are the resulting (incomplete) basis vectors for Ker T?

Now fill in the blanks: assign values in the positions of all the pivotal columns so that $T(\vec{\mathbf{v}_\mathbf{i}}) = 0$. The vectors $\vec{\mathbf{v}_\mathbf{i}}$ span the kernel, since assigning a value for each nonpivotal variable is precisely the technique for constructing the general solution to $T(\vec{\mathbf{v}}) = 0$.

18. Rank - nullity theorem(second part of proof 3.2)

    The matrix of $T : \mathbb{R}^n \to \mathbb{R}^m$ has n columns. We row-reduce it and find r pivotal columns and $n - r$ nonpivotal columns. The integer r is called the *rank* of the matrix. It is also equal to the number of linearly independent rows of the matrix.

    Each pivotal column gives rise to a basis vector for the image; so the dimension of Img T is r.

    Each nonpivotal column gives rise to a basis vector for the kernel; so the dimension of Ker T is $n - r$.

    Clearly, dim(Ker T) + dim(Img T) = n.

    In the special case of a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$, represented by a square $n \times n$ matrix, if the rank $r = n$ then

    - any equation $T(\vec{\mathbf{v}}) = \vec{\mathbf{b}}$ has a solution, since the image is n-dimensional.
    - any equation $T(\vec{\mathbf{v}}) = \vec{\mathbf{b}}$ has a unique solution, since the kernel is 0-dimensional.
    - T is invertible.

19. Linearly independent rows

    Hubbard (page 200) gives two arguments that the number of linearly independent rows of a matrix equals its rank. Here is yet another.

    Swap rows to put a nonzero row as the top row. Then swap a row that is linearly independent of the top row into the second position. Swap a row that is linearly independent of the top two rows into the third position. Continue until the top r rows are a linearly independent set, while each of the bottom $m - r$ rows is a linear combination of the top r rows.

    Continuing with elementary row operations, subtract appropriate multiples of the top r rows from each of the bottom rows in succession, reducing it to zero. (Easy in principle but hard in practice!). The top rows, still untouched, are linearly independent, so there is no way for row reduction to convert any of them to a zero row. In echelon form, the matrix will have r pivotal 1s: rank r.

    It follows that r is both the number of linearly independent columns and the number of linearly independent rows: the rank of A is equal to the rank of its transpose A^T.

20. Orthonormal basis:

If we have a dot product, then we can convert any spanning set of vectors into a basis. Here is the algorithm, sometimes called the "Gram-Schmidt process." We will apply it to the 3-dimensional subspace of $\mathbb{R}^4$ for which the components sum to zero.

Choose any vector $\vec{\mathbf{w}}_1$ and divide it by its length to make the first basis vector $\vec{\mathbf{v}}_1$.

If $\vec{\mathbf{w}}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}$, what is $\vec{\mathbf{v}}_1$?

Choose any vector $\vec{\mathbf{w}}_2$ that is linearly independent of $\vec{\mathbf{v}}_1$ and subtract off a multiple of $\vec{\mathbf{v}}_1$ to make a vector $\vec{\mathbf{x}}$ that is orthogonal to $\vec{\mathbf{v}}_1$. Divide this vector by its length to make the second basis vector $\vec{\mathbf{v}}_2$.

If $\vec{\mathbf{w}}_2 = \begin{bmatrix} 2 \\ -1 \\ -1 \\ 0 \end{bmatrix}$, calculate $\vec{\mathbf{x}} = \vec{\mathbf{w}}_2 - (\vec{\mathbf{w}}_2 \cdot \vec{\mathbf{v}}_1)\vec{\mathbf{v}}_1$

Choose any vector $\vec{\mathbf{w}}_3$ that is linearly independent of $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$, and subtract off multiples of $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$ to make a vector $\vec{\mathbf{x}}$ that is orthogonal to both $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$. Divide this vector by its length to make the third basis vector $\vec{\mathbf{v}}_3$. Continue until you can no longer find any vector that is linearly independent of your basis vectors.

Tedious computation gives $\vec{\mathbf{v}}_1 = \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}$, $\vec{\mathbf{v}}_2 = \begin{bmatrix} \frac{3}{2\sqrt{5}} \\ -\frac{1}{2\sqrt{5}} \\ -\frac{3}{2\sqrt{5}} \\ \frac{1}{2\sqrt{5}} \end{bmatrix}$, $\vec{\mathbf{v}}_3 = \begin{bmatrix} -\frac{1}{2\sqrt{5}} \\ -\frac{3}{2\sqrt{5}} \\ \frac{1}{2\sqrt{5}} \\ \frac{3}{2\sqrt{5}} \end{bmatrix}$.

A nice feature of an orthogonal basis (no need for it to be orthonormal) is that any set of orthogonal vectors is linearly independent.

Proof: assume $a_1\vec{\mathbf{v_1}} + a_2\vec{\mathbf{v_2}} + \cdots a_k\vec{\mathbf{v_k}} = \vec{\mathbf{0}}$.

Choose any $\vec{\mathbf{v_i}}$ and take the dot product with both sides of this equation. You get $a_i = 0$ for all i, which establishes independence.

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. Echelon form

   List the four requirements for a matrix to be in echelon form, illustrating each of them by pointing to the relevant features in the following matrix:

$$\begin{bmatrix} 0 & 1 & 2 & 0 & 1 & 3 & 0 & 4 \\ 0 & 0 & 0 & 1 & 4 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

2. Elementary matrices

   The matrix

$$A = \begin{bmatrix} 0 & 2 & 0 \\ 1 & 0 & 2 \\ 0 & 0 & 1 \end{bmatrix}$$

   can be reduced to the identity by multiplying successively on the left by three elementary matrices, i.e. $E_3E_2E_1A = I$.
   Exhibit E_1, E_2, and E_3 and show the action of each (if you write them down a column to the left of A you will not have to do any recopying).
   Use this procedure to prove that any invertible matrix can be expressed as a product of elementary matrices.

3. (Proof 3.1) – Equivalent descriptions of a basis:
   Prove that a maximal set of linearly independent vectors for a subspace of $\mathbb{R}^n$ is also a minimal spanning set for that subspace.

4. Dimension of a subspace (Hubbard, Proposition 2.4.19)
   First show that every subspace $E \subset \mathbb{R}^n$ has a basis.
   Then suppose that E has two different bases, $\vec{\mathbf{v}}_1, \cdots, \vec{\mathbf{v}}_m$ and $\vec{\mathbf{w}}_1, \cdots, \vec{\mathbf{w}}_p$.
   Using Hubbard Proposition 2.3.2, which proves that a matrix A is invertible only if it is square, prove that $m = p$.

5. (Proof 3.2) – Rank-nullity theorem
   Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Prove that Ker T and Img T are subspaces of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively and that

   $\dim(\text{Ker } T) + \dim(\text{Img } T) = n.$

   This is Hubbard, Theorem 2.5.8. You may use the results of Theorems 2.5.4 and 2.5.6, which show that, after row reducing T, you can easily construct a basis for Ker T and for Img T.

6. (Extra topic) Orthonormal basis

   Suppose that you have two vectors, $\vec{\mathbf{w}}_1$ and $\vec{\mathbf{w}}_2$, that span a two-dimensional subspace of $\mathbb{R}^n$. Describe the Gram-Schmidt procedure for converting this pair of vectors to an orthonormal basis $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$, for the same subspace. If time permits, explain how you would extend this procedure to a set of k vectors.

4 Workshop Problems

1. Row reduction and elementary matrices

   (a) By row reducing an appropriate matrix to echelon form, solve the system of equations

   $$2x + y + z = 2$$
   $$x + y + 2z = 2$$
   $$x + 2y + 2z = 1$$

   where all the coefficients and constants are elements of the finite field $\mathbb{Z}_3$. If there is no solution, say so. If there is a unique solution, specify the values of x, y, and z. If there is more than one solution, determine all solutions by giving formulas for two of the variables, perhaps in terms of the third one.

   (b) The matrix

   $$A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 2 & 3 & 4 \end{bmatrix}$$

   is not invertible. Nonetheless, there is a product E of four elementary matrices that will reduce it to echelon form. Find these four matrices and their product E. If you are willing to extend the definition of "elementary matrix" slightly, you can do the job with three matrices.

2. Some short proofs

   (a)
   - Show that type 3 elementary matrices are not strictly necessary, because it is possible to swap rows of a matrix by using only type 1 and type 2 elementary matrices. (If you can devise a way to swap the two rows of a 2×2 matrix, that it sufficient, since it is obvious how the technique generalizes.)
   - Prove that if a set of linearly independent vectors spans a vector space W, it is both a maximal linearly independent set and a minimal spanning set.

   (b)
   - Let $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ be linearly independent vectors in $\mathbb{R}^3$. Prove that if vector $\vec{\mathbf{w}}$ is orthogonal to $\vec{\mathbf{u}} \times \vec{\mathbf{v}}$, then $\vec{\mathbf{w}}$ is in the span of $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$. Hint: what happens if you make a matrix whose columns are $\vec{\mathbf{u}}, \vec{\mathbf{v}}$, and $\vec{\mathbf{w}}$ and row-reduce it?
   - Prove that in a vector space W, a minimal spanning set is a maximal linearly independent set.

3. Constructing a basis

   (a) Use Gram-Schmidt to construct an orthonormal basis for the two-dimensional subspace of $\mathbb{R}^3$ that is orthogonal to the vector $\begin{bmatrix} 4 \\ -1 \\ -1 \end{bmatrix}$. One vector in this subspace is $\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$.

   (b) Starting by doing row reduction, find a basis for the image and the kernel of the matrix

   $$A = \begin{bmatrix} 1 & 2 & 0 & 2 \\ 2 & 4 & 1 & 1 \\ 0 & 0 & -1 & 3 \end{bmatrix},$$

   Express the columns that are not in the basis for the image as linear combinations of the ones that are in the basis.

4. Problems to be solved by writing or editing R scripts.

   (a) The director of a budget office has to make changes to four line items in the budget, but her boss insists that they must sum to zero. Three of her subordinates make the following suggestions, all of which lie in the subspace of acceptable changes:

   $$\vec{\mathbf{w}}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ -6 \end{bmatrix}, \vec{\mathbf{w}}_2 = \begin{bmatrix} 3 \\ -2 \\ 2 \\ -3 \end{bmatrix}, \vec{\mathbf{w}}_3 = \begin{bmatrix} 3 \\ 1 \\ -2 \\ -2 \end{bmatrix}.$$

   The boss proposes $\vec{\mathbf{y}} = \begin{bmatrix} 1 \\ 1 \\ -2 \\ 0 \end{bmatrix}$, which is also acceptable, on the grounds that "it is simpler."

   Express $\vec{\mathbf{y}}$ as a linear combination of the $\vec{\mathbf{w}}_i$. Then convert the $\vec{\mathbf{w}}_i$ to an orthonormal basis $\vec{\mathbf{v}}_i$ and express $\vec{\mathbf{y}}$ as a linear combination of the $\vec{\mathbf{v}}_i$.

   (b) Find two different solutions to the following set of equations in $\mathbb{Z}_5$:
   $2x + y + 3z + w = 3$
   $3x + 4y + 3w = 1$
   $x + 4y + 2z + 4w = 2$

5 Homework

In working on these problems, you may collaborate with classmates and consult books and general online references. If, however, you encounter a posted solution to one of the problems, do not look at it, and email Paul, who will try to get it removed.

1. By row reducing an appropriate matrix to echelon form, solve the system of equations

$$2x + 4y + z = 2$$
$$3x + y = 1$$
$$3y + 2z = 3$$

over the finite field $\mathbb{Z}_5$. If there is no solution, say so. If there is a unique solution, specify the values of x, y, and z and check your answers. If there is more than one solution, express two of the variables in terms of an arbitrarily chosen value of the third one. For full credit you must reduce the matrix to echelon form, even if the answer becomes obvious!

2. (a) By using elementary matrices, find a vector $\vec{\mathbf{w}}$ that is not in the span of
$\vec{\mathbf{v}}_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$, $\vec{\mathbf{v}}_2 = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}$, and $\vec{\mathbf{v}}_3 = \begin{bmatrix} 2 \\ 4 \\ 0 \end{bmatrix}$

   (b) In the process, you will determine that the given three vectors are linearly dependent. Find a linear combination of them, with the coefficient of $\vec{\mathbf{v}}_3$ equal to 1, that equals the zero vector.

   (c) Find a 1×3 matrix $A = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix}$ such that $A\vec{\mathbf{v}}_1 = A\vec{\mathbf{v}}_2 = A\vec{\mathbf{v}}_3 = 0$, and use it to find the equation of the plane in which the given three vectors lie. You can check your answer to part (a) by showing that $\vec{\mathbf{w}}$ is not in that plane.

   (d) Let $\vec{\mathbf{a}} = A^T$. Compute the dot product of $\vec{\mathbf{a}}$ with the given three vectors and with $\vec{\mathbf{w}}$?. How does this provide another check on your answer to part (a)?

3. This problem illustrates how you can use row reduction to express a specifed vector as a linear combination of basis vectors.

   Your bakery uses flour, sugar, and chocolate to make cookies, cakes, and brownies. The list of ingredients for a batch of each product is described by a vector, as follows:

   Suppose $\vec{\mathbf{v}}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $\vec{\mathbf{v}}_2 = \begin{bmatrix} 4 \\ 2 \\ 7 \end{bmatrix}$, $\vec{\mathbf{v}}_3 = \begin{bmatrix} 7 \\ 8 \\ 11 \end{bmatrix}$.

   This means, for example, that a batch of cookies takes 1 pound of flour, 2 of sugar, 3 of chocolate.

   You are about to shut down for vacation and want to clear out your inventory of ingredients, described by the vector $\vec{\mathbf{w}} = \begin{bmatrix} 21 \\ 18 \\ 38 \end{bmatrix}$.

   Use row reduction to find a combination of cookies, cakes, and brownies that uses up the entire inventory.

4. Hubbard, exercises 2.3.8 and 2.3.11 (column operations: a few brief comments about the first problem will suffice for the second. These column operations will be used in the spring term to evaluate $n \times n$ determinants.)

5. (This result will be needed in Math 23b)

   Suppose that a $2n \times 2n$ matrix T has the following properties:

   - The first n columns are a linearly independent set.
   - The last n columns are a linearly independent set.
   - Each of the first n columns is orthogonal to each of the last n columns.

   Prove that T is invertible.

   Hint: Write $\vec{\mathbf{w}} = a\vec{\mathbf{u}} + \vec{\mathbf{v}}$, where $\vec{\mathbf{u}}$ is a linear combination of the first n columns and $\vec{\mathbf{v}}$ is a linear combination of the last n columns. Start by showing that $\vec{\mathbf{u}}$ is orthogonal to $\vec{\mathbf{v}}$. Then exploit the fact that if $\vec{\mathbf{w}} = \vec{\mathbf{0}}$, $\vec{\mathbf{w}} \cdot \vec{\mathbf{w}} = 0$.

6. (This result will be the key to proving the "implicit function theorem," key to many economic applications.)

   Suppose that $m \times n$ matrix C , where $n > m$, has m linearly independent columns and that these columns are placed on the left. Then we can split off a square matrix A and write $C = [A|B]$.

   (a) Let $\vec{\mathbf{y}}$ be an $(n-m)$-component vector of the "active variables," and let $\vec{\mathbf{x}}$ be the m-component vector of passive variables such that $C \begin{bmatrix} \vec{\mathbf{x}} \\ \vec{\mathbf{y}} \end{bmatrix} = \vec{\mathbf{0}}$.

   Prove that $\vec{\mathbf{x}} = -A^{-1}B\vec{\mathbf{y}}$.

   (b) Use this approach to solve the system of equations

   $5x + 2y + 3z + w = 0$

   $7x + 3y + z - 2w = 0$

   by inverting a 2×2 matrix, without using row reduction or any other elimination technique. The solution will express the "passive" variables x and y in terms of the "active" variables z and w.

7. (If you do the R problems, you get credit for this one automatically and can skip it. Just put a note in your pdf file.)
   Starting by doing row reduction, find a basis for the image and the kernel of the matrix

$$A = \begin{bmatrix} 1 & 2 & 0 & 3 \\ 2 & 0 & 4 & 2 \\ 2 & 1 & 3 & 3 \end{bmatrix},$$

   Then convert the basis for the image to an orthonormal basis.

   The remaining problems are to be solved by writing R scripts. You may use the rref() function whenever it works.

8. One of the seventeen problems on the first Math 25a problem set for 2014 was to find all the solutions of the system of equations

   $2x_1 - 3x_2 - 7x_3 + 5x_4 + 2x_5 = -2$
   $x_1 - 2x_2 - 4x_3 + 3x_4 + x_5 = -2$
   $2x_1 - 4x_3 + 2x_4 + x_5 = 3$
   $x_1 - 5x_2 - 7x_3 + 6x_4 + 2x_5 = -7$

   without the use of a computer.

   Solve this problem using R (like script 1.3A).

9. (Like script 1.3B, but set a finite field, so rref will not help!)
In R, the statement
`A<-matrix(sample(0:4, 24, replace = TRUE),4)`
was used to create a 4×6 matrix A with 24 entries in $\mathbb{Z}_5$. Each entry randomly has the value 0, 1, 2, 3, or 4.

Here is the resulting matrix:

$$A = \begin{bmatrix} 3 & 0 & 4 & 0 & 2 & 2 \\ 1 & 1 & 3 & 3 & 2 & 1 \\ 0 & 2 & 1 & 1 & 4 & 2 \\ 1 & 0 & 2 & 0 & 3 & 4 \end{bmatrix}.$$

Use row reduction to find a basis for the image of A and a basis for the kernel. Please check your answer for the kernel.

10. (Like script 1.3C)A neo-Cubist sculptor wants to use a basis for $\mathbb{R}^3$ with the following properties:

- The first basis vector $w_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ lies along the body diagonal of the cube.
- The second basis vector $w_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ lies along a face diagonal of the cube.
- The second basis vector $w_3 = \begin{bmatrix} 3 \\ 4 \\ 12 \end{bmatrix}$, has length 13.

Convert these three basis vectors to an orthonormal basis. Then make a 3×3 rotation matrix R by using this basis, and confirm that the transpose of R is equal to its inverse.
Note: R must have determinant 1, not -1. You may need to swap columns to make this so.

1. Terminology
   (a) Linear independence: a set of vectors $\{\vec{v_1}, ..., \vec{v_n}\}$ is linearly independent if
   $$\sum_{i=1}^{n} a_i \vec{v_i} = 0 \leftrightarrow \forall i[a_i = 0]$$
   (b) Subspace: a subspace $S \subset V$ is a space that is closed under addition and scalar multiplication—i.e. if $x, y \in S$, so is $ax + by$ for any $a, b \in \mathbb{R}$
      i. Is the set of points (x, y) s.t. $x+y = 0$ a subspace? Prove that it is, or give a counterexample.
      ii. Is the set of points (x, y) s.t. $x+y = 1$ a subspace? Prove that it is, or give a counterexample.
      iii. Is the set of points (x, y) s.t. $y = \sin(x)$ a subspace? Prove that it is, or give a counterexample.
   (c) Span: the span of a set of vectors is the space of all vectors that can be expressed as linear combinations of those vectors, i.e. the set of all $\vec{w}$ s.t. $\exists[a_1, ..., a_n]$ s.t. $\sum_{i=1}^{n} a_i \vec{v_i} = \vec{w}$
   (d) Basis: a basis for a vector space V is a set of linearly independent vectors that span V
   (e) Dimension of a vector space: the number of vectors in any basis for that space (this well-defined!)
   (f) Orthonormal basis: a basis in which all the vectors are unit vectors, and each basis vector is orthogonal to all the other basis vectors
   (g) Image: the subspace of vectors that are possible outputs of a matrix T (a subspace of the **codomain**)
   (h) Rank: the rank of a matrix is the dimension of the image
      i. Can calculate the rank as either the number of independent rows or the number of independent columns! ($rank(A) = rank(A^T)$)
   (i) Kernel: the "zero space" of a matrix, the subspace of vectors that T maps to the zero vector (a subspace of the **domain**)
   (j) Nullity: the dimension of the kernel
      i. THE KERNEL IS NEVER EMPTY!

2. Find a basis for the image and kernel of matrix A below.
   $$A = \begin{bmatrix} 1 & 2 & 1 & 2 & 0 \\ 3 & 6 & 0 & 3 & -3 \\ 0 & 0 & 2 & 2 & 2 \end{bmatrix}$$

3. True/false about images and kernels and ranks and functions
   (a) If A is an $n \times n$ matrix and $A\vec{x} = \vec{0}$, then $x = \vec{0}$.
   (b) If $A\vec{v} = A\vec{w}$, then $\vec{v} - \vec{w} \in ker(A)$.
   (c) If $m > n$, a function $f : \mathbb{R}^m \to \mathbb{R}^n$ cannot be one-to-one.
   (d) If $n > m$, a function $f : \mathbb{R}^m \to \mathbb{R}^n$ cannot be onto.
   (e) A function $f : \mathbb{R}^m \to \mathbb{R}^n$ is onto $\mathbb{R}^n$ if every vector in $\mathbb{R}^m$ maps onto some vector in $\mathbb{R}^n$.
   (f) All functions $f : \mathbb{R}^m \to \mathbb{R}^n$ map linearly independent vectors in $\mathbb{R}^m$ to linearly independent vectors in $\mathbb{R}^n$.
   (g) There exists a 2×2 matrix A such that $rank(A) = 0$.

(h) Let A and B be $n \times n$ matrices. If $\vec{v}$ is in $ker(B)$, then $\vec{v}$ is in $ker(AB)$.

(i) Let A and B be $n \times n$ matrices. If $\vec{v}$ is in $ker(A)$, then $\vec{v}$ is in $ker(AB)$.

(j) If a square matrix has two equal rows, then it is not invertible.

4. Using elementary matrices, find a vector not in the span of $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

5. Walk through how Gram-Schmidt works in the 2-dimensional case

MATHEMATICS 23a/E-23a, Fall 2018
Linear Algebra and Real Analysis I
Week 4 (Eigenvectors and Eigenvalues)

Author: Paul Bamberg
R scripts by Paul Bamberg
Last modified: August 11, 2018 by Paul Bamberg

Reading

- Hubbard, Section 2.7
- Hubbard, pages 474-475

Recorded Lectures

- Lecture 8 (Week 4, Class 1) (watch on October 2 or 3)
- Lecture 9 (Week 4, Class 2) (watch on October 4 or 5)

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

- 4.1 Prove that if $\vec{\mathbf{v}}_1, \cdots, \vec{\mathbf{v}}_n$ are eigenvectors of $A : \mathbb{R}^n \to \mathbb{R}^n$ with distinct eigenvalues $\lambda_1 \cdots \lambda_n$, they are linearly independent. Conclude that an $n \times n$ matrix cannot have more than n distinct eigenvalues.
- 4.2
  - For real $n \times n$ matrix A, prove that if all the polynomials $p_i(t)$ are simple and have real roots, then there exists a basis for $\mathbb{R}^n$ consisting of eigenvectors of A.
  - Prove that if there exists a basis for $\mathbb{R}^n$ consisting of eigenvectors of A, then all the polynomials $p_i(t)$ are simple and have real roots.

  Note - Theorem 2.7.9 in Hubbard is more powerful, because it applies to the complex case. The proof is the same. Our proof is restricted to the real case only because we are not doing examples with complex eigenvectors.

R Scripts

- 1.4A-EigenvaluesCharacteristic.R
  Topic 1 - Eigenvectors for a 2x2 matrix
  Topic 2 - Not every 2x2 matrix has real eigenvalues

- 1.4B-EigenvectorsAxler.R
  Topic 1 - Finding eigenvectors by row reduction
  Topic 2 - Eigenvectors for a 3 x 3 matrix

- 1.4C-Diagonalization.R
  Topic 1: Basis of real eigenvectors
  Topic 2 - Raising a matrix to a power
  Topic 3 - Wnat if the eigenvalues are complex?
  Topic 4 - What if there is no eigenbasis?

- 1.4X-EigenvectorApplications.R
  Topic 1 - The special case of a symmetric matrix
  Topic 2 - Markov Process (from script 1.1D)
  Topic 3 - Eigenvectors for a reflection
  Topic 4 - Sequences defined by linear recurrences

1 Executive Summary

1.1 Eigenvalues and eigenvectors

If $A\vec{\mathbf{v}} = \lambda\vec{\mathbf{v}}$, $\vec{\mathbf{v}}$ is called an eigenvector for A, and λ is the corresponding eigenvalue.

For example, if $A = \begin{bmatrix} -1 & 4 \\ -2 & 5 \end{bmatrix}$, we can check that $\vec{\mathbf{v}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector of A with eigenvalue 3.

If A is a 2×2 or 3×3 matrix, there is a quick, well-known way to find eigenvalues by using determinants.

Rewrite $A\vec{\mathbf{v}} = \lambda\vec{\mathbf{v}}$ as $A\vec{\mathbf{v}} = \lambda I\vec{\mathbf{v}}$, where I is the identity matrix.

Equivalently, $(A - \lambda I)\vec{\mathbf{v}} = \vec{\mathbf{0}}$

Suppose that λ is an eigenvalue of A. Then the eigenvector $\vec{\mathbf{v}}$ is a nonzero vector in the kernel of the matrix $(A - \lambda I)$.

It follows that the matrix $(A - \lambda I)$ is not invertible. But we have a formula for the inverse of a 2×2 or 3×3 matrix, which can fail only if the determinant is zero. Therefore a necessary condition for the existence of an eigenvalue is that $\det(A - \lambda I) = 0$.

The polynomial $\chi_A(\lambda) = \det(A - \lambda I)$ is called the **characteristic polynomial** of matrix A. It is easy to compute in the 2×2 or 3×3 case, where there is a simple formula for the determinant. For larger matrices $\chi_A(\lambda)$ is hard to compute efficiently, and this approach should be avoided.

Conversely, suppose that $\chi_A(\lambda) = 0$ for some real number λ. It follows that the columns of the matrix $(A - \lambda I)$ are linearly dependent. If we row reduce the matrix, we will find at least one nonpivotal column, which in turn implies that there is a nonzero vector in the kernel. This vector is an eigenvector.

This was the standard way of finding eigenvectors until 1995, but it has two drawbacks:

- It requires computation of the determinant of a matrix whose entries are polynomials. Efficient algorithms for calculating the determinant of large square matrices use row-reduction techniques, which might require division by a pivotal element that is a polynomial in λ.

- Once you have found the eigenvalues, finding the corresponding eigenvectors is a nontrivial linear algebra problem.

1.2 Finding eigenvalues - a simple example

Let $A = \begin{bmatrix} -1 & 4 \\ -2 & 5 \end{bmatrix}$. Then $A - \lambda I = \begin{bmatrix} -1-\lambda & 4 \\ -2 & 5-\lambda \end{bmatrix}$
and $\chi_A(\lambda) = \det(A - \lambda I) = (-1-\lambda)(5-\lambda) + 8 = \lambda^2 - 4\lambda + 3$.
Setting $\lambda^2 - 4\lambda + 3 = (\lambda - 1)(\lambda - 3) = 0$, we find two eigenvalues, 1 and 3.
Finding the corresponding eigenvectors still requires a bit of algebra.
For $\lambda = 1, A - \lambda I = \begin{bmatrix} -2 & 4 \\ -2 & 4 \end{bmatrix}$.

By inspection we see that $\vec{\mathbf{v}}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ is in the kernel of this matrix.

Check: $A\vec{\mathbf{v}}_1 = \begin{bmatrix} -1 & 4 \\ -2 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ – eigenvector with eigenvalue 1.

For $\lambda = 3, A - \lambda I = \begin{bmatrix} -4 & 4 \\ -2 & 2 \end{bmatrix}$, and $\vec{\mathbf{v}}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is in the kernel.

Check: $A\vec{\mathbf{v}}_2 = \begin{bmatrix} -1 & 4 \\ -2 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$ – eigenvector with eigenvalue 3.

1.3 A better way to find eigenvectors

Given matrix A, pick an arbitrary vector $\vec{\mathbf{w}}$. Keep computing $A\vec{\mathbf{w}}$, $A^2\vec{\mathbf{w}}$, $A^3\vec{\mathbf{w}}$, etc. until you find a vector that is a linear combination of its predecessors. This situation is easily detected by row reduction.
Now you have found a polynomial p of degree m such that $p(A)\vec{\mathbf{w}} = 0$. Furthermore, this is the nonzero polynomial of lowest degree for which $p(A)\vec{\mathbf{w}} = 0$.

Over the complex numbers, this polynomial is guaranteed to have a root λ by virtue of the "fundamental theorem of algebra" (Hubbard theorem 1.6.13). Over the real numbers or a finite field, it it will have a root in the field only if you are lucky. Assuming that the root exists, factor it out: $p(t) = (t - \lambda)q(t)$.

Now $p(A)\vec{\mathbf{w}} = (A - \lambda I)q(A)\vec{\mathbf{w}} = 0$.

Thus $q(A)\vec{\mathbf{w}}$ is an eigenvector with eigenvalue λ.

Again, let $A = \begin{bmatrix} -1 & 4 \\ -2 & 5 \end{bmatrix}$

As the arbitrary vector $\vec{\mathbf{w}}$ choose $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Then $A\vec{\mathbf{w}} = \begin{bmatrix} -1 \\ -2 \end{bmatrix}$ and $A^2\vec{\mathbf{w}} = \begin{bmatrix} -7 \\ -8 \end{bmatrix}$.

We need to express the third of these vectors, $A^2\vec{\mathbf{w}}$, as a linear combination of the first two. This is done by row reducing the matrix
$\begin{bmatrix} 1 & -1 & -7 \\ 0 & -2 & -8 \end{bmatrix}$ to $\begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 4 \end{bmatrix}$ to find that $A^2\vec{\mathbf{w}} = 4A\vec{\mathbf{w}} - 3I\vec{\mathbf{w}}$.

Equivalently, $(A^2 - 4A + 3I)\vec{\mathbf{w}} = 0$.
$p(A) = A^2 - 4A + 3I$ or $p(t) = t^2 - 4t + 3 = (t - 1)(t - 3)$: eigenvalues 1 and 3.

To get the eigenvector for eigenvalue 1, apply the remaining factor of $p(A)$, $A - 3I$, to $\vec{\mathbf{w}}$: $\begin{bmatrix} -4 & 4 \\ -2 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -4 \\ -2 \end{bmatrix}$. Divide by -2 to get $\vec{\mathbf{v}}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

To get the eigenvector for eigenvalue 3, apply the remaining factor of $p(A)$, $A - I$, to $\vec{\mathbf{w}}$: $\begin{bmatrix} -2 & 4 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ -2 \end{bmatrix}$. Divide by -2 to get $\vec{\mathbf{v}}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

In this case the polynomial $p(t)$ turned out to be the same as the characteristic polynomial, but that is not always the case.

- If we choose $\vec{\mathbf{w}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, we find $A\vec{\mathbf{w}} = 3\vec{\mathbf{w}}, p(A) = A - 3I, p(t) = t - 3$. We need to start over with a different $\vec{\mathbf{w}}$ to find the other eigenvalue.
- If we choose $A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$, then any vector is an eigenvector with eigenvalue 2. So $p(t) = t - 2$. But the characteristic polynomial is $(t - 2)^2$.
- If we choose $A = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$, the characteristic polynomial is $(t-2)^2$. But now there is only one eigenvector. If we choose $\vec{\mathbf{w}} = \vec{\mathbf{e}}_1$ we find $p(t) = t - 2$ and the eigenvector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. But if we choose a different $\vec{\mathbf{w}} = \vec{\mathbf{e}}_2$ we find $p(t) = (t - 2)^2$ and we fail to find a second, independent eigenvector.

1.4 When is there an eigenbasis?

Choose $\vec{\mathbf{w}}$ successively to equal $\vec{\mathbf{e}}_1\vec{\mathbf{e}}_2, \cdots, \vec{\mathbf{e}}_n$.

In searching for eigenvectors, we find successively polynomials $p_1(t), p_2(t), \cdots, p_n(t)$.

There is a basis of real eigenvectors if and only if each of the polynomials $p_i(t)$ has **simple real roots**, e.g. $p(t) = t(t-2)(t+4)(t-2.3)$. No repeated factors are allowed!

A polynomial like $p(t) = t^2 + 1$, although it has no repeated factors, has no real roots: $p(t) = (t+i)(t-i)$.

If we allow complex roots, then any polynomial can be factored into linear factors (Fundamental Theorem of Algebra, Hubbard page113).

There is a basis of complex eigenvectors if and only if each of the polynomials $p_i(t)$ has **simple roots**, e.g. $p(t) = t(t-i)(t+i)$. No repeated factors are allowed!

Our technique for finding eigenvectors works also for matrices over finite fields, but in that case it is entirely possible for a polynomial to have no linear factors whatever. In that case there are no eigenvectors and no eigenbasis. This is one of the few cases where linear algebra over a finite field is fundamentally different from linear algebra over the real or complex numbers.

1.5 Matrix Diagonalization

In the best case we can find a basis of n eigenvectors $\{\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \cdots, \vec{\mathbf{v}}_n\}$ with associated eigenvalues $\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$. Although the eigenvectors must be independent, some of the eigenvalues may repeat.

Create a matrix P whose columns are the eigenvectors. Since the eigenvectors form a basis, they are independent and the matrix P has an inverse P^{-1}

The matrix $D = P^{-1}AP$ is a diagonal matrix.

Proof: $D\vec{\mathbf{e}}_k = P^{-1}A(P\vec{\mathbf{e}}_k) = P^{-1}A\vec{\mathbf{v}}_k = P^{-1}\lambda_k\vec{\mathbf{v}}_k = \lambda_k P^{-1}\vec{\mathbf{v}}_k = \lambda_k\vec{\mathbf{e}}_k$.

The matrix A can be expressed as $A = PDP^{-1}$.

Proof: $A\vec{\mathbf{v}}_k = PD(P^{-1}\vec{\mathbf{v}}_k) = PD\vec{\mathbf{e}}_k = P(\lambda_k\vec{\mathbf{e}}_k) = \lambda_k P\vec{\mathbf{e}}_k = \lambda_k\vec{\mathbf{v}}_k$.

A diagonal matrix D is easy to raise to an integer power.

For example, if $D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$, then $D^k = \begin{bmatrix} \lambda_1^k & 0 & 0 \\ 0 & \lambda_2^k & 0 \\ 0 & 0 & \lambda_3^k \end{bmatrix}$

But now $A = PDP^{-1}$ is also easy to rasie to a power, because $A^k = PD^kP^{-1}$ (will be proved by induction)

The same result extends to kth roots of matrices,
where $B = A^{1/k}$ means that $B^k = A$.

1.6 Properties of an eigenbasis

- Even if all the eigenvalues are distinct, an eigenbasis is not unique. Any eigenvector in the basis can be multiplied by a nonzero scalar and remain an eigenvector.
- Eigenvectors that correspond to distinct eigenvalues are linearly independent (your proof 4.1)
- If the matrix A is symmetric, eigenvectors that correspond to distinct eigenvalues are orthogonal.

1.7 What if there is no eigenbasis?

We consider only the case where A is a 2×2 matrix. If a real polynomial $p(t)$ does not have two distinct real roots, then it either has a repeated real root or it has a pair of conjugate complex roots.

Case 1: Repeated root: $p(t) = (t - \lambda)^2$.

So $p(A) = (A - \lambda I)^2 = 0$.

Set $N = A - \lambda I$, and $N^2 = 0$. The matrix N is called **nilpotent**.

Now $A = \lambda I + N$, and $A^2 = (\lambda I + N)^2 = \lambda^2 I + 2\lambda N$.

It is easy to prove by induction that $A^k = (\lambda I + N)^k = \lambda^k I + k\lambda^{k-1} N$.

Case 2: Conjugate complex roots:

If a 2×2 real matrix A has eigenvalues $a \pm ib$, then it can be expressed in the form $A = PCP^{-1}$, where C is the conformal matrix $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ and P is a change of basis matrix. Since a conformal matrix is almost as easy as a diagonal matrix to raise to the nth power by virtue of De Moivre's theorem $(r(\cos\theta + i\sin\theta))^n = r^n(\cos n\theta + i \sin n\theta)$, this representation is often useful.

Here is an algorithm for constructing the matrices C and P:

Suppose that the eigenvalues of A are $a \pm ib$. Then A has no real eigenvectors, and for any real $\vec{\mathbf{w}}$ we will find the polynomial

$p(t) = (t - a - ib)(t - a + ib) = (t - a)^2 + b^2$

So $p(A) = (A - aI)^2 + b^2 I = 0$ or $(\frac{A-aI}{b})^2 = -I$.

Now we need to construct a new basis, which will not be a basis of eigenvectors but which will still be useful.

Set $\vec{\mathbf{v}}_1 = \vec{\mathbf{e}}_1, \vec{\mathbf{v}}_2 = (\frac{A-aI}{b})\vec{\mathbf{e}}_1$.

Then $(A - aI)\vec{\mathbf{v}}_1 = b\vec{\mathbf{v}}_2$ and $A\vec{\mathbf{v}}_1 = a\vec{\mathbf{v}}_1 + b\vec{\mathbf{v}}_2$.

Also, $(\frac{A-aI}{b})\vec{\mathbf{v}}_2 = (\frac{A-aI}{b})^2\vec{\mathbf{v}}_1 = -\vec{\mathbf{v}}_1$, so

$(A - aI)\vec{\mathbf{v}}_2 = -b\vec{\mathbf{v}}_1$ and $A\vec{\mathbf{v}}_2 = a\vec{\mathbf{v}}_2 - b\vec{\mathbf{v}}_1$.

With respect to the new basis, the matrix that represents A is the conformal matrix $C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$.

If we define P in the usual way with columns $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$, then $A = PCP^{-1}$, and the matrices P and C are real.

1.8 Applications of eigenvectors

- Markov processes

  Suppose that a system can be in one of two or more states and goes through a number of steps, in each of which it may make a transition from one state to another in accordance with specified "transition probabilities."

  For a two-state process, vector $\vec{\mathbf{v}}_n = \begin{bmatrix} p_n \\ q_n \end{bmatrix}$ specifies the probabilities for the system to be in state 1 or state 2 after n steps of the process, where $0 \leq p_n, q_n \leq 1$. and $p_n + q_n = 1$ The transition probabilities are spacified by a matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, where all the entries are between 0 and 1 and $a + c = b + d = 1$.

  After a large number of steps, the state of the system is speciifed by $\vec{\mathbf{v}}_n = A^n \vec{\mathbf{v}}_0$.

  The easy way to calculate A^n is by diagonalizing A. If there is a "stationary state" $\vec{\mathbf{v}}$ into which the system settles down, it corresponds to an eigenvector with eigenvalue 1, since $\vec{\mathbf{v}}_{n+1} = A\vec{\mathbf{v}}_n$ and $\vec{\mathbf{v}}_{n+1} = \vec{\mathbf{v}}_n = \vec{\mathbf{v}}$.

- Reflections

  If 2×2 matrix F represents reflection in a line through the origin with direction vector $\vec{\mathbf{v}}$, then $\vec{\mathbf{v}}$ must be an eigenvector with eigenvalue 1 and a vector perpendicular to $\vec{\mathbf{v}}$ must be an eigenvector with eigenvalue -1.

  If 3×3 matrix F represents reflection in a plane P through the origin with normal vector $\vec{\mathbf{N}}$, then $\vec{\mathbf{N}}$ must be an eigenvector with eigenvalue -1 and there must be a two-dimensional subspace of vectors in P, all with eigenvalue +1.

- Linear recurrences and Fibonacci-like sequences.

  In computer science, it is frequently the case that the first two terms of a sequence, a_0 and a_1, are specified, and subsequent terms are specified by a "linear recurrence" of the form $a_{n+1} = ba_{n-1} + ca_n$. The best-known example is the Fibonacci sequence (Hubbard, pages 220-221) where $a_0 = a_1 = 1$ and $b = c = 1$.

  Then $\begin{bmatrix} a_n \\ a_{n+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ b & c \end{bmatrix} \begin{bmatrix} a_{n-1} \\ a_n \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ b & c \end{bmatrix}^n \begin{bmatrix} a_0 \\ a_1 \end{bmatrix}$.

  The easy way to raise matrix $A = \begin{bmatrix} 0 & 1 \\ b & c \end{bmatrix}$ to the nth power is to diagonalize it.

- Solving systems of linear differential equations

  This topic, of crucial importance to physics, will be covered after we have done some calculus and infinite series.

2 Lecture Outline

1. Using the characteristic polynomial to find eigenvalues and eigenvectors

   If $A\vec{\mathbf{v}} = \lambda\vec{\mathbf{v}}$, $\vec{\mathbf{v}}$ is called an eigenvector for A, and λ is the corresponding eigenvalue.

   If A is a 2×2 or 3×3 matrix, there is a quick, well-known way to find eigenvalues by using determinants.

   Rewrite $A\vec{\mathbf{v}} = \lambda\vec{\mathbf{v}}$ as $A\vec{\mathbf{v}} = \lambda I\vec{\mathbf{v}}$, where I is the identity matrix.

   Equivalently, $(A - \lambda I)\vec{\mathbf{v}} = \vec{\mathbf{0}}$

   Suppose that λ is an eigenvalue of A. Then the eigenvector $\vec{\mathbf{v}}$ is a nonzero vector in the kernel of the matrix $(A - \lambda I)$.

   It follows that the matrix $(A - \lambda I)$ is not invertible. But we have a formula for the inverse of a 2×2 or 3×3 matrix, which can fail only if the determinant is zero. Therefore a necessary condition for the existence of an eigenvalue is that $\det(A - \lambda I) = 0$.

   The polynomial $\chi_A(\lambda) = \det(A - \lambda I)$ is called the **characteristic polynomial** of matrix A. It is easy to compute in the 2×2 or 3×3 case, where there is a simple formula for the determinant. For larger matrices $\chi_A(\lambda)$ is hard to compute efficiently, and this approach should be avoided.

   Conversely, suppose that $\chi_A(\lambda) = 0$ for some real number λ. It follows that the columns of the matrix $(A - \lambda I)$ are linearly dependent. If we row reduce the matrix, we will find at least one nonpivotal column, which in turn implies that there is a nonzero vector in the kernel. This vector is an eigenvector.

   While considering badminton as a Markov process, we constructed the transition matrix $A = \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix}$. Find its eigenvalues and eigenvectors.

2. Finding eigenvectors – a better approach

   This method is guaranteed to succeed only for the field of complex numbers, but the algorithm is valid for any field, and it finds the eigenvectors whenever they exist.

   Given matrix A, pick an arbitrary vector $\vec{\mathbf{w}}$. If you are really lucky, $A\vec{\mathbf{w}}$ is a multiple of $\vec{\mathbf{w}}$ and you have stumbled across an eigenvector. If not, keep computing $A^2\vec{\mathbf{w}}$, $A^3\vec{\mathbf{w}}$, etc. until you find a vector that is a linear combination of its predecessors. This situation is easily detected by row reduction.

   Now you have found a polynomial p of degree m such that $p(A)\vec{\mathbf{w}} = 0$. Furthermore, this is the nonzero polynomial of lowest degree for which $p(A)\vec{\mathbf{w}} = 0$. It is not necessarily the same as the characteristic polynomial.

   Over the complex numbers, this polynomial is guaranteed to have a root λ by virtue of the "fundamental theorem of algebra" (Hubbard theorem 1.6.13). Over the real numbers or a finite field, it it will have a root in the field only if you are lucky. Assuming that the root exists, factor it out:

   $p(t) = (t - \lambda)q(t)$

   Now $p(A)\vec{\mathbf{w}} = (A - \lambda I)q(A)\vec{\mathbf{w}} = 0$.

   Thus $q(A)\vec{\mathbf{w}}$ is an eigenvector with eigenvalue λ.

   Here is a 2×2 example where the calculation is easy.

   Let $A = \begin{bmatrix} -1 & 4 \\ -2 & 5 \end{bmatrix}$

   As the arbitrary vector $\vec{\mathbf{w}}$ choose $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Compute $A\vec{\mathbf{w}}$ and $A^2\vec{\mathbf{w}}$.

Use row reduction to express the third of these vectors, $A^2\vec{\mathbf{w}}$, as a linear combination of the first two.

$$\begin{bmatrix} 1 & -1 & -7 \\ 0 & -2 & -8 \end{bmatrix}$$

Write the result in the form $p(A)\vec{\mathbf{w}} = 0$.

Factor: $p(t)=$

To get the eigenvector for eigenvalue 1, apply the remaining factor of $p(A), A - 3I$, to $\vec{\mathbf{w}}$.

To get the eigenvector for eigenvalue 3, apply the remaining factor of $p(A), A - I$, to $\vec{\mathbf{w}}$.

Citing your source: This technique was brought to the world's attention by Sheldon Axler's 1995 article "Down with Determinants" (see Hubbard page 224). Unlike most of what is taught in undergraduate math, it should probably be cited when you use it in other courses. An informal comment like "Using Axler's method for finding eigenvectors..." would suffice.

3. Eigenvectors and eigenvalues in a finite field

   Consider the matrix $A = \begin{bmatrix} 3 & 2 \\ 3 & 3 \end{bmatrix}$ with entries from the finite field $\mathbb{Z}_5$.

   (a) Find the eigenvalues of A by solving the characteristic equation $\det(A - \lambda I) = 0$, then find the corresponding eigenvectors. Solving a quadratic equation over $\mathbb{Z}_5$ is easy – in a pinch, just try all five possible roots!

   (b) Find the eigenvalues of A by using the technique of example 2.7.8 of Hubbard. You will get the same equation for the eigenvalues, of course, but it will be more straightforward to find the eigenvectors.

   (c) Write down the matrix P whose columns are the basis of eigenvectors, and show that $P^{-1}AP$ is a diagonal matrix. Why is this reasonable?

4. When is there an eigenbasis?

   Choose $\vec{\mathbf{w}}$ successively to equal $\vec{\mathbf{e}}_1\vec{\mathbf{e}}_2, \cdots, \vec{\mathbf{e}}_n$.

   In searching for eigenvectors, we find successively polynomials $p_1(t), p_2(t), \cdots, p_n(t)$.

   There is a basis of real eigenvectors if and only if each of the polynomials $p_i(t)$ has **simple real roots**, e.g. $p(t) = t(t-2)(t+4)(t-3)$. No repeated factors are allowed!

   A polynomial like $p(t) = t^2 + 1$, although it has no repeated factors, has no real roots: $p(t) = (t+i)(t-i)$.

   If we allow complex roots, then any polynomial can be factored into linear factors (Fundamental Theorem of Algebra, Hubbard page 113).

   There is a basis of complex eigenvectors if and only if each of the polynomials $p_i(t)$ has **simple roots**, e.g. $p(t) = t(t-i)(t+i)$. No repeated factors are allowed!

   Here is a clever way to construct a matrix for which one of the polynomials $p_i(t)$ does not have simple roots and there is no basis of eigenvectors.

   Let $D = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$, $N = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$. The matrix N is a so-called "nilpotent" matrix: because its kernel is the same as its image, N^2 is the zero matrix.

   Show that the matrix $A = D + N$ has the property that if we choose any $\vec{\mathbf{w}}$ that is not in the kernel of N, then the polynomial $p(A)$ is $(A-2I)^2$ and so there is no basis of eigenvectors.

5. The easy case: n distinct real eigenvalues (Proof 4.1)

   If $\vec{\mathbf{v}}_1, \cdots, \vec{\mathbf{v}}_n$ are eigenvectors of $A : \mathbb{R}^n \to \mathbb{R}^n$ with distinct eigenvalues $\lambda_1 \cdots \lambda_n$, they are linearly independent.

   This proof could be done by induction, but there is an equivalent technique, using the "least number principle," that is a little bit easier.

   Suppose, for a contradiction, that the eigenvectors are linearly dependent.

   There exists a first eigenvector (the jth one) that is a linear combination of its predecessors:

$$\vec{\mathbf{v}}_j = a_1\vec{\mathbf{v}_1} + \cdots + a_{j-1}\vec{\mathbf{v}}_{j-1}.$$

   Multiply both sides by $A - \lambda_j I$. You get zero on the left, and on the right you get a linear combination where all the coefficients are nonzero because $\lambda_j - \lambda_i \neq 0$. This is in contradiction to the assumption that $\vec{\mathbf{v}}_j$ was the first one that is a linear combination of its predecessors.

   Since in $\mathbb{R}^n$ there cannot be more than n linearly independent vectors, there are at most n distinct eigenvalues.

6. Change of basis

Our "old" basis consists of the standard basis vectors $\vec{\mathbf{e}}_1$ and $\vec{\mathbf{e}}_2$.

Our "new" basis consists of one eigenvector for each eigenvalue of $A = \begin{bmatrix} -1 & 4 \\ -2 & 5 \end{bmatrix}$, with eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 3$.

Let's choose $\vec{\mathbf{v}}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\vec{\mathbf{v}}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

It would be all right to multiply either of these vectors by a constant or to reverse their order.

Write down the change of basis matrix P whose columns express the new basis vectors in term of the old ones.

Calculate the inverse change of basis matrix P^{-1} whose columns express the old basis vectors in terms of the new ones.

We are considering a linear transformation that is represented, relative to the standard basis, by the matrix A. What diagonal matrix D represents this linear transformation relative to the new basis of eigenvectors?

Confirm that $A = PDP^{-1}$. We have "diagonalized" the matrix A.

$$\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

7. Diagonalization and eigenvectors

Let P be a matrix whose columns are linearly independent eigenvectors of the $n \times n$ matrix A. (Such a matrix does not exist for every matrix A.) Denote by λ_i the eigenvalue corresponding to the eigenvector that is column i.

How do we know that the columns of P form a basis for $\mathbb{R}^n$?

Let $D = P^{-1}AP$.

Prove that D is the diagonal matrix

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

.

by applying $D = P^{-1}AP$ to standard basis vectors.

Conversely, let D have the specified form, and prove that $A\vec{\mathbf{v}}_i = \lambda_i \vec{\mathbf{v}}_i$

by applying $A = PDP^{-1}$ to eigenvectors.

8. Eigenvectors for a 3×3 matrix
   For Hubbard Example 2.7.8, the calculation is best subcontracted to R. The matrix is

$$A = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

   Since we have help with the computation, make the choice $\vec{\mathbf{w}} = \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}$.

   The matrix to row reduce is

$$\begin{bmatrix} 2 & -1 & 0 & 3 \\ 3 & -1 & -3 & -9 \\ 5 & 2 & 3 & 6 \end{bmatrix}$$

, different from the matrix in Hubbard.

   The result of row reduction is the same:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 4 \end{bmatrix}$$

   The rest of the work is easily done by hand.

   Using the last column, write the polynomial $p(t)$, and factor it.

   Find an eigenvector that corresponds to the smallest positive eigenvalue. It is not necessary to use the same $\vec{\mathbf{w}}$; any vector will do, as long as it is not in the subspace spanned by the other eigenvectors. Hubbard uses $\vec{\mathbf{e}}_1$. Use $\vec{\mathbf{e}}_3$ instead.

9. When is there an eigenbasis?

   This is a difficult issue in general. The simple case is where we are lucky and find a polynomial p of degree n that has n distinct roots. In that case we can find n eigenvectors, and it has already been proved that they are linearly independent. They form an eigenbasis. If the roots are real, the eigenvectors are elements of $\mathbb{R}^n$. If the roots are distinct but not all real, the eigenvectors are still a basis of $\mathbb{C}^n$.

   Suppose we try each standard basis vector in turn as $\vec{\mathbf{w}}$. Using $\vec{\mathbf{e}}_i$ leads to a polynomial p_i. If every p_i is a polynomial of degree $m_i < n$, the situation is more complicated. Theorem 2.7.9 in Hubbard states the result:

   There exists an eigenbasis of $\mathbb{C}^n$ if and only if all the roots of all the p_i are simple.

   Before doing the difficult proof, look the simplest examples of matrices that do not have n distinct eigenvalues.

   - Let $A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$. In this case every vector in $\mathbb{R}^2$ is an eigenvector with eigenvalue 2. There is only one eigenvalue, but any basis is an eigenbasis.

     If we choose $\vec{\mathbf{w}} = \vec{\mathbf{e}}_1$ and form the matrix whose columns are $\vec{\mathbf{w}}$ and $A\vec{\mathbf{w}}$,
     $\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$,

     the matrix is already in echelon form.

     What is p_1?

     What eigenvector do we find?

     What eigenvector do we find if we choose $\vec{\mathbf{w}} = \vec{\mathbf{e}}_2$?

     Key point: we found a basis of eigenvectors, even though there was only one eigenvalue, and the polynomial $(t-2)^2$ never showed up.

- Let $A = \begin{bmatrix} 2 & 0 \\ 1 & 2 \end{bmatrix}$. In this case there is only one eigenvalue and there is no eigenbasis.
  What happens if we choose $\vec{\mathbf{w}} = \vec{\mathbf{e}}_2$?

  If we choose $\vec{\mathbf{w}} = \vec{\mathbf{e}}_1$,
  confirm that $\begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & 4 \end{bmatrix}$
  row reduces to $\begin{bmatrix} 1 & 0 & -4 \\ 0 & 1 & 4 \end{bmatrix}$.

  What is p_1?

  What happens when we carry out the procedure that usually gives an eigenvector?

  Key point: There was only one eigenvalue, the polynomial $(t-2)^2$ showed up, and we were unable to find a basis of eigenvectors.

10. An instructive 3×3 example

The surprising case, and the one that makes the proof difficult, is the one where there exists a basis of eigenvectors but there are fewer than n distinct eigenvalues. A simple example is $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

Here each standard basis vector is an eigenvector. For the first one the eigenvalue is 1; for the second and third, it is 2.

A less obvious example is

$A = \begin{bmatrix} 2 & 1 & -1 \\ 0 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix}$

When we use row reduction to find the eigenvectors,we obtain the following results:

Using $\vec{\mathbf{w}} = \vec{\mathbf{e}}_1$, we get $p_1(t) = t - 2$ and find an eigenvector

$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ with eigenvalue 2.

Using $\vec{\mathbf{w}} = \vec{\mathbf{e}}_2$, we get $p_2(t) = (t-1)(t-2)$ and find two eigenvectors:

$\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ with eigenvalue 1, $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ with eigenvalue 2.

At this point we have found three linearly independent eigenvectors and we have a basis.

If we use $\vec{\mathbf{w}} = \vec{\mathbf{e}}_3$, we get $p_3(t) = (t-1)(t-2)$ and find two eigenvectors:

$\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ with eigenvalue 1, $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ with eigenvalue 2.

In general, if we use some arbitrary $\vec{\mathbf{w}}$, we will get $p(t) = (t-1)(t-2)$ and we will find the eigenvector with eigenvalue 1 along with some linear combination of the eigenvectors with eigenvalue 2.

Key points about this case:

- The polynomial $p_i(t)$, in order to be simple, must have degree less than n.
- We need to use more than one standard basis vector in order to find a basis of eigenvectors.

11. Proof that if all roots are simple there is an eigenbasis

    Assume that whenever we choose $\vec{\mathbf{w}} = \vec{\mathbf{e}}_i$, the polynomial p_i of degree m_i has simple roots. The columns of the matrix that we row reduce are $\vec{\mathbf{e}}_i, A\vec{\mathbf{e}}_i, \cdots A^{m_i}\vec{\mathbf{e}}_i$. The image of this matrix has three properties.

    - It is a subspace E_i of $\mathbb{R}^n$.
    - It includes m_i eigenvectors. Since these correspond to distinct eigenvalues, they are linearly independent, and therefore they span E_i.
    - It includes $\vec{\mathbf{e}}_i$.

    Now take the union of all the E_i. This union has the following properties:

    - It includes each standard basis vector $\vec{\mathbf{e}}_i$, so it is all of $\mathbb{R}^n$.
    - It is spanned by the union of the sets of eigenvectors. In general there will be more than n vectors in this set. Use them as columns of a matrix. The image of this matrix is all of $\mathbb{R}^n$. We can find a basis for the image consisting of n columns, which are all eigenvectors.

12. Proof that if there is an eigenbasis, each p_i has simple roots.

    There are k distinct eigenvalues, $\lambda_1, \cdots, \lambda_k$. It is entirely possible that $k < n$, since different eigenvectors may have the same eigenvalue.

    Since there is a basis of eigenvectors, we can express each $\vec{\mathbf{e}}_i$ as a linear combination of eigenvectors.

    Define $p_i(t) = \prod(t - \lambda_j)$. The product extends just over the set of eigenvalues that are associated with the eigenvectors needed to express $\vec{\mathbf{e}}_i$ as a linear combination, so there may be fewer than k factors.

    Form $p_i(A) = \prod(A - \lambda_j I)$. The factors can be in any order. If $\vec{\mathbf{w}}$ is any eigenvector whose eigenvalue λ_j is included in the product, then $(A - \lambda_j I)\vec{\mathbf{w}} = 0$ and so $p_i(A)\vec{\mathbf{w}} = \vec{\mathbf{0}}$. Since those eigenvectors from a basis for a subspace that includes $\vec{\mathbf{e}}_i$, it follows that $p_i(A)\vec{\mathbf{e}}_i = \vec{\mathbf{0}}$.

    If we form a nonzero polynomial $p_i'(t)$ of lower degree by omitting one factor from the product, then $p_i'(A)\vec{\mathbf{e}}_i \neq \vec{\mathbf{0}}$, since the eigenvectors that correspond to the omitted eigenvalue do not get killed off.

    So $p_i(t)$ is the nonzero polynomial of lowest degree for which $p_i(A)\vec{\mathbf{e}}_i = \vec{\mathbf{0}}$, and by construction it has simple roots.

13. Proof 4.2, first half

    Assume that whenever we choose $\vec{\mathbf{w}} = \vec{\mathbf{e}}_i$, the polynomial p_i of degree m_i has simple real roots. Consider the subspace E that is the image of the matrix whose columns are
    $\vec{\mathbf{e}}_1, A\vec{\mathbf{e}}_1, \cdots A^{m_1}\vec{\mathbf{e}}_1, \vec{\mathbf{e}}_2, A\vec{\mathbf{e}}_2, \cdots A^{m_2}\vec{\mathbf{e}}_2, \cdots, \vec{\mathbf{e}}_n, A\vec{\mathbf{e}}_n, \cdots A^{m_n}\vec{\mathbf{e}}_n$.
    Prove that $E = \mathbb{R}^n$ (easy) and that there exists a basis for E that consists entirely of eigenvectors(harder).

    Theorem 2.7.9 in Hubbard is more powerful, because it applies to the complex case. The proof is the same. Our proof is restricted to the real case only because we are not doing examples with complex eigenvectors.

14. Proof 4.2, second half

    Assume that there is a basis of $\mathbb{R}^n$ consisting of eigenvectors of $n \times n$ matrix A, but that A has only $k \leq n$ distinct real eigenvalues. Prove that for any basis vector $\vec{\mathbf{w}} = \vec{\mathbf{e}}_i$, the polynomial $p_i(t)$ has simple roots.

15. Fibonacci numbers by matrices

The usual way to generate the Fibonacci sequence is to set $a_0 = 1, a_1 = 1$, then calculate $a_2 = a_0 + a_1 = 2$, $a_3 = a_1 + a_2 = 3$, etc.

In matrix notation this can be written

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

and more generally

$$\begin{bmatrix} a_n \\ a_{n+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^n \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Use this approach to determine a_2 and a_3, doing the matrix multiplication first.

$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

Determine a_6 and a_7 by using the square of the matrix that was just constructed.

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

We have found a slight computational speedup, but it would be nicer to have a general formula for $\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^n$.

16. Powers of a diagonal matrix.

For a 2×2 diagonal matrix, $\begin{bmatrix} c_1 & 0 \\ 0 & c_2 \end{bmatrix}^n = \begin{bmatrix} c_1^n & 0 \\ 0 & c_2^n \end{bmatrix}$.

Now suppose that we want to compute A^n. If there is a basis of eigenvectors, we can construct the matrix P, whose columns are eigenvectors, such that

$$P^{-1}AP = \begin{bmatrix} c_1 & 0 \\ 0 & c_2 \end{bmatrix} = D.$$

Prove by induction that

$$(P^{-1}AP)^n = P^{-1}A^nP.$$

For the Fibonacci example, where $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$,

$D \begin{bmatrix} (1+\sqrt{5})/2 & 0 \\ 0 & (1+\sqrt{5})/2 \end{bmatrix}$ and $P = \begin{bmatrix} 2 & 2 \\ 1+\sqrt{5} & 1-\sqrt{5} \end{bmatrix}$

You can check that $PD^2P^{-1} = A^2$
(messy because of the irrational numbers!)
Example 2.7.1 in the textbook computes $A^n = PD^nP^{-1}$ and gets a formula for the nth Fibonacci number that is useful in computer science:

$$a^n = \frac{5+\sqrt{5}}{10}(\frac{1+\sqrt{5}}{2})^n + \frac{5-\sqrt{5}}{10}(\frac{1-\sqrt{5}}{2})^n$$

17. Dealing with complex eigenvalues in the 2×2 case

If the polynomial $p(t)$ for a matrix A has two distinct complex roots, it is possible to compute A^n by diagonalizing the matrix and calculating $A^n = PD^nP^{-1}$. However, both the eigenvectors and eigenvalues will be complex, and we know that A^n is a real matrix. It seems a shame to have to do complex arithmetic to compute a matrix of real numbers.

Fortunately, there is another type of 2×2 matrix that is almost as easy to raise to a power as a diagonal matrix – a conformal matrix.

Recall that the matrix $C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} r\cos\theta & -r\sin\theta \\ r\sin\theta & r\cos\theta \end{bmatrix}$ represents the complex number $a + bi = re^{i\theta}$.

Its nth power is $C^n = \begin{bmatrix} r^n\cos n\theta & -r^n\sin n\theta \\ r^n\sin n\theta & r^n\cos n\theta \end{bmatrix}$.

Given a 2×2 matrix A with complex eigenvalues $a \pm bi$, we just need to invent a matrix P such that $A = PCP^{-1}$.

Here is an algorithm for constructing the matrices C and P:

Suppose that the eigenvalues of A are $a \pm ib$, with $b \neq 0$. Then no real $\vec{\mathbf{w}}$ can be an eigenvector, and for any real $\vec{\mathbf{w}}$ we will find the polynomial

$p(t) = (t - a - ib)(t - a + ib) = (t - a)^2 + b^2$

So $p(A) = (A - aI)^2 + b^2I = 0$ or

$$(\frac{A - aI}{b})^2 = -I.$$

Now we need to construct a new basis, which will not be a basis of eigenvectors but which will still be useful.

Set $\vec{\mathbf{v}}_1 = \vec{\mathbf{e}}_1, \vec{\mathbf{v}}_2 = (\frac{A-aI}{b})\vec{\mathbf{e}}_1$.

Then $(A - aI)\vec{\mathbf{v}}_1 = b\vec{\mathbf{v}}_2$ and $A\vec{\mathbf{v}}_1 = a\vec{\mathbf{v}}_1 + b\vec{\mathbf{v}}_2$.

Also, $(\frac{A-aI}{b})\vec{\mathbf{v}}_2 = (\frac{A-aI}{b})^2\vec{\mathbf{v}}_1 = -\vec{\mathbf{v}}_1$, so

$(A - aI)\vec{\mathbf{v}}_2 = -b\vec{\mathbf{v}}_1$ and $A\vec{\mathbf{v}}_2 = a\vec{\mathbf{v}}_2 - b\vec{\mathbf{v}}_1$.

With respect to the new basis, the matrix that represents A is the conformal matrix $C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$.

If we define P in the usual way with columns $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$, then $A = PCP^{-1}$, and the matrices P and C are real.

18. Conformal matrices and complex numbers – an example

(a) Show that the polynomial $p(t)$ for the matrix $A = \begin{bmatrix} 7 & -10 \\ 2 & -1 \end{bmatrix}$ has roots $3 \pm 2i$.

(b) Show that $(\frac{A-3I}{2})^2 = -I$.

(c) Choose a new basis with $\vec{\mathbf{v}}_1 = \vec{\mathbf{e}}_1, \vec{\mathbf{v}}_2 = (\frac{A-3I}{2})\vec{\mathbf{e}}_1$.
Use these basis vectors as the columns of matrix P.
Confirm that $A = PCP^{-1}$, where C is conformal and P is real.

19. What to do if there is no basis of eigenvectors

    In the 2×2 case, suppose there is no basis of eigenvectors, just a single eigenvector $\vec{\mathbf{v}}$ with eigenvalue λ. If we choose $\vec{\mathbf{w}} \neq \vec{\mathbf{v}}$, we will get a polynomial whose roots are not simple: $p(t) = (t - \lambda)^2$.

    So $p(A) = (A - \lambda I)^2 = 0$. In other words, $N = A - \lambda I$ is a nilpotent matrix, with $N^2 = 0$.

    It is easy to raise $A = \lambda I + N$ to the nth power.

    Prove by induction that
    $A^k = (\lambda I + N)^k = \lambda^k I + k\lambda^{k-1} N$.

    As an example, let $A = \begin{bmatrix} 5 & -1 \\ 4 & 1 \end{bmatrix}$ and compute A^4

20. What about non-integer powers?

    If $A = PDP^{-1}$ or $A = PCP^{-1}$ (basis of real or complex eigenvectors), then it is possible to raise A to a non-integer power! For example, you can compute $B = A^{\frac{1}{2}}$ so that $B^2 = A$.. Be careful, though – there may be many answers or none.

    - In the diagonal case where $A = PDP^{-1}$ the matrix D must have non-negative entries (the eigenvalues cannot be negative.) For a positive eigenvalue, there are two possible square roots.
    - In the conformal case where $A = PCP^{-1}$ the conformal matrix C has two different square roots, one with angle $\frac{1}{2}\theta$ and one with angle $\frac{1}{2}\theta + \pi$.

21. A couple of short but interesting proofs

Suppose that S is a symmetric matrix with n distinct real eigenvalues. Prove that you can construct an orthonormal basis of eigenvectors.

Suppose that matrix AB has a nonzero eigenvalue λ. Prove that λ is also an eigenvalue for BA.

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. Define *eigenvector* and *eigenvalue* for a matrix A, and prove that λ (which may be a complex number) is an eigenvalue of A if and only if it is a root of the characteristic polynomial $\chi_A(\lambda) = \det(A - \lambda I)$.

2. (Proof 4.1)
   Prove that if $\vec{\mathbf{v}}_1, \cdots, \vec{\mathbf{v}}_n$ are eigenvectors of $A : \mathbb{R}^n \to \mathbb{R}^n$ with distinct eigenvalues $\lambda_1 \cdots \lambda_n$, they are linearly independent. Conclude that an $n \times n$ matrix cannot have more than n distinct eigenvalues.

3. Suppose that for $n \times n$ matrix A there is a basis of n eigenvectors $\{\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \cdots, \vec{\mathbf{v}}_n\}$ with associated eigenvalues $\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$. Although the eigenvectors must be independent, some of the eigenvalues may repeat.

   Create a matrix P whose columns are the eigenvectors. Since the eigenvectors form a basis, they are independent and the matrix P has an inverse P^{-1}. Prove the following:

   (a) Matrix P is invertible.

   (b) The matrix $D = P^{-1}AP$ is a diagonal matrix. .

   (c) The matrix A can be expressed as $A = PDP^{-1}$.

   (d) The matrix A^n can be expressed as $A^n = PD^nP^{-1}$ (use induction; do not write $PDP^{-1} \cdots PDP^{-1}$.)

4. For the matrix $A = \begin{bmatrix} 3 & 2 \\ 0 & 3 \end{bmatrix}$, apply our standard technique for finding eigenvectors, first choosing $\vec{\mathbf{w}} = \vec{\mathbf{e}}_1$, then choosing $\vec{\mathbf{w}} = \vec{\mathbf{e}}_2$. Show that $p_2(t)$ does not have simple roots, that you find only one eigenvector, and that A cannot be "diagonalized" (written in the form $A = PDP^{-1}$.)

   Write A in the form $A = \lambda I + N$, where N is called "nilpotent" because its square is the zero matrix. Prove by induction that there is still a simple formula for the powers of A, namely $A^n = \lambda^n I + n\lambda^{n-1}N$.

5. (Proof 4.2)

   - For real $n \times n$ matrix A, prove that if all the polynomials $p_i(t)$ are simple and have real roots, then there exists a basis for $\mathbb{R}^n$ consisting of eigenvectors of A.
   - Prove that if there exists a basis for $\mathbb{R}^n$ consisting of eigenvectors of A, then all the polynomials $p_i(t)$ are simple and have real roots.

4 Workshop Problems

1. Some interesting examples with 2×2 matrices

   (a) Since a polynomial equation with real (or complex) coefficients always has a root (the "fundamental theorem of algebra"), a real matrix is guaranteed to have at least one complex eigenvalue. No such theorem holds for polynomial equations with coefficients in a finite field, so zero eigenvalues is a possibility. This is one of the few results in linear algebra that depends on the underlying field.

   Consider the matrix $A = \begin{bmatrix} 3 & 1 \\ n & 3 \end{bmatrix}$ with entries from the finite field $\mathbb{Z}_5$.

   By considering the characteristic equation, find values of n that lead to 2, 1, or 0 distinct eigenvalues. For the case of 1 eigenvalue, find an eigenvector.

   Hint: After writing the characteristic equation with n isolated on the right side of the equals sign, make a table of the value of $t^2 + 4t + 4$ for each of the five possible eigenvalues. That table lets you determine how many solutions there are for each of the five possible values of n. When the characteristic polynomial is the square of a linear factor, there is only one eigenvector and it is easy to construct.

   (b) Extracting square roots by diagonalization.

   The matrix $A = \begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix}$

   conveniently has two eigenvalues that are perfect squares. Find a basis of eigenvectors and construct a matrix P such that $P^{-1}AP$ is a diagonal matrix.

   Thereby find two independent square roots of A, i.e. find matrices B_1 and B_2 such that $B_1^2 = B_2^2 = A$, with $B_2 \neq \pm B_1$. Hint: use the negative square root of one of the eigenvalues, the positive square root of the other.

   If you take Physics 15c next year, you may encounter this technique when you study "coupled oscillators."

2. Some proofs. In doing these, you may use the fact that an eigenbasis exists if and only if all the $p_i(t)$ have simple roots.

   (a) Suppose that a 5×5 matrix has a basis of eigenvectors, but that its only eigenvalues are 1 and 2. Using Hubbard Theorem 2.7.9, show that you must make at least three different choices of $\vec{\mathbf{e}}_i$ in order to find all the eigenvectors.

   (b) An alternative approach to proof 4.1 – use induction.

   Identify a base case (easy). Then show that, if a set of $k-1$ eigenvectors with distinct eigenvalues is linearly independent and you add to the set an eigenvector $\vec{\mathbf{v}}_k$ with an eigenvalue λ_k that is different from any of the preceding eigenvalues, the resulting set of k eigenvectors with distinct eigenvalues is linearly independent.

3. Examples where there is no basis of eigenvectors

   (a) The matrix $A = \begin{bmatrix} 1 & -1 \\ 4 & -3 \end{bmatrix}$ has only a single eigenvalue and only one independent eigenvector.

   Find the eigenvalue and eigenvector, show that $A = D + N$ where D is diagonal and N is nilpotent, and use the formula from seminar topic 4 to calculate A^3 without ever multiplying A by itself (unless you want to check your answer).

   (b) Find two eigenvectors for the matrix $A = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ -2 & 2 & 0 \end{bmatrix}$. and confirm that using each of the three standard basis vectors in turn will not produce a third independent eigenvector.

   Clearly the columns of A are not independent; so 0 is an eigenvalue. This property makes the algebra fairly easy.

4. Problems with 3×3 matrices, to be solved by writing or editing R scripts

   (a) Sometimes you don't find all the eigenvectors on the first try.
   The matrix $A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
   has three real, distinct eigenvalues, and there is a basis of eigenvectors. Find what polynomial equation for the eigenvalues arises from each of the following choices, and use it to construct as many eigenvectors as possible.:

   - $\vec{\mathbf{w}} = \vec{\mathbf{e}_1}$.
   - $\vec{\mathbf{w}} = \vec{\mathbf{e}_3}$.
   - $\vec{\mathbf{w}} = \vec{\mathbf{e}_1} + \vec{\mathbf{e}_3}$.

   (b) Use the technique of example 2.7.8 in Hubbard to find the eigenvalues and eigenvectors of the matrix $A = \begin{bmatrix} 3 & 4 & -4 \\ 1 & 3 & -1 \\ 3 & 6 & -4 \end{bmatrix}$

5 Homework

1. Consider the sequence of numbers described, in a manner similar to the Fibonacci numbers, by

   $b_3 = 2b_1 + b_2$
   $b_4 = 2b_2 + b_3$

   $b_{n+2} = 2b_n + b_{n+1}$

   (a) Write a matrix B to generate this sequence in the same way that Hubbard generates the Fibonacci numbers.

   (b) By considering the case $b_1 = 1, b_2 = 2$ and the case $b_1 = -1, b_2 = 1$, find the eigenvectors and eigenvalues of B.

   (c) Express the vector $\begin{bmatrix}1\\1\end{bmatrix}$ as a linear combination of the two eigenvectors, and thereby find a formula for b_n if $b_1 = 1, b_2 = 1$.

2. (This is similar to group problem 1b.)

   Consider the matrix $A = \begin{bmatrix}-10 & 9\\-18 & 17\end{bmatrix}$.

   (a) By using a basis of eigenvectors, find a matrix P such that $P^{-1}AP$ is a diagonal matrix.

   (b) Find a cube root of A, i.e. find a matrix B such that $B^3 = A$.

3. (a) Prove that if $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$ are eigenvectors of matrix A, both with the same eigenvalue λ, then any linear combination of $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$ is also an eigenvector.

   (b) Suppose that A is a 3×3 matrix with a basis of eigenvectors but with only two distinct eigenvalues. Prove that for any $\vec{\mathbf{w}}$, the vectors $\vec{\mathbf{w}}$, $A\vec{\mathbf{w}}$, and $A^2\vec{\mathbf{w}}$ are linearly dependent. (This is another way to understand why all the polynomials $p_i(t)$ are simple when A has a basis of eigenvectors but a repeated eigenvalue.)

4. Harvard graduate Ivana Markov, who concentrated in English and mathematics with economics as a secondary field, just cannot decide whether she wants to be a poet or an investment banker, and so her career path is described by the following Markov process:

   - If Ivana works as a poet in year n, there is a probability of 0.9 that she will feel poor at the end of the year and take a job as an investment banker for year $n+1$. Otherwise she remains a poet.
   - If Ivana works as an investment banker in year n, there is a probability of 0.7 that she will feel overworked and unfulfilled at the end of the year and take a job as a poet for year $n+1$. Otherwise she remains an investment banker.

   Thus, if $\begin{bmatrix} p_n \\ q_n \end{bmatrix}$ describes the probabilities that Ivana works as a poet or a banker respectively in year n, the corresponding probabilities for year $n+1$ are given by $\begin{bmatrix} p_{n+1} \\ q_{n+1} \end{bmatrix} = A \begin{bmatrix} p_n \\ q_n \end{bmatrix}$, where $A = \begin{bmatrix} 0.1 & 0.7 \\ 0.9 & 0.3 \end{bmatrix}$

   (a) Find the eigenvalues and eigenvectors of A.

   (b) Construct the matrix P whose columns are the eigenvectors, invert it, and thereby express the vector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ as a linear combination of the eigenvectors.

   (c) Suppose that in year 0 Ivana works as a poet, so that $\begin{bmatrix} p_0 \\ q_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Find an explicit formula for $\begin{bmatrix} p_n \\ q_n \end{bmatrix}$ and use it to determine $\begin{bmatrix} p_{10} \\ q_{10} \end{bmatrix}$. What happens in the limit of large n?

5. In general, the square matrix A that represents a Markov process has the property that all the entries are between 0 and 1 and each column sums to 1. Prove that such a matrix A has an eigenvalue of 1 and that there is a "stationary vector" that is transformed into itself by A. You may use the fact, which we have proved so far only for 2×2 and 3×3 matrices, that if a matrix has a nonzero vector in its kermel, its determinant is zero.

6. (a) Prove by induction (no "$\cdots$" allowed!) that if $F = PCP^{-1}$, then $F^n = PC^nP^{-1}$ for all positive integers n.

   (b) Suppose that 2×2 real matrix F has complex eigenvalues $re^{\pm i\theta}$. Show that, for integer n, F^n is a multiple of the identity matrix if and only if $n\theta = m\pi$ for some integer m. Hint: write $F = PCP^{-1}$ where C is conformal. This hint also helps with the rest of the problem.

   (c) If $F = \begin{bmatrix} 3 & 7 \\ -1 & -1 \end{bmatrix}$, find the smallest n for which F^n is a multiple of the identity. Check your answer by matrix multiplication.

   (d) If $G = \begin{bmatrix} -2 & -15 \\ 3 & 10 \end{bmatrix}$, use half-angle formulas to find a matrix A for which $A^2 = G$. Check your answer by matrix multiplication.

7. (You may use this problem as a third R script problem, in which case you will automatically get credit for it as an ordinary homework problem)

   Use the technique of example 2.7.8 in Hubbard to find the eigenvalues and eigenvectors of the following two matrices. One has a repeated eigenvalue and will require you to use the technique with two different basis vectors.

   (a) $A = \begin{bmatrix} 3 & 4 & -4 \\ 1 & 3 & -1 \\ 3 & 6 & -4 \end{bmatrix}$

   (b) $B = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 3 & -1 \\ 1 & 2 & 0 \end{bmatrix}$

   Optional problems that require writing or editing R scripts

8. The matrix $A = \begin{bmatrix} 5 & 1 & 1 \\ -1 & 3 & -1 \\ 0 & 0 & 4 \end{bmatrix}$ has only one eigenvalue, 4, and so its characteristic polynomial must be $(t-4)^3$.

   (a) Show that A has a two-dimensional subspace of eigenvectors but that there is no other eigenvector.

   (b) Write $A = D+N$ where D is diagonal and N is nilpotent, and confirm that N^2 is the zero matrix.

9. Here is a symmetric matrix, which is guaranteed to have an orthonormal basis of eigenvectors. For once, the numbers have not been rigged to make the eigenvalues be integers.

$$A = \begin{bmatrix} 4 & -1 & 1 \\ -1 & 3 & 2 \\ 1 & 2 & -3 \end{bmatrix}$$

Express A in the form PDP^{-1}, where D is diagonal and P is an isometry matrix whose columns are orthogonal unit vectors.

A similar example is in script 1.4X.

10. Problem 7 above.

1. Finding eigenvectors and eigenvalues

   (a) Using the characteristic polynomial method, find the eigenvectors and eigenvalues of

   $$A = \begin{bmatrix} 1 & 2 \\ -1 & 4 \end{bmatrix}$$

   (b) Use Axler's method to find the eigenvectors and eigenvalues of A again.

2. Some quick useful facts and special cases of eigenvectors and eigenvalues

   (a) A triangular matrix has eigenvalues along its main diagonal:

   (b) square of a triangular matrix and the eigenvalues

   (c) Symmetric matrices have eigenvectors that are orthogonal

   (d) eigenvelues of rotations through 90 and 180 degrees

3. Walk through Proof 4.2: For real $n \times n$ matrix A, prove that all the polynomials $p_i(t)$ are simple and have real roots iff there exists a basis for $\mathbb{R}^n$ consisting of eigenvectors of A.

4. If λ is an eigenvalue A with $\vec{v}$ the corresponding eigenvector, prove that $\lambda - s$ is an eigenvalue of $A - sI$ for any scalar s, with corresponding eigenvector $\vec{v}$.

MATHEMATICS 23a/E-23a, Fall 2018
Linear Algebra and Real Analysis I
Week 6 (Series, Convergence, Power Series)

Authors: Paul Bamberg and Kate Penner (based on their course MATH S-322)
R scripts by Paul Bamberg
Last modified: August 13, 2018 by Paul Bamberg (special offer on HW)

Reading from Ross

- Chapter 2, sections 10 and 11(pp. 56-77) (monotone and Cauchy sequences, subsequences, introduction to lim sup and lim inf)
- Chapter 2, sections 14 and 15 (pp. 95-109) (series and convergence tests)
- Chapter 4, section 23 (pp.187-192) (convergence of power series)

Recorded Lectures

- Lecture 12 (Week 6, Class 1) (watch on October 16 or 17)
- Lecture 13 (Week 6, Class 2) (watch on October 18 or 19)

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

- 6.1 Bolzano-Weierstrass
  - Prove that any bounded increasing sequence converges. (You may assume without additional proof the corresponding result, that any bounded decreasing sequence converges.)
  - Prove that every sequence (s_n) has a monotonic subsequence.
  - Prove the Bolzano-Weierstrass Theorem: every bounded sequence has a convergent subsequence.
- 6.2 The Root Test
  Consider the infinite series $\sum a_n$ and $\limsup |a_n|^{1/n}$, referred to as α. Prove the following statements about $\sum a_n$:
  - The series converges absolutely if $\alpha < 1$.
  - The series diverges if $\alpha > 1$.
  - If $\alpha = 1$, then nothing can be deduced conclusively about the behavior of the series.

Additional proofs(may appear on quiz, students wiill post pdfs or videos

- 6.3 (Cauchy sequences) A Cauchy sequence is defined as a sequence where $\forall \epsilon > 0$, $\exists N$ s.t. $\forall m, n > N \implies |s_n - s_m| < \epsilon$

  - Prove that any Cauchy sequence is bounded.
  - Prove that any convergent sequence is Cauchy.
  - Prove that any Cauchy sequence of real numbers is convergent. You will need to use something that follows from the completeness of the real numbers. This could be the Bolzano-Weierstrass theorem, or it could the fact that, for a sequence of real numbers, if $\liminf s_n = \limsup s_n = s$, then $\lim s_n$ is defined and

    $$\lim s_n = s$$

- 6.4 (Ross, p.188, Radius of Convergence)
  Consider the power series $\sum a_n x^n$. Refer to $\limsup |a_n|^{1/n}$ as β and $1/\beta$ as R. If $\beta = 0, R = +\infty$ and if $\beta = +\infty, R = 0$.)
  Prove the following:

  - If $|x| < R$, the power series converges.
  - If $|x| > R$, the power series diverges.

R Scripts

- Script 2.2A-MoreSequences.R
  Topic 1 – Cauchy Sequences
  Topic 2 – Lim sup and lim inf of a sequence

- Script 2.2B-Series.R
  Topic 1 – Series and partial sums
  Topic 2 – Passing and failing the root test
  Topic 3 – Why the harmonic series diverges

1 Executive Summary

1.1 Monotone sequences

A sequence (s_n) is **increasing** if $s_n \leq s_{n+1}\ \forall n$.
A sequence (s_n) is **strictly increasing** if $s_n < s_{n+1}\ \forall n$.
A sequence (s_n) is **decreasing** if $s_n \geq s_{n+1}\ \forall n$.
A sequence (s_n) is **strictly decreasing** if $s_n > s_{n+1}\ \forall n$.
A sequence that is either increasing or decreasing is called a **monotone** sequence.
All bounded monotone sequences converge.
For an unbounded increasing sequence, $\lim_{n\to\infty} s_n = +\infty$.
For an unbounded decreasing sequence, $\lim_{n\to\infty} s_n = -\infty$.

1.2 Supremum, infimum, maximum, minimum

The **supremum** of a subset S (which is a subset of some set T) is the least element of T that is greater than or equal to all of the elements that are in the subset S. The supremum of the subset S definitely lives in the set T. It may also be in S, but that is not a requirement.
The **supremum** of a sequence the least upper bound of its set of elements.
The **maximum** is the largest value attained within a set or sequence.
It is easy to find examples of sets or sequences for which no supremum exists, or for which a supremum exists but a maximum does not.

The **infimum** of a sequence is the "greatest lower bound," or the greatest element of T that is less than or equal to all of the elements that are in the subset S. It is not the same as a **minimum**, because the minimum must be achieved in S, while the infimum may be an element of only T.

1.3 Cauchy sequences

A sequence is a *Cauchy sequence* if

$$\forall \epsilon > 0, \exists N \text{ s.t. } \forall m, n > N,\ |s_n - s_m| < \epsilon$$

Both convergent and Cauchy sequences must be bounded.
A convergent sequence of real numbers or of rational numbers is Cauchy.
A Cauchy sequence *of real numbers* is convergent.
It is easy to invent a Cauchy sequence of rational numbers whose limit is an irrational number.
Off the record: quantum mechanics is done in a "Hilbert space," one of the requirements for whichis that every Cauchy sequence is convergent. Optimization problems in economics are frequently formualated in a "Banach space," which has the same requirement.

1.4 lim inf and lim sup

Given any bounded sequence, the "tail" of the sequence, which consists of the infinite number of elements beyond the Nth element, has a well-defined supremum and infimum.

Let us combine the notion of limit with the definitions of supremum and infimum. The "limit infimum" and "limit supremum" are written and defined as follows:

$$\liminf s_n = \lim_{N\to\infty} \inf\{s_n : n > N\}$$

$$\limsup s_n = \lim_{N\to\infty} \sup\{s_n : n > N\}$$

The limit supremum is defined in a parallel manner, only considering the supremum of the sequences instead of the infimum.

Now that we know the concepts of $\liminf$ and $\limsup$, we find the following properties hold:

- If $\lim s_n$ is defined as a real number or $\pm\infty$, then

$$\liminf s_n = \lim s_n = \limsup s_n$$

- If $\liminf s_n = \limsup s_n$, then $\lim s_n$ is defined and

$$\lim s_n = \liminf s_n = \limsup s_n$$

- For a Cauchy sequence of real numbers, $\liminf s_n = \limsup s_n$, and so the sequence converges.

1.5 Subsequences and the Bolzano-Weierstrass theorem

A subsequence is a sequence obtained by selecting an infinite number of terms from the "parent" sequence in order.
If (s_n) converges to s, then any subsequence selected from it also converges to s.

Given any sequence, we can construct from it a monotonic subsequence, either an increasing whose limit is $\limsup s_n$, a decreasing sequence whose limit is $\liminf s_n$, or both. If the original sequence is bounded, such a monotonic sequence must converge, even if the original sequence does not.
This construction proves one of the most useful results in all of mathematics, the Bolzano-Weierstrass theorem:
Every bounded sequence has a convergent subsequence.

1.6 Infinite series, partial sums, and convergence

Given an infinite series Σa_n we define the **partial sum**

$$s_n = \sum_{k=m}^{n} a_k$$

The lower limit m is usually either 0 or 1.

The series $\sum_{k=m}^{\infty} a_k$ is said to **converge** when the limit of its partial sums as $n \to \infty$ equals some number S. If a series does not converge, it is said to **diverge**. The sum $\sum a_n$ has no meaning unless its sequence of partial sums either converges to a limit S or diverges to either $+\infty$ or $-\infty$.
A series with all positive terms will either converge or diverge to $+\infty$.
A series with all negative terms will either converge or diverge to $-\infty$.
For a series with both positive and negative terms, the sum $\sum a_n$ may have no meaning.
A series is called **absolutely convergent** if the series $\sum |a_n|$ converges.
Absolutely convergent series are also convergent.

1.7 Familiar examples

A **geometric series** is of the form

$$a + ar + ar^2 + ar^3 + \dots$$

If $|r| < 1$, then

$$\sum_{n=0}^{\infty} ar^n = \frac{a}{1-r}$$

A **p-series** is of the form

$$\sum_{n=1}^{\infty} \frac{1}{n^p}$$

for some positive real number p. It converges if $p > 1$, diverges if $p \leq 1$.

1.8 Cauchy criterion

. We say that a series satisfies the Cauchy criterion if the sequence of its partial sums is a Cauchy sequence. Writing this out with quantifiers, we have

$$\forall \epsilon > 0, \exists N \text{ s.t. } \forall m, n > N,\ |s_n - s_m| < \epsilon$$

Here is a restatement of the Cauchy criterion, which proves more useful for some proofs:

$$\forall \epsilon > 0, \exists N \text{ s.t. } \forall n \geq m > N,\ |\sum_{k=m}^{n} a_k| < \epsilon$$

A series converges if and only if it satisfies the Cauchy criterion.

1.9 Convergence tests

- **Limit of the terms.** If a series converges, the limit of its terms is 0.
- **Comparison Test.** Consider the series $\sum a_n$ of all positive terms.
  If $\sum a_n$ converges and $|b_n| \leq a_n$ for all n then $\sum b_n$ also converges.
  If $\sum a_n$ diverges to $+\infty$ and $|b_n| > a_n$ for all n, then $\sum b_n$ also diverges to $+\infty$
- **Ratio Test.** Consider the series $\sum a_n$ of nonzero terms.
  This series converges if $\limsup |\frac{a_{n+1}}{a_n}| < 1$
  This series diverges if $\liminf |\frac{a_{n+1}}{a_n}| > 1$
  If $\liminf |\frac{a_{n+1}}{a_n}| \leq 1 \leq \limsup |\frac{a_{n+1}}{a_n}|$, then we have no information and need to perform another test to determine convergence.
- **Root Test.** Consider the series $\sum a_n$, and evaluate $\limsup |a_n|^{1/n}$.
  If $\limsup |a_n|^{1/n} < 1$, the series $\sum a_n$ converges absolutely.
  If $\limsup |a_n|^{1/n} > 1$, the series $\sum a_n$ diverges.
  If $\limsup |a_n|^{1/n} = 1$, the test gives no information.
- **Integral Test.** Consider a series of nonnegative terms for which the other tests seem to be failing. In the event that we can find a function $f(x)$, such that $f(n) = a_n \ \forall n$, we may look at the behavior of this function's integral to tell us whether the series converges.
  If $\lim_{n\to\infty} \int_1^n f(x)dx = +\infty$, then the series will diverge.
  If $\lim_{n\to\infty} \int_1^n f(x)dx < +\infty$, then the series will converge.
- **Alternating Series Test.** If the absolute value of the each term in an alternating series is decreasing and has a limit of zero, then the series converges.

1.10 Convergence tests for power series

Power series are series of the form

$$\sum_{n=0}^{\infty} a_n x^n$$

where the sequence (a_n) is a sequence of real numbers. A power series defines a function of x whose domain is the set of values of x for which the series converges. That, of course, depends on the coefficients (a_n). There are three possibilities:
– Converges $\forall x \in \mathbb{R}$.
– Converges only for $x = 0$.
– Converges $\forall x$ in some interval, centered at 0. The interval may be open $(-R, R)$, closed $[-R, R]$, or a mix of the two like $[-R, R]$. The number R is called the *radius of convergence*. Frequently the series converges absolutely in the interior of the interval, but the convergence at an endpoint is only conditional.

2 Lecture Outline

1. Supremum, infimum, maximum, minimum
   The **supremum** of a subset S (which is a subset of some set T) is the least element of T that is greater than or equal to all of the elements that are in the subset S. The supremum of the subset S definitely lives in the set T. It may also be in S, but that is not a requirement.
   The **supremum** of a sequence the least upper bound of its set of elements.
   The **maximum** is the largest value attained within a set or sequence.

   Invent a sequence for which no supremum exists.

   Invent a sequence for which a supremum exists but a maximum does not.

   The **infimum** of a subset $T \subset S$ is the "greatest lower bound," or the greatest element of T that is less than or equal to all of the elements that are in the subset S. It is not the same as a **minimum**, because the minimum must be achieved in S, while the infimum may be an element only of T.

2. Monotone sequences
   Terminology used in the latest edition of Ross (available through HOLLIS):

   (a) A sequence (s_n) is **increasing** if $s_n \leq s_{n+1}\ \forall n$.

   (b) A sequence (s_n) is **strictly increasing** if $s_n < s_{n+1}\ \forall n$.

   (c) A sequence (s_n) is **decreasing** if $s_n \geq s_{n+1}\ \forall n$.

   (d) A sequence (s_n) is **strictly decreasing** if $s_n > s_{n+1}\ \forall n$.

   (e) A sequence that is either increasing or decreasing is called a **monotone** sequence.

   If you own an earlier edition of Ross, beware:
   He used "nondecreasing" for what is now called "increasing" .
   He used "increasing" for what is now called "strictly increasing."

   Useful and easy-to-prove results for increasing sequences of real numbers.
   A bounded increasing sequence converges to its least upper bound.

   For an unbounded increasing sequence, $\lim s_n = +\infty$.

3. Defining a sequence recursively (model for group problems, set 1)

   John's rich parents hope that a track record of annual gifts to Harvard will enhance his chance of admission. On the day of his birth they set up a trust fund with a balance $s_0 = 1$ million dollars. On each birthday they add another million dollars to the fund, and the trustee immediately donates 1/3 of the fund to Harvard in John's name. After the donation, the balance is therefore

$$s_{n+1} = \frac{2}{3}(s_n + 1).$$

   - Find the annual fund balance up through s_2.
   - Use induction to show $s_n < 2$ for all n.
   - Show that (s_n) is an increasing sequence.
   - Show that $\lim s_n$ exists and find $\lim s_n$.

4. (Ross, p. 62, convergent & Cauchy sequences)
   A Cauchy sequence is defined as a sequence where
   $\forall \epsilon > 0$, $\exists N$ s.t. $\forall m, n > N \implies |s_n - s_m| < \epsilon$.

   (a) Prove that any Cauchy sequence is bounded.

   (b) Prove that any convergent sequence is Cauchy.

5. (Ross, pp. 60-62, limits of the supremum and infimum)
The limit of the supremum, written "lim sup" is defined as follows:

$$\limsup s_n = \lim_{N\to\infty} \sup\{s_n : n > N\}$$

The limit of the infimum, written "lim inf" is defined as follows:

$$\liminf s_n = \lim_{N\to\infty} \inf\{s_n : n > N\}$$

(We do not restrict s_n to be a bounded sequence, so if it is not bounded above, $\limsup s_n = +\infty$, and if it is not bounded below, $\liminf s_n = -\infty$) With these definitions, $\limsup s_n$ and $\liminf s_n$ exist for every sequence (s_n). These are difficult concepts, but we will need them in order to make correct statements of convergence tests for infinite series.

An easy example: If $s_n = (0, \frac{1}{2}, -2, \frac{3}{4}, -4, \frac{7}{8}, -8, \cdots)$

What is $\limsup s_n$?

What is $\liminf s_n$?

A slightly harder example: If $s_n = (\frac{3}{2}, -\frac{3}{2}, \frac{4}{3}, -\frac{4}{3}, \frac{5}{4}, -\frac{5}{4}, \cdots)$

What is $\limsup s_n$?

What is $\liminf s_n$?

6. Existence of a limit in terms of lim sup and lim inf

   Let (s_n) be a sequence in $\mathbb{R}$. Prove that if $\liminf s_n = \limsup s_n = s$, then $\lim s_n$ is defined and

$$\lim s_n = s.$$

7. Examples of subsequences

   Consider the sequence

$$s_n = \frac{n+2}{n+1} \sin(\frac{n\pi}{4}),$$

   give three examples of a subsequence, find the lim sup and the lim inf, and determine whether it converges.

   To get started, write out the first few terms.

8. Determine the lim sup and the lim inf for this sequence.

9. Invent a subsequence for which the lim sup is positive but less than 1 and the lim inf is 0.

10. Invent a subsequence that converges to a negative number.

11. (Ross, p. 64, convergent and Cauchy sequences)

    Using the result of the preceding proof, which relies on the completeness axiom for the real numbers, prove that any Cauchy sequence of real numbers is convergent.

12. (Convergent subsequences, Bolzano Weierstrass)
Given a sequence $(s_n)_{n\in\mathbb{N}}$, a subsequence of this sequence is a sequence $(t_k)_{k\in\mathbb{N}}$, where for each k, there is a positive integer n_k such that

$$n_1 < n_2 < \ldots < n_k < n_{k+1} \ldots$$

and $t_k = s_{n_k}$. So (t_k) is just a sampling of some, or all, of the (s_n) terms, with order preserved.
A term s_n is called *dominant* if it is greater than any term that follows it.

(a) Use the concept of dominant term to prove that every sequence (s_n) has a monotonic subsequence.

(b) Prove the Bolzano-Weierstrass Theorem: every bounded sequence has a convergent subsequence.

13. Infinite series, partial sums, and convergence
Given an infinite series Σa_n we define the **partial sum**

$$s_n = \sum_{k=m}^{n} a_k$$

The lower limit m is usually either 0 or 1.

The series $\sum_{k=m}^{\infty} a_k$ is said to **converge** when the limit of its partial sums as $n \to \infty$ equals some number S. If a series does not converge, it is said to **diverge**. The sum $\sum a_n$ has no meaning unless its sequence of partial sums either converges to a limit S or diverges to either $+\infty$ or $-\infty$.
A series with all positive terms will either converge or it will diverge to $+\infty$.
A series with all negative terms will either converge or it will diverge to $-\infty$.
For a series with both positive and negative terms, the sum $\sum a_n$ may have no meaning.
A series is called **absolutely convergent** if the series $\sum |a_n|$ converges.
Absolutely convergent series are always convergent.

A series is called **conditionally convergent** if the series $\sum a_n$ converges but the series $\sum |a_n|$ diverges.

14. A cautionary tale about conditional convergence.

    What is the fallacy in the following argument?

    - $$\log_e 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \cdots .$$
    - $$\frac{1}{2}\log_e 2 = \frac{1}{2} - \frac{1}{4} + \frac{1}{6} - \frac{1}{8} + \cdots .$$
    - $$\frac{3}{2}\log_e 2 = 1 - \frac{1}{4} - \frac{1}{4} + \frac{1}{3} - \frac{1}{8} - \frac{1}{8} + \frac{1}{5} - \frac{1}{12} - \frac{1}{12} + \cdots = \log_e 2.$$
    - $$\frac{3}{2} = 1; 3 = 2; 1 = 0.$$

    In fact, the general situation is even worse that this example suggests. In a conditionally convergent series

    - The sum of the positive terms is $+\infty$.
    - The sum of the negative terms is $-\infty$.

    The Riemann series theorem states that it is possible to reorder the terms of a conditionally convergent series to achieve any of the following:

    - The sequence of partial sums converges to any desired value a.
    - The sequence of partial sums diverges to $+\infty$.
    - The sequence of partial sums diverges to $-\infty$.

    Suppose that the terms a_n are a conditionally convergent infinite series of revenue items (positive) and expense items (negative) for your startup company. How could you exploit the Riemann series theorem to create alternative business plans with any desired long-term outcome?
    (Beware: CFOs who do this sometimes end up in prison!)

15. Non-negative terms: order does not matter

If all the a_i are non-negative and the series converges, the order of terms in the series is irrelevant. We can express the sum in a way that is independent of the order of the terms as

$$S' = \sup \sum_{i \in A \subset \mathbb{N}} a_i$$

where the supremum is over all finite subsets of $\mathbb{N}$.

We need to prove that S' is equal to S, the limit of the sequence $S_0, S_1, ...$

The sum S is the limit of a nondecreasing sequence, so it can be expressed as

$$S = \sup S_N$$

How does this establish that $S' \geq S$?

Any finite subset A is a subset of S_N for some N. How does this establish that $S' \leq S$?

But if $S' \geq S$ and $S' \leq S$, then $S' = S'$, and the proof is complete.

On the homework you can prove that if a series has both positive and negative terms but $\sum_i |a_i|$ converges, then the order in which the terms are summed is irrelevant. Such a series is absolutely convergent, not conditionally convergent.

16. (Ross, p. 96, Example 1, geometric series (refers also to p. 98))

Prove that

$$\sum_{k=0}^{\infty} ar^k = \frac{a}{1-r} \text{ if } |r| < 1,$$

and that the series diverges if $|r| \geq 1$.

For the sake of novelty, do the first part of the proof by using the least-number principle insted of by induction.

Given a repeating decimal, you can write this number as a geometric series. Write the repeating decimal $0.363636363\cdots$ as a geometric series, and use the formula

$$\frac{a}{1-r}$$

to show that it is equal to $4/11$.

17. Cauchy criterion

We say that a series satisfies the Cauchy criterion if the sequence of its partial sums is a Cauchy sequence. Writing this out with quantifiers, we have

$$\forall \epsilon > 0, \exists N \text{ s.t. } \forall m, n > N,\ |s_n - s_m| < \epsilon$$

Here is a restatement of the Cauchy criterion, which proves more useful for some proofs:

$$\forall \epsilon > 0, \exists N \text{ s.t. } \forall n \geq m > N,\ |\sum_{k=m}^{n} a_k| < \epsilon$$

A series converges if and only if it satisfies the Cauchy criterion.

Use the Cauchy criterion to prove two simple but useful convergence tests.

- **Limit of the terms.** For a series to converge, the limit of its terms must be 0.

- **Comparison Test.** Consider the series $\sum a_n$ of all positive terms.
  If $\sum a_n$ converges and $|b_n| \leq a_n\ \forall n$ then $\sum b_n$ also converges.
  If $\sum a_n$ diverges to $+\infty$ and $b_n > a_n\ \forall n$, then $\sum b_n$ diverges to $+\infty$.

18. Clever proofs for p-series.

(a) Prove that $\sum \frac{1}{n} = +\infty$ by showing that the sequence of partial sums is not a Cauchy sequence.

(b) Evaluate

$$\sum_{n=2}^{\infty} \frac{1}{n(n-1)}$$

by exploiting the fact that this is a "telescoping series."

(c) Prove that

$$\sum_{n=2}^{\infty} \frac{1}{n^2}$$

is convergent.

19. (Ross, p.99, The Root Test)
Consider the infinite series $\sum a_n$ and $\limsup |a_n|^{1/n}$, referred to as α.
Prove the following statements about $\sum a_n$:
(you may assume the Comparison Test as proven)

- The series converges absolutely if $\alpha < 1$.
- The series diverges if $\alpha > 1$.
- If $\alpha = 1$, then nothing can be deduced conclusively about the behavior of the series.

20. (Ross, pp. 99-100, The Ratio Test)
Let $\sum a_n$ be an infinite series of nonzero terms.
Prove the following, assuming the Root Test as proven.
We will use without proof the following result from Ross (theorem 12.2):

$$\liminf |\frac{s_{n+1}}{s_n}| \leq \liminf |s_n|^{\frac{1}{n}} \leq \limsup |s_n|^{\frac{1}{n}} \leq \limsup |\frac{s_{n+1}}{s_n}|$$

- If $\limsup |a_{n+1}/a_n| < 1$, then the series converges absolutely.
- If $\liminf |a_{n+1}/a_n| > 1$, then the series diverges.
- If $\liminf |a_{n+1}/a_n| \leq 1 \leq \limsup |a_{n+1}/a_n|$, then the test gives no information.

21. A case where the root test outperforms the ratio test
(Ross, Example 8 on page 103)

$$\sum_{n=0}^{\infty} 2^{(-1)^n - n} = 2 + \frac{1}{4} + \frac{1}{2} + \frac{1}{16} + \frac{1}{8} + \frac{1}{64} + \cdots .$$

(a) Show that the ratio test fails totally.

(b) Show that the root test correctly concludes that the series is convergent.

(c) Find a simpler argument using the comparison test.

22. (Ross, p.188, Radius of Convergence)
Consider the power series $\sum a_n x^n$. Let us refer to $\limsup |a_n|^{1/n}$ as β and $1/\beta$ as R.
Limiting cases: if $\beta = 0, R = +\infty$ and if $\beta = +\infty, R = 0$.
Prove the following:

- If $|x| < R$, the power series converges.
- If $|x| > R$, the power series diverges.

(You may recognize R here as the radius of convergence.)

23. Tests that are useful when the root test and the ratio test fail, which is often the case at the endpoints of the interval of convergence of a power series.

Integral Test. Consider a series of nonnegative terms for which the other tests seem to be failing. In the event that we can find a function $f(x)$, such that $f(n) = a_n \ \forall n$, we may look at the behavior of this function's integral to tell us whether the series converges.
If $\lim_{n\to\infty} \int_1^n f(x)dx = +\infty$, then the series will diverge.
If $\lim_{n\to\infty} \int_1^n f(x)dx < +\infty$, then the series will converge.
Use this test to provide an alternate proof that $\sum \frac{1}{n}$ diverges but $\sum \frac{1}{n^2}$ converges.

Alternating Series Test. If the absolute value of the each term in an alternating series a_n is decreasing and has a limit of zero, then the series converges.
Ross proves this on page 108 by using the Cauchy criterion. Prove it instead by considering the sequence of partial sums, (s_n), and showing that $\liminf s_n = \limsup s_n$.

24. (Model for group problems, set 3 – to be done as time permits) Find the radius of convergence and the exact interval of convergence for the series

$$\sum_{n=0}^{\infty} \frac{n}{2^n} x^{3n} \text{ by using the Root Test.}$$

Then use appropriate tests for the following examples.

$$\sum (\frac{2^n}{n!}) x^n$$

$$\sum n! x^n.$$

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 6.1 – Bolzano-Weierstrass)

   - Prove that any bounded increasing sequence converges. (You may assume without additional proof the corresponding result, that any bounded decreasing sequence converges.)
   - Prove that every sequence (s_n) has a monotonic subsequence.
   - Prove the Bolzano-Weierstrass Theorem: every bounded sequence has a convergent subsequence.

2. For sequence (s_n), define $\limsup s_n$, and illustrate your definition for these two sequences, neither of which converges:

   - $(s_n) = (2, 0, \frac{3}{2}, 0, \frac{4}{3}, 0, \frac{5}{4}, 0, \cdots)$
   - $(t_n) = (0, 0, \frac{1}{2}, 0, \frac{2}{3}, 0, \frac{3}{4}, 0, \cdots)$

   Then prove the following properties of lim and lim sup:

   - If $\lim s_n = s$, then $\forall \epsilon > 0$,
     there are **only finitely many** s_n for which $s_n \geq s + \epsilon$
     and there are **only finitely many** s_n for which $s_n \leq s - \epsilon$.
   - If $\limsup s_n = s$, then $\forall \epsilon > 0$,
     there are **only finitely many** s_n for which $s_n \geq s + \epsilon$
     and there are **infinitely many** s_n for which $s_n > s - \epsilon$.

3. (Proof 6.2 – The Root Test)
   Consider the infinite series $\sum a_n$ and $\limsup |a_n|^{1/n}$, referred to as α.
   Prove the following statements about $\sum a_n$:

   - The series converges absolutely if $\alpha < 1$.
   - The series diverges if $\alpha > 1$.
   - If $\alpha = 1$, then nothing can be deduced conclusively about the behavior of the series.

4. (Proof 6.3 – Cauchy sequences)
   A Cauchy sequence is defined as a sequence where $\forall \epsilon > 0$, $\exists N$ s.t. $\forall m, n > N \implies |s_n - s_m| < \epsilon$

   - Prove that any Cauchy sequence is bounded.
   - Prove that any convergent sequence is Cauchy.
   - Prove that any Cauchy sequence of real numbers is convergent. You will need to use something that follows from the completeness of the real numbers. This could be the Bolzano-Weierstrass theorem, or it could the fact that, for a sequence of real numbers, if $\liminf s_n = \limsup s_n = s$, then $\lim s_n$ is defined and $\lim s_n = s$.

     A simpler strategy than the one that Ross uses is to prove the contrapositive: assume that (s_n) is not convergent, so that
     $\limsup s_n - \liminf s_n = 3\epsilon > 0$,
     and prove that (s_n) is not Cauchy.

5. (Proof 6.4 – Radius of Convergence)
   Consider the power series $\sum a_n x^n$. Refer to $\limsup |a_n|^{1/n}$ as β and $1/\beta$ as R. If $\beta = 0, R = +\infty$ and if $\beta = +\infty, R = 0$.)
   Prove the following:

   - If $|x| < R$, the power series converges.
   - If $|x| > R$, the power series diverges.

6. (Extra topic – see pages 21 and 22 in the lecture outline)
   State the Cauchy criterion for convergence of series, and use it to prove the following results, which are traditionally proved by the integral test:

   - The series

     $$\sum_{k=1}^{\infty} \frac{1}{k} \text{ diverges. (Hint: consider } \sum_{k=2^n+1}^{2^{n+1}} \frac{1}{k}.)$$

   - The series

     $$\sum_{k=1}^{\infty} \frac{1}{k^2} \text{ converges.}$$

     $$\text{Hint: } \sum_{k=m}^{n} \frac{1}{k^2} < \sum_{k=m}^{n} \frac{1}{k(k-1)} = \sum_{k=m}^{n} (\frac{1}{k-1} - \frac{1}{k}).$$

4 Workshop Problems

1. Working with lim sup and lim inf

   (a) (Ross, 12.4)
   Show that $\limsup(s_n + t_n) \leq \limsup s_n + \limsup t_n$ for bounded sequences (s_n) and (t_n), and invent an example where $\limsup(s_n + t_n) < \limsup s_n + \limsup t_n$.
   Here is the hint from page 82 of Ross : first show that

   $$\sup\{s_n + t_n : n > N\} \leq \sup\{s_n : n > N\} + \sup\{t_n : n > N\}$$

   You may use Ross Exercise 9.9 c, which can be stated as
   If $(a_n) \to a, (b_n) \to b$, and $\exists N_0$ s.t. $\forall n > N_0, a_n \leq b_n$, then $a \leq b$.
   A useful trick might be to replace t_n on the left-hand side by $v_N = \sup\{t_n : n > N\}$, which does not depend on n.

   (b) The following famous series, known as Gregory's series but discovered by the priest-mathematicians of southwest India long before James Gregory (1638-1675) was born, converges to $\frac{\pi}{4}$.

   $$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} + \cdots .$$

   i. For the sequence of partial sums (s_n), find an increasing subsequence and a decreasing subsequence.

   ii. Prove convergence by showing that $\limsup s_n = \liminf s_n$.
   Hint: Show that $\sup\{s_n : n > N\} - \inf\{s_n : n > N\}$ is equal to the magnitude of one term in the series.

   iii. Prove that the series is not absolutely convergent by showing that it fails the Cauchy test with $\epsilon = 1/4$.

2. Sequences, defined recursively

If someone in your group is skillful with R, you can use it to calculate a lot of terms of the sequence. By modifying script 2.2C, you can easily plot the first 20 or so terms. It you come up with a good R script, please upload it to the solutions page.

Once you have shown that $\lim s_n$ exists, you can assert that $\lim s_{n+1} = \lim s_n = s$.

(a) (Ross, 10.9) Let $s_1 = 1$ and $s_{n+1} = (\frac{n}{n+1})s_n^2$ for $n > 1$.

- Find s_2, s_3, s_4 if working by hand. If using R, use a for loop to go at least as far as s_{20}.
- Show that $\lim s_n$ exists.
- Prove that $\lim s_n = 0$.

(b) (Ross, 10.12) Let $t_1 = 1$ and $t_{n+1} = [1 - \frac{1}{(n+1)^2}]t_n$ for $n > 1$.

- Find t_2, t_3, t_4 if working by hand. If using R, use a for loop to go at least as far as t_{20}.
- Show that $\lim t_n$ exists.
- Use induction to show $t_n = \frac{n+1}{2n}$ for all n.
- Find $\lim t_n$.

This last set of problems should be done using LaTeX or in the Canvas editor. They provide good practice with summations, fractions, and exponents.

3. Applying convergence tests to power series (Ross, 23.1 and 23.2)

   Find the radius of convergence R and the exact interval of convergence.

   In each case, you can apply the root test (works well with powers) or the ratio test (works well with factorials) to get an equation that can be solved for x to get the radius of convergence R. Since you have an x^n, the root test, which you may not have encountered in AP calculus, is especially useful. At the endpoints you may need to apply something like the alternating series test or the integral test.

   Remember that $\lim n^{1/n} = 1$.

   (a)
   $$\sum (\frac{3^n}{n \cdot 4^n}) x^n \text{ and } \sum \sqrt{n} x^n.$$

   (b)
   $$\sum (\frac{(-1)^n}{n^2 4^n}) x^n \text{ and } \sum \frac{3^n}{\sqrt{n}} x^n.$$

5 Homework

Special offer – if you do the entire problem set, with one problem omitted, in LaTeX, you will receive full credit for the omitted problem. Alternatively, if you work all the problems in LaTeX, we will convert your lowest score to a perfect score.

1. Ross, 10.2 (Prove all bounded decreasing sequences converge.)

2. Ross, 10.6,

3. Ross, 11.8.

4. Suppose that (s_n) is a Cauchy sequence and that the subsequence $(s_1, s_2, s_4, s_8, s_{16}, \cdots)$ converges to s. Prove that $\lim s_n = s$. Hint – use the standard bag of tricks: the triangle inequality, epsilon-over-2, etc.

5. If a series $\sum_i a_i$ has both positive and negative terms but $\sum_i |a_i|$ converges, then the series is said to be absolutely convergent. If S_+ denotes the least upper bound for subsets of positive terms and S_- denotes the greatest lower bound for subsets of negative terms, then the sum $S = \sum_i a_i$ can be written as $S = S_+ + S_-$. This formula makes it clear that the order in which the terms are summed is irrelevant. As defined, S_- is a negative number.

   (a) Start with equation 0.5.7, which is proved at the bottom of page 20 of Hubbard. Re-express this formula by writing the sums on the right in terms of S_+ and S_-, and thereby show that $S = S_+ + S_-$.

   (b) Hubbard's proof of Theorem 0.5.8 treats positive and negative terms in an unsymmetrical way. Create a new, more symmetrical version of the proof by using Hubbard's b_m and also defining

   $$c_m = \sum_{n=1}^{m} (|a_n| - a_n).$$

   In your proof, be careful to change the order only in finite sums, and make it clear how you are using Theorem 0.5.7. You can use the result of theorem 1.5.16, which you proved in the previous problem, in the case $n = 1$.

6. Ross, 14.3 (Determining whether a series converges. Apologies to those who have already done hundreds of these in a high-school course.)

7. Ross, 14.8.

8. Ross, 15.6.

9. Ross, 23.4. You might find it useful to have R generate some terms of the series.

10. Ross, 23.5.

1. Is it true that if x_n is convergent and y_n is divergent , then $x_n y_n$ is divergent? If yes, provide a proof, if not provide a counter example. (adapted from Feher)
2. Calculate the limit of $(2^n - n)^{1/n}$ (Adapted from Feher).
3. Prove the squeeze lemma sometimes referred to as the sandwich theorem.
4. Prove that $\limsup x_n + y_{n_n} \leq \limsup x_n + \limsup y_n$(Workshop problems)
5. A sequence a_n has only one convergent subsequence that converges to a. Is it necessarily true that $\lim_{n\to\infty} a_n = a$? If so, provide a proof, if not provide a counterexample.(Adapted from 2)
6. Prove that $\frac{(n^2+100n+86)\sin n^3}{n^2+n+1}$ has a convergent subsequence. (Adapted from Source 2)

MATHEMATICS 23a/E-23a, Fall 2018
Linear Algebra and Real Analysis I
Week 7 (Limits and continuity of functions)

Authors: Paul Bamberg and Kate Penner (based on their course MATH S-322)
R scripts by Paul Bamberg
Last modified: August 13, 2018 by Paul Bamberg

Reading from Ross

- Chapter 3, sections 17 and 18. (continuity)
- Chapter 3, sections 19 and 20 (uniform continuity and limits of functions)

Recorded Lectures

- Lecture 14 (Week 7, Class 1) (watch on October 23 or 24)
- Lecture 15 (Week 7, Class 2) (watch on October 25 or 26)

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

- 7.1 Suppose that $a < b$, f is continuous on $[a, b]$, and $f(a) < y < f(b)$. Prove that there exists at least one $x \in [a, b]$ such that $f(x) = y$.

  Use Ross's "no bad sequence" definition of continuity, not the epsilon-delta definition.

- 7.2 Using the Bolzano-Weierstrass theorem, prove that if function f is continuous on the closed interval $[a, b]$, then f is uniformly continuous on $[a, b]$.

Additional proofs(may appear on quiz, students will post pdfs or videos)

- 7.3 Prove that if f and g are real-valued functions that are continuous at $x_0 \in \mathbb{R}$, then $f + g$ is continuous at x_0. Do the proof twice: once using the "no bad sequence" definition of continuity and one using the epsilon-delta definition of continuity.

- 7.4 (Ross, page 146; uniform continuity and Cauchy sequences)

  Prove that if f is uniformly continuous on a set S and (s_n) is a Cauchy sequence in S, then $(f(s_n))$ is a Cauchy sequence. Invent an example where f is continuous but not uniformly continuous on S and $(f(s_n))$ is not a Cauchy sequence.

R Scripts

- Script 2.3A-Continuity.R
  Topic 1 - Two definitions of continuity
  Topic 2 – Uniform continuity

- Script 2.3B-IntermediateValue.R
  Topic 1 - Proving the intermediate value theorem
  Topic 2 - Corollaries of the IVT

1 Executive Summary

1.1 Two equivalent definitions of continuity

- Continuity in terms of sequences

  This definition is not standard: Ross uses it, but many authors use the equivalent epsilon-delta definition. Here is some terminology that students find useful when discussing the concept:

  - If $\lim x_n = x_0$ and $\lim f(x_n) = f(x_0)$, we call x_n a "good sequence."
  - If $\lim x_n = x_0$ but $\lim f(x_n) \neq f(x_0)$, we call x_n a "bad sequence."

  Then "function f is continuous at x_0" means "every sequence is a good sequence"; i.e. "there are no bad sequences."

- The more conventional definition:

  Let f be a real-valued function with domain $U \subset \mathbb{R}$. Then f is continuous at $x_0 \in U$ if and only if

  $\forall \epsilon > 0, \exists \delta > 0$ such that if $x \in U$ and $|x - x_0| < \delta, |f(x) - f(x_0)| < \epsilon$.

- Which definition to use?

  To prove that a function is continuous, it is often easier to use the second version of the definition. Start with a specified ϵ, and find a δ (not "the δ") that does the job. However, as Ross Example 1a on page 125 shows, the first definition, combined with the limit theorems that we have already proved, can let us prove that an arbitrary sequence is good.

  To prove that a function is discontinuous, the first definition is generally more useful. All you have to do is to construct one bad sequence.

1.2 Useful properties of continuous functions

- New continuous functions from old ones.
  - If f is continuous at x_0, then $|f|$ is continuous at x_0.
  - If f is continuous at x_0, then kf is continuous at x_0.
  - If f and g are continuous at x_0, then $f + g$ is continuous at x_0.
  - If f and g are continuous at x_0, then fg is continuous at x_0.
  - If f and g are continuous at x_0 and $g(x_0) \neq 0$, then $\frac{f}{g}$ is continuous at x_0.
  - If f is continuous at x_0 and g is continuous at $f(x_0)$, then the composite function $g \circ f$ is continuous at x_0.

  Once you know that the identity function and elementary functions like nth root, sine, cosine, exponential, and logarithm as continuous (Ross has not yet defined most of these functions!), you can state the casual rule

  "If you can write a formula for a function that does not involve division by zero, that function is continuous everywhere."

- Theorems about a continuous function on a closed interval $[a, b]$ (an example of a "compact set"), easy to prove by using the Bolzano-Weierstrass theorem.
  - f is a *bounded* function.
  - f achieves its maximum and minimum values on the interval (i.e. they are not just approached as limiting values).

- The Intermediate Value Theorem and some of its corollaries.
  It is impossible to do calculus without either proving these theorems or stating that they are obvious!

  Now f is assumed continuous on an interval I that is not necessarily closed (e.g. $\frac{1}{x}$ on (0, 1])
  - IVT: If $a < b$ and y lies between $f(a)$ and $f(b)$, there exists at least one x in (a, b) for which $f(x) = y$.
  - The image of an interval I is either a single point or an interval J.
  - If f is a strictly increasing function on I, there is a continuous strictly increasing inverse function $f^{-1} : J \to I$.
  - If f is a strictly decreasing function on I, there is a continuous strictly decreasing inverse function $f^{-1} : J \to I$.
  - If f is one-to-one on I, it is either strictly increasing or strictly decreasing.

1.3 Continuity versus uniform continuity

It's all a matter of the order of quantifiers. For continuity, y is agreed upon before the epsilon-delta game is played. For uniform continuity, a challenge is made using some $\epsilon > 0$, then a δ has to be chosen that meets the challenge independent of y.

For function f whose domain is a set S:

- Continuity: $\forall y \in S, \forall \epsilon > 0$,
  $\exists \delta > 0$ such that $\forall x \in S$, $|x - y| < \delta$ implies $|f(x) - f(y)| < \epsilon$.

- Uniform continuity: $\forall \epsilon > 0$,
  $\exists \delta > 0$ such that $\forall x, y \in S$,$|x - y| < \delta$ implies $|f(x) - f(y)| < \epsilon$.

- On $[0, \infty]$ (not a bounded set), the squaring function is continuous but not uniformly continuous.

- On $(0, 1)$ (not closed) the function $f(x) = \frac{1}{x}$ is continuous but not uniformly continuous.

- On a closed, bounded interval $[a, b]$, continuity implies uniform continuity. The proof uses the Bolzano-Weierstrass theorem.

- By definition, if a function is continuous at $s \in S$ and (s_n) converges to s, then $(f(s_n))$ converges to $f(s)$. If (s_n) is merely Cauchy, we know that it converges, but not what it converges to. To guarantee that $(f(s_n))$ is also Cauchy, we must require f to be *uniformly* continuous.

- On an open interval (a, b) a function can be continuous without being uniformly continuous. However, if we can *extend* f to a function $\overline{f}$, defined so that $\overline{f}$ is continuous at a and b, then $\overline{f}$ is uniformly continuous on $[a, b]$ and f is uniformly continuous on (a, b). The most familiar example is $f(x) = \frac{\sin x}{x}$ on $(0, \infty)$, extended by defining $\overline{f}(0) = 1$.

- Alternative criterion for uniform continuity (sufficient but not necessary): f is differentiable on (a, b), with f' bounded on (a, b).

1.4 Limits of functions

1. Definitions of "limit"

   - Ross's definition of limit, consistent with the definition of continuity: S is a subset of $\mathbb{R}$, f is a function defined on S, and a and L are real numbers, ∞ or $-\infty$. Then $\lim_{x\to a^S} f(x) = L$ means
     for every sequence (x_n) in S with limit a, we have $\lim(f(x_n)) = L$.
   - The conventional epsilon-delta definition:
     f is a function defined on $S \subset \mathbb{R}$, a is a real number in the closure of S (not $\pm\infty$) and L is a real number (not $\pm\infty$). $\lim_{x\to a} f(x) = L$ means $\forall\epsilon > 0, \exists\delta > 0$ such that if $x \in S$ and $|x - a| < \delta$, then $|f(x) - L| < \epsilon$.

2. Useful theorems about limits, useful for proving differentiation rules.

   Note: a can be $\pm\infty$ but L has to be finite.

   Suppose that $L_1 = \lim_{x\to a^S} f_1(x)$ and $L_2 = \lim_{x\to a^S} f_2(x)$ exist and are finite.

   Then

   - $\lim_{x\to a^S}(f_1 + f_2)(x) = L_1 + L_2$.
   - $\lim_{x\to a^S}(f_1 f_2)(x) = L_1 L_2$.
   - $\lim_{x\to a^S}(\frac{f_1}{f_2})(x) = \frac{L_1}{L_2}$, provided $L_2 \neq 0$ and $f_2(x) \neq 0$ for $x \in S$.

3. Limit of the composition of functions

   Suppose that $L = \lim_{x\to a^S} f(x)$ exists and is finite.
   Then $\lim_{x\to a^S}(g \circ f)(x) = g(L)$ provided

   - g is defined on the set $\{f(x) : x \in S\}$.
   - g is defined at L
     (which may just be a limit point of the set $\{f(x) : x \in S\}$.)
   - g is continuous at L.

4. One-sided limits

   We can modify either definition to provide a definition for $L = \lim_{x\to a^+} f(x)$.

   - With Ross's definition, choose the set S to include only values that are greater than a.
   - With the conventional definition, consider only $x > a$: i.e.
     $a < x < a + \delta$ implies $|f(x) - L| < \epsilon$.

   It is easy to prove that

   $\lim_{x\to a} f(x) = L$ if and only if $\lim_{x\to a^+} f(x) = \lim_{x\to a^-} f(x) = L$.

2 Lecture outline

1. Continuity defined in terms of sequences (Ross, page 124)

   For specified x_0 and function f, define the following terminology:

   - If $\lim x_n = x_0$ and $\lim f(x_n) = f(x_0)$, we call x_n a "good sequence."
   - If $\lim x_n = x_0$ but $\lim f(x_n) = f(x_0)$, we call x_n a "bad sequence."

   Then Ross's definition of continuity is "every sequence is a good sequence."

   Prove the following, which is the more conventional definition:

   Let f be a real-valued function with domain $U \subset \mathbb{R}$. Then f is continuous at $x_0 \in U$ if and only if

   $\forall \epsilon > 0, \exists \delta > 0$ such that if $x \in U$ and $|x - x_0| < \delta, |f(x) - f(x_0)| < \epsilon$.

2. Using the "bad sequence" criterion to show that a function is discontinuous.

   The "signum function" $\text{sgn}(x)$ is defined as $\frac{x}{|x|}$ for $x \neq 0$, 0 for $x = 0$.

   Invent a "bad sequence," none of whose elements is zero, to prove that $\text{sgn}(x)$ is discontinuous at 0, then show that for any positive x, no such bad sequence can be constructed.

   Restate this proof that $\text{sgn}(x)$ is discontinous at $x = 0$, continuous for positive x, in terms of the epsilon-delta definition.

3. Proving that a function is continuous

   Using sequences, prove that the function $f(x) = x^2 - 2x + 1$ is continuous at any $x_0 \in \mathbb{R}$. It is important to realize that this argument is valid "for all sequences."

4. Proof 7.3 – two ways to prove a theorem about continuity (Ross, page 128)

   Prove that if f and g are real-valued functions that are continuous at $x_0 \in \mathbb{R}$, then $f + g$ is continuous at x_0.

   Approach 1 (the easy way). Use Ross's definition of continuity in terms of sequences, and invoke the appropriate therem about sequences.

   Approach 2(the hard way). Use the epsilon-delta definition of continuity.

5. An important theorem of calculus that is often just stated without proof (Ross, page 133)

   Let f be a continuous real-valued function on a closed interval $[a, b]$.
   Using the Bolzano-Weierstrass theorem, prove that f is bounded and that f achieves its maximum value: i.e. $\exists y_0 \in [a, b]$ such that $f(x) \leq f(y_0)$ for all $x \in [a, b]$.

6. Proof 7.1 – the intermediate value theorem (Ross, page 134)

   Suppose that $a < b$, f is continuous on $[a, b]$, and $f(a) < y < f(b)$. Prove that there exists at least one $x \in [a, b]$ such that $f(x) = y$.

   Use Ross's "no bad sequence" definition of continuity, not the epsilon-delta definition. Constructing the appropriate sequences requires some care.

7. Using the intermediate value theorem

   Prove that the function

$$C(x) = 1 - \frac{x^2}{2} + \frac{x^4}{24}$$

   is equal to zero for one and only one value $x \in [1, 2]$.

   This result will be useful when we define π without trigonometry.

8. (Ross, pages 135 and 136; existence of inverse function. This is a composite of Corollary 18.3 and Theorem 18.4)

   Let f be a strictly increasing function on some interval I. Then $f(I)$ is an interval J and there exists a continuous strictly inverse function f^{-1} such that $f^{-1}(f(x)) = x$ for

   A diagram is a great help in visualizing this theorem!

9. Continuity versus uniform continuity (Ross, page 143)

   It's all a matter of the order of quantifiers. For continuity, y is agreed upon before the epsilon-delta game is played. For uniform continuity, a challenge is made using some $\epsilon > 0$, then a δ has to be chosen that meets the challenge independent of y.

   For function f whose domain is a set S:

   - Continuity: $\forall y \in S, \forall \epsilon > 0$,
     $\exists \delta > 0$ such that $\forall x \in S$, $|x - y| < \delta$ implies $|f(x) - f(y)| < \epsilon$.
   - Uniform continuity: $\forall \epsilon > 0$,
     $\exists \delta > 0$ such that $\forall x, y \in S$,$|x - y| < \delta$ implies $|f(x) - f(y)| < \epsilon$.

   Two standard counterexamples, using sets that are not both closed and bounded.

   - On $[0, \infty]$ (not a bounded set), the squaring function is continuous, Show that it is not uniformly continuous.

   - On $(0, 1)$ (not closed) the function $f(x) = \frac{1}{x}$ is continuous. Show that it is not uniformly continuous.

10. Proof 7.2 – when continuity implies uniform continuity (Ross, theorem 19.2)

    Using the Bolzano-Weierstrass theorem, prove that if function f is continuous on the closed interval $[a, b]$, then f is uniformly continuous on $[a, b]$.

11. Proof 7.4 – uniform continuity in terms of sequences (Ross, page 146)

    Prove that if f is uniformly continuous on a set S and (s_n) is a Cauchy sequence in S, then $(f(s_n))$ is a Cauchy sequence.

    Invent an example where f is continuous but not uniformly continuous on S and $(f(s_n))$ is not a Cauchy sequence.

12. Uniform continuity (or lack thereof)

    Let $f(x) = x^2 + \frac{1}{x^2}$.

    Determine whether f is or is not uniformly continuous on each of the following intervals:

    (a) $[1, 2]$

    (b) $(0, 1]$

    (c) $[2, \infty)$

    (d) $(1, 2)$

13. One way to show uniform continuity on an interval that is not closed

    On an open interval (a, b) a function can be continuous without being uniformly continuous. However, if we can *extend* f to a function $\overline{f}$, defined so that $\overline{f}$ is continuous at a and b, then $\overline{f}$ is uniformly continuous on $[a, b]$ and f is uniformly continuous on (a, b).

    Show that on the open interval $(0, \pi)$ the function

    $$f(x) = \frac{1 - \cos x}{x^2}$$

    is uniformly continuous by using the "extension" approach.

14. Ross's non-standard but excellent definition of limit (Ross, page 156)

    S is a subset of $\mathbb{R}$, f is a function defined on S, and a and L are real numbers, ∞ or $-\infty$.

    Then $\lim_{x \to a^S} f(x) = L$ means

    for every sequence (x_n) in S with limit a, we have $\lim(f(x_n)) = L$.

    Suppose that $L_1 = \lim_{x \to a^S} f_1(x)$ and $L_2 = \lim_{x \to a^S} f_2(x)$ exist and are finite.

    Prove that $\lim_{x \to a^S}(f_1 + f_2)(x) = L_1 + L_2$ and
    $\lim_{x \to a^S}(f_1 f_2)(x) = L_1 L_2$.

15. The conventional definition of limit (Ross, page 159;)

    Let f be a function defined on $S \subset \mathbb{R}$, let a be in the closure of S, and let a be a real number.

    Prove that $\lim_{x \to a} f(x) = L$ if and only if

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

    if $x \in S$ and $|x - a| < \delta$, then $|f(x) - L| < \epsilon$.

16. Evaluating limits by brute force

    (a) Use the epsilon-delta definition of limit to prove that $\lim_{x \to 0} \sqrt{|x|} = 0$.

    (b) Use the sequence definition of limit to show that $\lim_{x \to 0} \frac{x}{|x|}$ does not exist.

17. Useful rules for evaluation of limits

Note:

Suppose that $L_1 = \lim_{x \to a^S} f_1(x)$ and $L_2 = \lim_{x \to a^S} f_2(x)$ exist and are finite. a can be $\pm\infty$.

- $\lim_{x \to a^S}(f_1 + f_2)(x) = L_1 + L_2$.
- $\lim_{x \to a^S}(f_1 f_2)(x) = L_1 L_2$.
- $\lim_{x \to a^S}(\frac{f_1}{f_2})(x) = \frac{L_1}{L_2}$, provided $L_2 \neq 0$ and $f_2(x) \neq 0$ for $x \in S$.

Prove the second of these (limit of product = product of limits) by using the corresponding theorem about sequences.

18. Limit of the composition of functions

Suppose that $L = \lim_{x \to a^S} f(x)$ exists and is finite.
Then $\lim_{x \to a^S}(g \circ f)(x) = g(L)$ provided

- g is defined on the set $\{f(x) : x \in S\}$.
- g is defined at L
  (which may just be a limit point of the set $\{f(x) : x \in S\}$.)
- g is continuous at L.

Combine all of these rules and you can pretty much conclude that if you can write a formula for a function that does not involves division by zero, the function is continuous.

19. Limits that involve roots

    Use the sum and product rules for limits to evaluate

$$\lim_{x \to 1} \frac{x^{\frac{1}{3}} - 1}{x - 1}$$

20. Limits that involve trig functions

Using the definition of the sine and tangent functions from right-triangle trigonomtry and the principle that if region R_1 of the plane is a subregion of region R_2, then R_1 has a smaller area than R_2, prove that, for angle $\theta \geq 0$,

$$\sin\theta \leq \theta \leq \tan\theta.$$

Using the squeeze lemma for limits (proof left to the homework), prove that $\lim_{x\to 0} \frac{\sin x}{x} = 1$.

By a clever rewrite to express everything in terms of $\frac{\sin x}{x}$, evaluate

$$\lim_{x\to 0} \frac{\sin 2x - 2\sin x}{x^3}$$

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 7.3)
   Prove that if f and g are real-valued functions that are continuous at $x_0 \in \mathbb{R}$, then $f + g$ is continuous at x_0. Do the proof twice: once using the "no bad sequence" definition of continuity and one using the epsilon-delta definition of continuity.

2. (Ross, page 133)

   Let f be a continuous real-valued function on a closed interval $[a, b]$.
   Using the Bolzano-Weierstrass theorem, prove that f is bounded and that f achieves its maximum value: i.e. $\exists y_0 \in [a, b]$ such that $f(x) \leq f(y_0)$ for all $x \in [a, b]$.

3. (Proof 7.1)
   Suppose that $a < b$, f is continuous on $[a, b]$, and $f(a) < y < f(b)$. Prove that there exists at least one $x \in [a, b]$ such that $f(x) = y$.

   Use Ross's "no bad sequence" definition of continuity, not the epsilon-delta definition.

4. (Proof 7.2)
   Using the Bolzano-Weierstrass theorem, prove that if function f is continuous on the closed interval $[a, b]$, then f is uniformly continuous on $[a, b]$.

5. (Proof 7.4 – uniform continuity and Cauchy sequences)
   Prove that if f is uniformly continuous on a set S and (s_n) is a Cauchy sequence in S, then $(f(s_n))$ is a Cauchy sequence. Invent an example where f is continuous but not uniformly continuous on S and $(f(s_n))$ is not a Cauchy sequence.

6. (Extra topic) Here is Ross's definition of limit:
   S is a subset of $\mathbb{R}$, f is a function defined on S, and a and L are real numbers, ∞ or $-\infty$.

   Then $\lim_{x \to a^S} f(x) = L$ means

   for every sequence (x_n) in S with limit a, we have $\lim(f(x_n)) = L$.

   For the case where a is in the closure of S and L is finite, prove the conventional definition of limit:
   $\lim_{x \to a^S} f(x) = L$ if and only if

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

   if $x \in S$ and $|x - a| < \delta$, then $|f(x) - L| < \epsilon$.

4 Workshop Problems

1. Proofs about continuity

   (a) Prove that if f is continuous at $x_0 \in \mathbb{R}$, and g is continuous at $f(x_0)$, then the composite function $g \circ f$ is continuous at x_0.

   Do two different versions of the proof:

   - Use the "no bad sequence definition" (this is easy, but in case you get stuck, it is also Ross Theorem 7.5)
   - Use the epsilon-delta definition. Using ϵ in the codomain of g and δ in the codomain of f, you will need a third Greek letter in the domain of f. Ross often uses η (eta) in this role.

   (b)
   - The Heaviside function $H(x)$ is defined by
     $H(x) = 0$ for $x < 0, H(x) = 1$ for $x \geq 0$.
     Using the "no bad sequence" definition, prove that H is discontinuous at $x = 0$. (Hint: try a sequence of negative numbers.)
   - Using the epsilon-delta definition of continuity, prove that $f(x) = x^3$ is continuous for arbitrary x_0. (Hint: first deal with the special case $x_0 = 0$, then notice that for small enough δ, $|x| < 2|x_0|$.

2. Intermediate-value theorem

   (a) Using the intermediate-value theorem

   As a congressional intern, you are asked to propose a tax structure for families with incomes in the range 2 to 4 million dollars inclusive. Your boss, who feels that proposing a tax rate of exactly 50% for anyone would be political suicide, wants a function $T(x)$ with the following properties:

   - It is continuous.
   - Its domain is [2,4].
   - Its codomain is [1,2].
   - There is no x for which $2T(x) = x$.

   Prove that this set of requirements cannot be met by applying the intermediate-value theorem to the function $x-2T(x)$, which is negative if the tax rate exceeds 50%.

   Then prove "from scratch" that this set of requirements cannot be met, essentially repeating the proof of the IVT. Hint: Consider the least upper bound s of the set of incomes $S \in [2, 4]$ for which the tax rate is greater than 50 %, and construct sequences (s_n) and (t_n) that converge to s from below and above.

   (b) Continuous functions on an interval that is not closed

   Let $S = [0, 1)$. Invent a sequence $x_n \in S$ that converges to a number $x_0 \notin S$. Hint: try $x_1 = \frac{1}{2}, x_2 = \frac{3}{4}$.

   Then, using this sequence, invent an unbounded continuous function on S and invent a bounded continuous function on S that has no maximum. Explain why you could not do either of these things if S were [0,1].

3. Calculation of limits

(a)
- For i, use the sequence definition of limit and the squeeze lemma.
- For ii, rewrite the expression so that it becomes easy to use the sum and product rules for limits.
- For iii, rewrite the expression so that it becomes easy to use the fact that $\lim_{x\to 0} \frac{\sin x}{x} = 1.$ along with the limit rules.

i. Prove that
$$\lim_{x\to 0} x \sin \frac{1}{x} = 0.$$

ii. Evaluate
$$\lim_{h\to 0} \frac{(x+h)^{\frac{3}{2}} - x^{\frac{3}{2}}}{h}$$

iii. Evaluate
$$\lim_{x\to 0} \frac{\cos 2x - 1}{x^2}$$

(b)
- For i, use the sequence definition of limit and invent sequences that would lead to two different values for L.
- For ii, rewrite the expression so that it becomes easy to use the sum and product rules for limits.
- For iii, rewrite the expression so that it becomes easy to use the fact that $\lim_{x\to 0} \frac{\sin x}{x} = 1.$ along with the limit rules.

i. Show that $\lim_{x\to 0} \sin \frac{1}{x}$ does not exist.

ii. Evaluate
$$\lim_{x\to\infty} (\sqrt{x+1} - \sqrt{x})$$

iii. Evaluate
$$\lim_{x\to 0} \frac{\tan x - \sin x}{x^3}$$

5 Homework

Special offer – if you do the entire problem set, with one problem omitted, in LaTeX, you will receive full credit for the omitted problem. Alternatively, if you work all the problems in LaTeX, we will convert your lowest score to a perfect score.

1. Ross, exercises 19.2(b) and 19.2(c). Be sure that you prove *uniform* continuity, not just continuity!

2. Ross, exercise 19.4.

3. Ross, exercises 20.16 and 20.17. This squeeze lemma is a cornerstone of elementary calculus, and it is nice to be able to prove it!

4. Ross, exercise 20.18. Be sure to indicate where you are using various limit theorems.

5. Ross, exercise 17.4. It is crucial that the value of δ is allowed to depend on x.

6. Ross, exercises 17-13a and 17-14. These functions will be of interest when we come to the topic of integration in the spring term.

7. Ross, exercise 18-4. To show that something exists, describe a way to construct it.

8. Ross, exercise 18-10. You may use the intermediate-value theorem to prove the result.

1. Prove that the "no bad sequence" definition of continuity holds iff the ϵ-δ definition holds.
   (a) Pro tip: I'd usually use the ϵ-δ definition to show continuity and the "no bad sequence" definition to show discontinuity. It's often not too bad to come up with a bad sequence for showing discontinuity, but showing that something is continuous by the "no bad sequence" criterion is often not too fun.
   (b) Using the "no bad sequence" definition, show that $f(x) = x + 1$ is convergent for any $x \in (0, 1)$.
   (c) Using the ϵ-δ definition, show that $f(x) = x + 1$ is convergent for any $x \in (0, 1)$.
   (d) Using the "no bad sequence" definition, show the following sequence (defined on [0,2]) is divergent at $x = 1$:
   $$f(x) = \begin{cases} x+1 & 0 \le x \le 1 \\ x & 1 < x \le 2 \end{cases}$$
   (e) Using the ϵ-δ definition, show the following sequence (defined on [0,2]) is divergent at $x = 1$:
   $$f(x) = \begin{cases} x+1 & 0 \le x \le 1 \\ x & 1 < x \le 2 \end{cases}$$

2. Continuity and uniform continuity
   (a) To prove that $f(x)$ is continuous **at a particular** x_0:
   $\forall \epsilon > 0, \exists \delta > 0$ s.t. $\forall x, |x - x_0| < \delta \to |f(x) - f(x_0)| < \epsilon$
   (b) To prove that $f(x)$ is continuous **everywhere**:
   $\forall x_0, \forall \epsilon > 0, \exists \delta > 0$ s.t. $\forall x, |x - x_0| < \delta \to |f(x) - f(x_0)| < \epsilon$
   (c) To prove that $f(x)$ is **uniformly continuous**:
   $\forall \epsilon > 0, \exists \delta > 0$ s.t. $\forall x, \forall x_0, |x - x_0| < \delta \to |f(x) - f(x_0)| < \epsilon$
   (d) Proof 7.2: If a function f is continuous on a closed interval, it is uniformly continuous on that interval.
   (e) It is a **sufficient but not necessary** criterion for uniform continuity of f on (a, b) that f be differentiable on (a, b), with f' bounded on (a, b).
   (f) Prove that $f(x) = \frac{1}{x}$ is **not uniformly continuous** on $(0, \infty)$.
   (g) Prove that $\sqrt{x}$ is **uniformly continuous** on $[0, \infty)$. (You may use that, as part of a pset problem, you prove(d) that it is continuous on $[0, \infty)$.) Hint: Break the domain into two (overlapping) pieces, [0,2] and $[1, \infty)$.
   (h) Ross 18.6: Prove that $x = \cos(x)$ for some $x \in (0\frac{pi}{2})$.

3. Prove that a function $f(x) : \mathbb{R} \to \mathbb{R}$ is injective iff it is either strictly increasing or strictly decreasing.

4. Ross 17.12: Let f be a continuous real-valued function with domain (a, b). Show that if $f(r) = 0$ for each rational number $r \in (a, b)$, then $f(x) = 0$ $\forall x \in (a, b)$.

MATHEMATICS 23a/E-23a, Fall 2018
Linear Algebra and Real Analysis I
Week 8 (Derivatives, Inverse functions, Taylor series)

Authors: Paul Bamberg and Kate Penner (based on their course MATH S-322)
R scripts by Paul Bamberg
Last modified: November 3, 2018 by Paul Bamberg (fixed proof 8.2)

Reading from Ross

- Chapter 5, sections 28 and 29 (pp.223-240)
- Chapter 5, sections 30 and 31, but only up through section 31.7.
- Chapter 7, section 37 (logarithms and exponentials)

Recorded Lectures

- Lecture 16 (Week 8, Class 1) (watch on October 30 or 31)
- Lecture 17 (Week 8, Class 2) (watch on November 1 or 29)

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

- 8.1 (Ross, pp.233-234, Rolle's Theorem and the Mean Value Theorem)

  - Prove Rolle's Theorem: if f is a continuous function on $[a, b]$ that is differentiable on (a, b) and satisfies $f(a) = f(b)$, then there exists at least one x in (a, b) such that $f'(x) = 0$.

  - Using Rolle's Theorem, prove the Mean Value Theorem: if f is a continuous function on $[a, b]$ that is differentiable on (a, b), then there exists at least one x in (a, b) such that

$$f'(x) = \frac{f(b) - f(a)}{b - a}$$

- 8.2 Differentiating an inverse function
  Suppose that f is a one-to-one continuous function on open interval I (either strictly increasing or strictly decreasing) Let open interval $J = f(I)$, and define the inverse function $f^{-1} : J \to I$ for which
  $(f^{-1} \circ f)(x) = x$ for $X \in I$; $f \circ f^{-1}(y) = y$ for $y \in J$.
  Let $g = f^{-1}$, and define $y_0 = f(x_0)$.
  Take it as proved that g is continuous at y_0.
  Prove that, if f if differentiable at x_0 and $f'(x_0) \neq 0$, then

$$\lim_{y \to y_0} \frac{g(y) - g(y_0)}{y - y_0} = \frac{1}{f'(x_0)}.$$

Additional proofs(may appear on quiz, students will post pdfs or videos)

- 8.3 (Ross, pp. 228, The Chain Rule – easy special case)
  Assume the following:
  - Function f is differentiable at a.
  - Function g is differentiable at $f(a)$.
  - There is an open interval J containing a on which f is defined and $f(x) \neq f(a)$ (without this restriction, you need the messy Case 2 on page 229).
  - Function g is defined on the open interval $I = f(J)$, which contains $f(a)$.

  Using the sequential definition of a limit, prove that the composite function $g \circ f$ is defined on J and differentiable at a and that

$$(g \circ f)'(a) = g'(f(a)) \cdot f'(a).$$

- 8.4 Taylor's Theorem with remainder:
  Let f be defined on (a, b) with $a < 0 < b$.
  Suppose that the nth derivative $f^{(n)}$ exists on (a, b).
  Define the remainder

$$R_n(x) = f(x) - \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} x^k.$$

  Prove, by repeated use of Rolle's theorem, that for each $x \neq 0$ in (a, b), there is some y between 0 and x for which

$$R_n(x) = \frac{f^{(n)}(y)}{n!} x^n.$$

R Scripts

- Script 2.4A-Taylor Series.R
  Topic 1 - Convergence of the Taylor series for the cosine function
  Topic 2 - A function that is not the sum of its Taylor series
  Topic 3 - Illustrating Ross's proof of Taylor series with remainder.

- Script2.4B-LHospital.R Topic 1 - Illustration of proof 6 from Week 8

- Script 2.4C-SampleProblems.R

1 Executive Summary

1.1 The Derivative - Definition and Properties

- A function f is differentiable at some point a if the limit

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

exists and is finite. It is referred to as $f'(a)$. If a function is differentiable at a point a, then it is continuous at a as well.

- Derivatives, being defined in terms of limits, share many properties with limits. Given two functions f and g, both differentiable at some point a, the following properties hold:
  - scalar multiples: $(cf)'(a) = c \cdot f'(a)$
  - sums of functions: $(f + g)'(a) = f'(a) + g'(a)$
  - Product Rule: $(fg)'(a) = f(a)g'(a) + f'(a)g(a)$
  - Quotient Rule: $(f/g)'(a) = [g(a)f'(a) - f(a)g'(a)]/g^2(a)$ if $g(a) \neq 0$
- The most memorable derivative rule is **The Chain Rule**, which states that if f is differentiable at some point a, and g is differentiable at $f(a)$, then their composite function $g \circ f$ is also differentiable at a, and

$$(g \circ f)'(a) = g'(f(a)) \cdot f'(a)$$

1.2 Increasing and decreasing functions

The terminology is the same as what we used for sequences. It applies to functions whether or not they are differentiable or even continuous.

- A function f is **strictly increasing** on an interval I if $x_1, x_2 \in I$ and $x_1 < x_2 \implies f(x_1) < f(x_2)$
- A function f is **strictly decreasing** on an interval I if $x_1, x_2 \in I$ and $x_1 < x_2 \implies f(x_1) > f(x_2)$
- A function f is **increasing** on an interval I if $x_1, x_2 \in I$ and $x_1 < x_2 \implies f(x_1) \leq f(x_2)$
- A function f is **decreasing** on an interval I if $x_1, x_2 \in I$ and $x_1 < x_2 \implies f(x_1) < f(x_2)$

1.3 Behavior of differentiable functions

These justify our procedures when we are searching for the critical points of a given function. They are the main properties we draw on when reasoning about a function's behavior.

- If f is defined on an open interval, achieves its maximum or minimum at some x_0, and is differentiable there, then $f'(x_0) = 0$.
- Rolle's Theorem. If f is continuous on some interval $[a, b]$ and differentiable on (a, b) with $f(a) = f(b)$, then there exists at least one $x \in (a, b)$ such that $f'(x) = 0$ (Rolle's Theorem).
- Mean Value Theorem. If f is continuous on some interval $[a, b]$ and differentiable on (a, b), then there exists at least one $x \in (a, b)$ such that

$$f'(x) = \frac{f(b) - f(a)}{b - a}$$

- If f is differentiable on (a, b) and $f'(x) = 0\ \forall x \in (a, b)$, then f is a constant function on (a, b).
- If f and g are differentiable functions on (a, b) such that $f' = g'$ on (a, b), then there exists a constant c such that
  $\forall x \in (a, b)\ f(x) = g(x) + c$

1.4 Inverse functions and their derivatives

- Review of a corollary of the intermediate value theorem: If function f is continuous and one-to-one on a interval I(which means it must be either strictly increasing or strictly decreasing), then there is a continuous inverse function f^{-1}, whose domain is the interval
  $J = f(I)$, such that $f \circ f^{-1}$ and $f^{-1} \circ f$ are both the identity function.
- Not quite a proof: Since $(f \circ f^{-1})(y) = y$, the chain rule states that $f'(f^{-1}(y))(f^{-1})'(y) = y$ and, if $f'(f^{-1}(y)) \neq 0$,

$$(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}.$$

- Example: if $f(x) = \tan x$ with $I = (\frac{-\pi}{2}, \frac{\pi}{2})$, then $f^{-1}(y) = \arctan y$ and

$$(\arctan)'(y) = \frac{1}{(\tan)'(\arctan y)} = \frac{1}{\sec^2(\arctan y)} = \frac{1}{1 + \tan^2(\arctan y)} = \frac{1}{1 + y^2}$$

- The problem: we need to prove that f' is differentiable.

1.5 Defining the logarithm and exponential functions

Define the natural logarithm as an antiderivative:

$$L(y) = \int_1^y \frac{1}{t} dt, \text{ and define } e \text{ so that } \int_1^e \frac{1}{t} dt = 1.$$

¿From this definition it is easy to prove that $L'(y) = \frac{1}{y}$ and not hard to prove that $L(xy) = L(x) + L(y)$.

Now the exponential function can be defined as the inverse function, so that $E(L(y)) = y$. From this definition it follows that $E(x + y) = E(x) + E(y)$ and that $E'(x) = E(x)$.

1.6 L'Hospital's rule

- Suppose that f and g are differentiable functions and that

$$\lim_{x \to a+} \frac{f'(x)}{g'(x)} = L; \lim_{x \to a^+} f(x) = \lim_{x \to a^+} g(x) = 0; g'(a) < 0.$$

  Then

$$\lim_{x \to a+} \frac{f(x)}{g(x)} = L.$$

- Replace $x \to a^+$ by $x \to a^-$ or $x \to a$ or $x \to \pm\infty$ and the result is still valid. It is also possible to have $\lim_{x \to a^+} f(x) = \lim_{x \to a^+} g(x) = \infty$. The restriction to $g'(a) < 0$ is just to make the proof easier; the result is also true if $g'(a) > 0$.

- Once you understand the proof in one special case, the proof in all the other cases is essentially the same.

- Here is the basic strategy: given that

$$\lim_{x \to a} \frac{f'(x)}{g'(x)} = L,$$

  use the mean value theorem to construct an interval (a, α) on which

$$|\frac{f(x)}{g(x)} - L| < \epsilon.$$

1.7 Taylor series

- If a function f is *defined* by a convergent power series, i.e.

$$f(x) = \sum_{k=0}^{\infty} a_k x^k \text{ for } |x| < R,$$

then it is easy to show that

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k \text{ for } |x| < R.$$

The challenge is to extend this formula to functions that are differentiable many times but that are not defined by power series, like trig functions defined geometrically, or the function $\sqrt{(1+x)}$.

- Taylor's theorem with remainder – version 1
By the mean value theorem, $f(x) - f(0) = f'(y)x$ for some $y \in (0, x)$.

The generalization is that

$$f(x) - f(0) - f'(0)x - \frac{f''(0)}{2!}x^2 - \cdots - \frac{f^{(n-1)}(0)}{(n-1)!}x^{n-1} = \frac{f^{(n)}(y)}{n!}x^n$$

for some y between 0 and x. It is proved by induction, using Rolle's theorem n times.

- If the right hand side approaches zero in the limit of large n, then the Taylor series converges to the function. This is true if all the derivatives $f^{(n)}$ are bounded by a single constant C. This criterion is sufficient to establish familiar Taylor expansions like
$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \cdots$
$\cos x = 1 - \frac{x^2}{2} + \frac{x^4}{4!} + \cdots$

- Taylor's theorem with remainder – version 2
The fundamental theorem of calculus says that $f(x) - f(0) = \int_0^x f'(t)dt$.

The generalization is that

$$f(x) - f(0) - f'(0)x - \frac{f''(0)}{2!}x^2 - \cdots - \frac{f^{(n-1)}(0)}{(n-1)!}x^{n-1} = \int_0^x \frac{(x-t)^{n-1}}{(n-1)!} f^{(n)}(t)dt.$$

It is proved by induction, using integration by parts, but not by us!

- A famous counterexample.
The function $f(x) = e^{-\frac{1}{x}}$ for $x > 0$ and $f(x) = 0$ for $x \leq 0$ has the property that the remainder does not approach a limit of zero. It does not equal the sum of its Taylor series.

2 Lecture Outline

1. The Derivative - Definition and Properties

   A function f is differentiable at some point a if the limit

   $$\lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

   exists and is finite. It is referred to as $f'(a)$. If a function is differentiable at a point a, then it is continuous at a as well.

2. Derivatives, being defined in terms of limits, share many properties with limits. Given two functions f and g, both differentiable at some point a, the following properties hold:

   - scalar multiples: $(cf)'(a) = c \cdot f'(a)$
   - sums of functions: $(f + g)'(a) = f'(a) + g'(a)$
   - Product Rule: $(fg)'(a) = f(a)g'(a) + f'(a)g(a)$
   - Quotient Rule: $(f/g)'(a) = [g(a)f'(a) - f(a)g'(a)]/g^2(a)$ if $g(a) \neq 0$

3. (Ross, p.226, Sum and Product Rule for Derivatives)
   Consider two functions f and g. Prove that if both functions are differentiable at some point a, then $(f+g)$ and fg are differentiable at a as well, and:

   - $(f+g)'(a) = f'(a) + g'(a)$
   - $(fg)'(a) = f(a)g'(a) + f'(a)g(a)$

4. Proving differentiation formulas by induction

   Using the product rule, which we just proved, show by induction that for the function $f(x) = x^n$

   $f'(x) = nx^{n-1}$ for all $n > 0$.

   Using the product rule, show by induction that for the function $g(x) = x^{-n}$

   $g'(x) = -nx^{-(n+1)}$ for all $n > 0$

5. (Proof 8.3 – Ross, pp. 228, The Chain Rule – easy special case)
Assume the following:

   - Function f is differentiable at a.
   - Function g is differentiable at $f(a)$.
   - There is an open interval J containing a on which f is defined and $f(x) \neq f(a)$ (without this restriction, you need the messy Case 2 on page 229).
   - Function g is defined on the open interval $I = f(J)$, which contains $f(a)$.

   Using the sequential definition of a limit, prove that the composite function $g \circ f$ is defined on J and differentiable at a and that

$$(g \circ f)'(a) = g'(f(a)) \cdot f'(a).$$

6. Calculating derivatives by using the chain rule

   Let $f(x) = \sqrt[3]{x}$.

   (a) Calculate $f'(x)$ using the definition of the derivative.

   (b) Calculate $f'(x)$ by applying the chain rule to $(f(x))^3 = x$.

   (c) Using the chain rule, which we just proved, show that for the function $f(x) = x^{m/n}$ with $m, n > 0$

$$f'(x) = \frac{m}{n} x^{\frac{m}{n}-1}.$$

7. Behavior of differentiable functions

   These justify our procedures when we are searching for the critical points of a given function. They are the main properties we draw on when reasoning about a function's behavior.

   - If f is defined on an open interval, achieves its maximum or minimum at some x_0, and is differentiable there, then $f'(x_0) = 0$.
   - Rolle's Theorem. If f is continuous on some interval $[a, b]$ and differentiable on (a, b) with $f(a) = f(b)$, then there exists at least one $x \in (a, b)$ such that $f'(x) = 0$ (Rolle's Theorem).
   - Mean Value Theorem. If f is continuous on some interval $[a, b]$ and differentiable on (a, b), then there exists at least one $x \in (a, b)$ such that
   $$f'(x) = \frac{f(b) - f(a)}{b - a}$$
   - If f is differentiable on (a, b) and $f'(x) = 0\ \forall x \in (a, b)$, then f is a constant function on (a, b).
   - If f and g are differentiable functions on (a, b) such that $f' = g'$ on (a, b), then there exists a constant c such that
   $\forall x \in (a, b)\ f(x) = g(x) + c$

8. The derivative at a maximum or minimum (Ross, page 232)
   Prove that if f is defined on an open interval containing x_0, if f has its maximum of minimum at x_0, and if f is differentiable at x_0, then $f'(x_0) = 0$.

9. (Proof 8.1 – Ross, pp.233-234, Rolle's Theorem and Mean Value Theorem) Prove Rolle's Theorem: if f is a continuous function on $[a, b]$ that is differentiable on (a, b) and satisfies $f(a) = f(b)$, then there exists at least one x in (a, b) such that $f'(x) = 0$.

Using Rolle's Theorem, prove the Mean Value Theorem: if f is a continuous function on $[a, b]$ that is differentiable on (a, b), then there exists at least one x in (a, b) such that

$$f'(x) = \frac{f(b) - f(a)}{b - a}$$

10. Using the Mean Value Theorem

    Suppose f is differentiable on $\mathbb{R}$ and $f(0) = 0$, $f(1) = 1$, and $f(2) = 1$. Show that $f'(x) = 1/2$ for some $x \in (0, 2)$.
    Then, by applying the Intermediate Value Theorem and Rolle's Theorem to $g(x) = f(x) - \frac{1}{4}x$, show that $f'(x) = \frac{1}{4}$ for some $x \in (0, 2)$.

11. Increasing and decreasing functions
    The terminology is the same as what we used for sequences. It applies to functions whether or not they are differentiable or even continuous.

    - A function f is **strictly increasing** on an interval I if $x_1, x_2 \in I$ and $x_1 < x_2 \implies f(x_1) < f(x_2)$
    - A function f is **strictly decreasing** on an interval I if $x_1, x_2 \in I$ and $x_1 < x_2 \implies f(x_1) > f(x_2)$
    - A function f is **increasing** on an interval I if $x_1, x_2 \in I$ and $x_1 < x_2 \implies f(x_1) \leq f(x_2)$
    - A function f is **decreasing** on an interval I if $x_1, x_2 \in I$ and $x_1 < x_2 \implies f(x_1) < f(x_2)$

    Prove that if f is a differentiable function on an interval (a, b) and $f'(x) > 0 \; \forall x \in (a, b)$, then f is strictly increasing.

12. Defining inverse functions

A function $f(x)$, defined on an interval I, has an inverse function $g(y)$ for which $f(g(y)) = y)$ and $g(f(x)) = x)$ only if it is monotone (either increasing or decreasing) on I.
If I is an open interval (a, b) on which f is differentiable, the inverse function is differentiable everywhere on the interval $J = f(I)$ if and only if f is either strictly increasing or strictly decreasing.

(a) Sketch a graph for the case $f(x) = \sin x, I = [\frac{-\pi}{2}, \frac{\pi}{2}]$ to show how the arc sine function is defined.

(b) Although $f(x) = \sin x$ is strictly increasing on the closed interval I, the derivative rule requires an open interval! Show that $g(y) = \arcsin(y)$ is differentiable on the open interval $(-1, 1)$.

13. Proof 8.2 – Differentiating an inverse function
Suppose that f is a one-to-one continuous function on open interval I (either strictly increasing or strictly decreasing) Let open interval $J = f(I)$, and define the inverse function $f^{-1} : J \to I$ for which

$(f^{-1} \circ f)(x) = x$ for $X \in I$; $f \circ f^{-1}(y) = y$ for $y \in J$.

Let $g = f^{-1}$, and define $y_0 = f(x_0)$.
Take it as proved that g is continuous at y_0.

Prove that, if f if differentiable at x_0, then

$$\lim_{y \to y_0} \frac{g(y) - g(y_0)}{y - y_0} = \frac{1}{f'(x_0)}.$$

14. Applying the inverse-function rule

    The function $g(y) = \arctan y^2, y \geq 0$ is continuous and strictly increasing, hence invertible.
    Calculate its derivative by finding a formula for the inverse function $f(x)$, which is easy to differentiate, then using the rule for the derivative of an inverse function. You can confirm your answer by using the known derivative of the arctan function.

15. (L'Hospital's Rule; based on Ross, 30.2, but simplified to one special case)
Suppose that f and g are differentiable functions and that

$$\lim_{z \to a+} \frac{f'(z)}{g'(z)} = L; f(a) = 0, g(a) = 0; g'(a) > 0.$$

Choose $x > a$ so that for $a < z \leq x, g(z) > 0$ and $g'(z) > 0$.
(You do not have to prove that this can always be done!)

By applying Rolle's Theorem to $h(z) = f(z)g(x) - g(z)f(x)$,

prove that

$$\lim_{x \to a+} \frac{f(x)}{g(x)} = L.$$

16. Using L'Hospital's rule – tricks of the trade

(a) Conversion to a quotient – evaluate

$$\lim_{x \to 0+} x \log_e x^2.$$

(b) Evaluate

$$\lim_{x \to 0} \frac{xe^x - \sin x}{x^2}$$

both by using L'Hospital's rule and by expansion in a Taylor series.

17. Taylor series

- If a function f is *defined* by a convergent power series, i.e.

$$f(x) = \sum_{k=0}^{\infty} a_k x^k \text{ for } |x| < R,$$

then it is easy to show that

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k \text{ for } |x| < R.$$

The challenge is to extend this formula to functions that are differentiable many times but that are not defined by power series, like trig functions defined geometrically, or the function $\sqrt{(1+x)}$.

- Taylor's theorem with remainder
By the mean value theorem, $f(x) - f(0) = f'(y)x$ for some $y \in (0, x)$.
The generalization is that

$$f(x) - f(0) - f'(0)x - \frac{f''(0)}{2!}x^2 - \cdots - \frac{f^{(n-1)}(0)}{(n-1)!}x^{n-1} = \frac{f^{(n)}(y)}{n!}x^n$$

for some y between 0 and x. It is proved by induction, using Rolle's theorem n times.

- If the right hand side approaches zero in the limit of large n, then the Taylor series converges to the function. This is true if, as is often the case, all the derivatives $f^{(n)}$ are bounded by a single constant C.

18. (Proof 8.4 – Ross, page 250; version 1 of Taylor's Theorem with remainder, setting $c = 0$)

    Let f be defined on (a, b) with $a < 0 < b$. Suppose that the nth derivative $f^{(n)}$ exists on (a, b).

    Define the remainder

$$R_n(x) = f(x) - \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} x^k.$$

    Prove, by repeated use of Rolle's theorem, that for each $x \neq 0$ in (a, b), there is some y between 0 and x for which

$$R_n(x) = \frac{f^{(n)}(y)}{n!} x^n.$$

19. Calculating a Taylor series

(a) Derive the Taylor series for the function $f(x) = \cos x$.

(b) Prove that the series converges for all x.

(c) Use an appropriate form of remainder to prove that it converges to the cosine function.

20. A function that is not equal to the sum of its Taylor series

(a) Use L'Hospital's rule to show that for all $n > 0$,

$$\lim_{x \to 0} x^{-n} e^{-\frac{1}{x}} = 0.$$

(b) Show that the function defined by

$$f(x) = 0 \text{ for } x \leq 0, f(x) = e^{-\frac{1}{x}} \text{ for } x > 0$$

is not equal to the sum of its Taylor series.

21. (Ross, pp. 342-343; defining the natural logarithm)

Define

$$L(y) = \int_1^y \frac{1}{t} dt.$$

Prove from this definition the following properties of the natural logarithm:

- $$L'(y) = \frac{1}{y} \text{ for } y \in (0, \infty).$$
- $L(yz) = L(y) + L(z)$ for $y, z \in (0, \infty)$.
- $\lim_{y \to \infty} L(y) = +\infty$.

22. Definition and properties of the exponential function

Denote the function inverse to L by E, i.e.

$$(E(L(y)) = y \text{ for } y \in (0, \infty)$$
$$L(E(x)) = x \text{ for } x \in \mathbb{R}$$

Prove from this definition the following properties of the exponential function E:

- $E'(x) = E(x)$ for $x \in \mathbb{R}$.
- $E(u + v) = E(u)E(v)$ for $u, v \in \mathbb{R}$.

23. Hyperbolic functions, defined by their Taylor series

$$\sinh x = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots ; \cosh x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots$$

- Calculate $\sinh' x$ and $\cosh' x$, and prove that $\cosh^2 x - \sinh^2 x = 1$.
- Use Taylor's theorem to prove that
  $\sinh(a + x) = \sinh a \cosh x + \cosh a \sinh x$.

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 8.3 – The Chain Rule – easy special case)
   Assume the following:

   - Function f is differentiable at a.
   - Function g is differentiable at $f(a)$.
   - There is an open interval J containing a on which f is defined and $f(x) \neq f(a)$ (without this restriction, you need the messy Case 2 on page 229).
   - Function g is defined on the open interval $I = f(J)$, which contains $f(a)$.

   Using the sequential definition of a limit, prove that the composite function $g \circ f$ is defined on J and differentiable at a and that

   $$(g \circ f)'(a) = g'(f(a)) \cdot f'(a).$$

2. (Proof 8.1 – Rolle's Theorem and the Mean Value Theorem)

   - Prove Rolle's Theorem: if f is a continuous function on $[a, b]$ that is differentiable on (a, b) and satisfies $f(a) = f(b)$, then there exists at least one x in (a, b) such that $f'(x) = 0$.

   - Using Rolle's Theorem, prove the Mean Value Theorem: if f is a continuous function on $[a, b]$ that is differentiable on (a, b), then there exists at least one x in (a, b) such that

   $$f'(x) = \frac{f(b) - f(a)}{b - a}$$

3. (Proof 8.2 – Differentiating an inverse function)
   Suppose that f is a one-to-one continuous function on open interval I (either strictly increasing or strictly decreasing) Let open interval $J = f(I)$, and define the inverse function $f^{-1} : J \to I$ for which

   $(f^{-1} \circ f)(x) = x$ for $X \in I$; $f \circ f^{-1}(y) = y$ for $y \in J$.

   Let $g = f^{-1}$, and define $y_0 = f(x_0)$.
   Take it as proved that g is continuous at y_0.

   Prove that, if f if differentiable at x_0 and $f'(x_0) \neq 0$, then

$$\lim_{y \to y_0} \frac{g(y) - g(y_0)}{y - y_0} = \frac{1}{f'(x_0)}.$$

4. Suppose that function f is *defined* by a convergent power series, i.e.

$$f(x) = \sum_{k=0}^{\infty} a_k x^k \text{ for } |x| < R,$$

   Prove that in this case the Taylor series formula

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k \text{ for } |x| < R$$

   is correct.

5. (Proof 8.4 – Taylor's Theorem with remainder)
   Let f be defined on (a, b) with $a < 0 < b$.
   Suppose that the nth derivative $f^{(n)}$ exists on (a, b).

   Define the remainder

$$R_n(x) = f(x) - \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} x^k.$$

   Prove, by repeated use of Rolle's theorem, that for each $x \neq 0$ in (a, b), there is some y between 0 and x for which

$$R_n(x) = \frac{f^{(n)}(y)}{n!} x^n.$$

6. (Extra topic) The derivative at a maximum or minimum (Ross, page 232)
   Prove that if f is defined on an open interval containing x_0, if f has its maximum or minimum at x_0, and if f is differentiable at x_0, then $f'(x_0) = 0$.

4 Workshop Problems

1. Proving differentiation rules

   (a) Trig functions

   - Prove that $(\sin x)' = \cos x$ from scratch using the fact that

   $$\lim_{x \to 0} \frac{\sin x}{x} = 1$$

   - Let $f(x) = \csc x$ so that $\sin x f(x) = 1$. Use the product rule to prove that

   $$(\csc x)' = -\csc x \cot x.$$

   (b) Non-integer exponents

   - Negative rational exponent: Let $f(x) = x^{-m/n}$, so that $(f(x))^n x^m = 1$.
     Prove that

     $$f'(x) = \frac{-m}{n} x^{\frac{-m}{n}-1}.$$

   - Irrational exponent:
     Let p be any real number and define $f(x) = x^p = E(pL(x))$.
     Prove that $f'(x) = px^{p-1}$.

2. MVT, L'Hospital, inverse functions

(a)
- When a local minimum is also a global minimum
  Suppose that f is twice differentiable on (a, b), with $f'' > 0$, and that there exists $x \in (a, b)$ for which $f'(x) = 0$, so that x is a local minimum of f. Consider $y \in (x, b)$. By using the mean value theorem twice, prove that $f(y) > f(x)$. This, along with a similar result for $y \in (a, x)$, establishes that x is also the global minimum of f on (a, b).
- Evaluate the limit

$$\lim_{x \to 0} \frac{1 - \cos x}{e^x - x - 1}$$

  by using L'Hospital's rule, then confirm your answer by expanding both numerator and denominator in a Taylor series.

(b)
- Applying the inverse-function rule
  The function $g(y) = \arcsin \sqrt{y}, 0 < y < 1$ is important in the theory of random walks.
  Calculate its derivative by finding a formula for the inverse function $f(x)$, which is easy to differentiate, then using the rule for the derivative of an inverse function. You can confirm your answer by using the known derivative of the arcsin function.
- Evaluate the limit

$$\lim_{x \to 0} \frac{\csc x - \cot x}{x}.$$

  It takes a little bit of algebraic work to rewrite this in a form to which L'Hospital's rule can be applied.

3. Taylor series

   (a) Using the Taylor series for the trig functions
   Define functions $S(x)$ and $C(x)$ by the power series

$$S(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots ; C(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots$$

   - Calculate $S'(x)$ and $C'(x)$, and prove that $S^2(x) + C^2(x) = 1$.
   - Use Taylor's theorem to prove that
     $C(a + x) = C(a)C(x) - S(a)S(x)$.

   (b) Using the remainder to prove convergence
   Define $f(x) = \log_e(1 + x)$ for $x \in (-1, \infty)$.
   Using the remainder formula

$$R_n(x) = \frac{f^{(n)}(y)}{n!}x^n$$

   prove that

$$\log_e 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \cdots .$$

   Show that the remainder does not go to zero if you set $x = -1$.

5 Homework

Again, if you do the entire assignment in TeX, you may omit one problem and receive full credit for it.

1. Ross, 28.2
2. Ross, 28.8
3. Ross, 29.12
4. Ross, 29.18
5. Ross, exercises 30-1(d) and 30-2(d). Do these two ways: once by using L'Hospital's rule, once by replacing each function by the first two or three terms of its Taylor series.
6. Ross, 30-4. Use the result to convert exercise 30-5(a) into a problem that involves a limit as $y \to \infty$.
7. One way to define the exponential function is as the sum of its Taylor series:

   $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$.

   Using this definition and Taylor's theorem, prove that $e^{a+x} = e^a e^x$.
8. Ross, exercise 31.5. For part (a), just combine the result of example 3 (whose messy proof you need not study) with the chain rule.
9. Ross, exercise 37.9.

1. Using the mean value theorem, prove that if a function f is differentiable on (a, b) and $f'(x) = 0$ for all $x \in (a, b)$, then f is a constant function on (a, b).

2. Taylor series practice: Compute the Taylor series (around $x = 0$, expressed in closed summation form) for the polynomial $f(x) = e^{-x}$. (Do the whole Taylor series process for this. Afterwards, see if you can come up with an easy way to come up with this Taylor series given the Taylor series for e^x, which you might have memorized.)

3. Use the remainder formula (Taylor's Theorem with Remainder) to prove that $e = 1 + 1 + \frac{1}{2} + \frac{1}{6} + ...$ (Do this without explicitly referencing the Taylor series for e^x.)

4. Calculate the derivative of inverse of $f(x)$ in the following cases, you may assume the result from proof 8.2, and that the inverse functions are differentiable.

   (a) $f(x) = cos(x)$

   (b) $f(x) = tan(x)$

   (c) $f(x) = e^x$

5. Use Taylor's theorem with remainder and the function $f(x) = \sqrt{1 + x}$ to show that

$$\lim_{n \to \infty} \sqrt{n + \sqrt{n}} - \sqrt{n} = \frac{1}{2} \tag{1}$$

   (Source:mathcs.org)

MATHEMATICS 23a/E-23a, Fall 2018
Linear Algebra and Real Analysis I
Week 9 (Topology, sequences and series of vectors and matrices)

Author: Paul Bamberg
R scripts by Paul Bamberg
Last modified: August 13, 2018 by Paul Bamberg

Reading

- Hubbard, Section 1.5. The only topology that is treated is the "open-ball topology."

Alas, Hubbard does not mention either finite topology or differential equations. I have included a set of notes on these topics that I wrote for Math 121.

Recorded Lectures

- Lecture 18 (Week 9, Class 1) (watch on November 6 or 7)
- Lecture 19 (Week 9, Class 2) (watch on November 8 or 9)

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

- 9.1
  - Define "Hausdorff space," and prove that in a Hausdorff space the limit of a sequence is unique.
  - Prove that $\mathbb{R}^n$, with the topology defined by open balls, is a Hausdorff space.
- 9.2 Starting from the triangle inequality for two vectors, prove the triangle inequality for m vectors in $\mathbb{R}^n$, then prove the "infinite triangle inequality:"

$$|\sum_{i=1}^{\infty} \vec{\mathbf{a}_\mathbf{i}}| \leq \sum_{i=1}^{\infty} |\vec{\mathbf{a}_\mathbf{i}}|$$

  You may assume that the series $\sum_{i=1}^{\infty} \vec{\mathbf{a}_\mathbf{i}}$ is "absolutely summable" (the infinite series of lengths on the right is convergent) but you must prove that this series is "summable" (infinite sum of vectors on the left is convergent.) As on page 100 of the textbook, you may use theorems 0.5.8 (if $\sum_{n=1}^{\infty} |a_n|$ converges, then so does $\sum_{n=1}^{\infty} a_n$) and 1.5.13 (a sequence of vectors in $\mathbb{R}^n$ converges if and only if each component converges).

R Scripts

- Script 3.1A-FiniteTopology.R
  Topic 1 - The "standard" Web site graph, used in notes and examples
  Topic 2 - Drawing a random graph to create a different topology on the same set

- Script 3.1B-SequencesSeriesRn.R
  Topic 1 - A convergent sequence of points in $\mathbb{R}^2$
  Topic 2 - A convergent infinite series of vectors
  Topic 3 - A convergent geometric series of matrices

- Script 3.1C-DiffEquations.R
  Topic 1 - Two real eigenvalues
  Topic 2 - A repeated real eigenvalue
  Topic 3 - Complex conjugate eigenvalues

1 Executive Summary

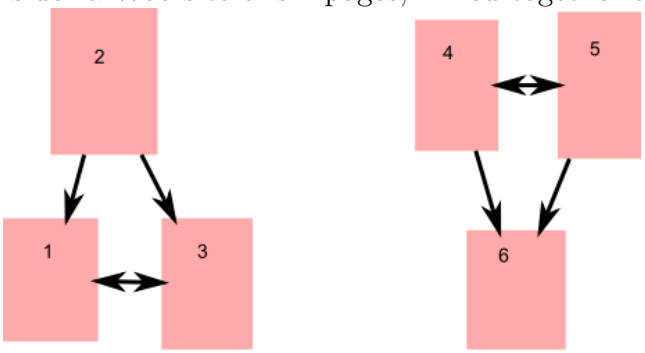
1.1 Axioms of Topology

In topology, we start with a set X and single out some of its subsets as "open sets." The only requirement on a topology is that the collection of open sets satisfies the following rules (axioms)

- The empty set and the set X are both open.
- The union of any finite or infinite collection number of open sets is open.
- The intersection of two open sets is open. It follows by induction that the intersection of n open sets is open, but the intersection of infinitely many open sets is not necessarily open.

1.2 A Web-site model for finite topology

A model for a set of axioms is a set of real-world objects that satisfy the axioms. Consider a Web site of six pages, linked together as follows:



In this model, an "open set" is defined by the property that no page in the set can be reached by a link from outside the set. We need to show that this definition is consistent with the axioms for open sets.
–The empty set is open. Since it contains no pages, it contains no page that can be reached by an outside link.
–The set X of all six pages is open, because there is no other page on the site from which an outside link could come.
–If sets A and B are open, no page in either can be reached by an outside link, and so their union is also open.
–If sets A and B are open, so is their intersection $A \cap B$. Proof by contraposition:

Suppose that $A \cap B$ is not open. Then it contains a page that can be reached by an outside link. If that link comes from A, then B is not open. If that link comes from B, then A is not open. If that link comes from outside both A and B, then both A and B are not open.

1.3 Topology in $\mathbb{R}$ and $\mathbb{R}^n$

The usual way to introduce a topology for the set $\mathbb{R}$ is to decree that any open interval is an open set and so is the empty set. Equivalently, we can decree that the set of points for which
$|x - x_0| < \epsilon$, with $\epsilon > 0$, is an open set. Notice that the infinite intersection of the open sets $(-1/n, 1/n)$ is the single point 0, a closed set!

The usual way to introduce a topology for the set $\mathbb{R}^n$ is to decree that any "open ball," the set of points for which $|\mathbf{x} - \mathbf{x_0}| < \epsilon$, with $\epsilon > 0$, is an open set.

1.4 More concepts of general topology

These definitions are intuitively reasonable for $\mathbb{R}$ and $\mathbb{R}^n$, but they also apply to the Web-site finite topology,

- Closed sets

  A closed set A is one whose complement $A^c = X - A$ is open. Careful: this is different from "one that is not open." There are lots of sets that are neither open nor closed, and there are sets that are both open and closed.

- A *neighborhood* of a point is any set that has as a subset an open set containing the point. A neighborhood does not have to be open.

- The *closure* of set $A \subset \mathbb{R}^n$, denoted $\overline{A}$, is "the smallest closed set that contains A," i.e. the intersection of all the closed sets that contain A

- The *interior* of a set $A \subset \mathbb{R}^n$, denoted $\mathring{A}$, is "the largest open set that is contained in A," i.e. the union of all the open subsets of A.

- The *boundary* of A, denoted ∂A, is the set of all points $\mathbf{x}$ with the property that any neighborhood of $\mathbf{x}$ includes points of A and also includes points of the complement A^c.
  The boundary of A is the difference between the closure of A and its interior.

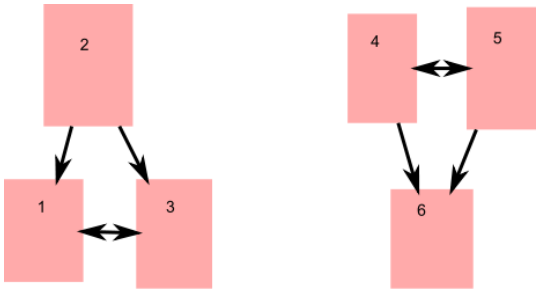
1.5 A topological definition of convergence

Sequence s_n converges to a limit s if for every open set A containing s, $\exists N$ such that $\forall n > N$, $a_n \in A$. In other words, the points of the sequence eventually get inside A and stay there.
Specialize to $\mathbb{R}$ and $\mathbb{R}^n$.

A sequence a_n of real numbers converges to a limit a if $\forall \epsilon > 0, \exists N$ such that $\forall n > N$, $|a - a_n| < \epsilon$. (open sets defined as open intervals)

A sequence $\mathbf{a_1}, \mathbf{a_2}, ...$ in $\mathbb{R}^n$ converges to the limit $\mathbf{a}$ if $\forall \epsilon > 0, \exists M$ such that if $m > M$, $|\mathbf{a_m} - \mathbf{a}| < \epsilon$. (open sets defined by open balls)
The sequence converges if and only if the sequences of coordinates all converge.

1.6 Something special about the open ball topology



For the Web diagram above, the sequence (6,5,4,6,5,4,5,4,5,4,...) converges both to 4 and to 5. Both {456} and {45} are open sets (no incoming links) but {4} {5}, {46}, and {56} are not.

This cannot happen in $\mathbb{R}^n$. If the sequence $\mathbf{a_1}, \mathbf{a_2}, ...$ in $\mathbb{R}^n$ converges to $\mathbf{a}$ and same sequence also converges to the limit $\mathbf{b}$, we can prove that $\mathbf{a} = \mathbf{b}$.

Why? The open ball topological space is *Hausdorff*. Given any two distinct points a and b, we can find open sets A and B with $a \in A$, $b \in B$, and $A \cap B = \emptyset$. In a Hausdorff space, the limit of a sequence is unique.

1.7 Infinite sequences and series of vectors and matrices

- We need something that can be made "less than ϵ." For vectors the familiar length is just fine. The "infinite triangle inequality" (proof 9.2) states that

$$|\sum_{i=1}^{\infty} \vec{\mathbf{a_i}}| \leq \sum_{i=1}^{\infty} |\vec{\mathbf{a_i}}|$$

- We define the "length of a matrix" by viewing the matrix as a vector.

  Since an $m \times n$ matrix A is an element of $\mathbb{R}^{mn}$, we can view it as a vector and define its length $|A|$ as the square root of the sum of the squares of all its entries. This definition has the following useful properties:

  - $|A\vec{\mathbf{b}}| \leq |A||\vec{\mathbf{b}}|$
  - $|AB| \leq |A||B|$

  Let A be a square matrix, and define its exponential by

$$\exp(At) = \sum_{r=0}^{\infty} \frac{(A)^r t^r}{r!}.$$

  Denoting the length of matrix A by $|A|$, we have

$$|\exp(At)| \leq \sum_{r=0}^{\infty} \frac{(|A|t)^r}{r!}.$$

  or $|\exp(At)| \leq \exp(|A|t) + \sqrt{n} - 1$, so the series is convergent for all t.

1.8 Calculating the exponential of a matrix

–If $D = \begin{bmatrix} b & 0 \\ 0 & c \end{bmatrix}$, then $Dt = \begin{bmatrix} bt & 0 \\ 0 & ct \end{bmatrix}$ and

$$\exp(Dt) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} bt & 0 \\ 0 & ct \end{bmatrix} + \frac{1}{2}\begin{bmatrix} (bt)^2 & 0 \\ 0 & (ct)^2 \end{bmatrix} + \cdots = \begin{bmatrix} e^{bt} & 0 \\ 0 & e^{ct} \end{bmatrix}$$

–If there is a basis of eigenvectors for A,
then $A = PDP^{-1}$, $A^r = PD^rP^{-1}$., and $\exp(At) = P\exp(Dt)P^{-1}$.
–Replace D by a conformal matrix $C = aI + bJ$ where $J^2 = -I$ and
$\exp(Ct) = \exp(aIt)\exp(bJt)$ can be expressed in terms of $\sin t$ and $\cos t$.
–If $A = bI + N$, and $N^2 = 0$, $\exp(At) = \exp bt \exp(Nt) = \exp bt(I + Nt)$.

1.9 Solving systems of linear differential equations

We put a dot over a quantity to denote its time derivative.
The solution to the differential equation $\dot{x} = kx$ is $x = \exp(kt)x_0$.
Suppose that there is more than one variable, for example
$\dot{x} = x + y$
$\dot{y} = -2x + 4y$.
If we set $\vec{\mathbf{v}} = \begin{bmatrix} x \\ y \end{bmatrix}$ then this pair of equations becomes
$\dot{\vec{\mathbf{v}}} = A\vec{\mathbf{v}}$, where $A = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}$
The solution is the same as in the single-variable case: $\vec{\mathbf{v}} = \exp(At)\vec{\mathbf{v}}_0$
Proof:

$$\exp At = \sum_{r=0}^{\infty} \frac{A^r t^r}{r!}.$$

$$\frac{d}{dt}\exp At = \sum_{r=1}^{\infty} \frac{rA^r t^{r-1}}{r!}.$$

Set $s = r - 1$.

$$\frac{d}{dt}\exp At = \sum_{s=0}^{\infty} \frac{A^{s+1}t^s}{s!} = A\sum_{s=0}^{\infty} \frac{A^s t^s}{s!} = A\exp At.$$

So

$$\dot{\vec{\mathbf{v}}} = \frac{d}{dt}\exp At\vec{\mathbf{v}}_0 = A\exp At\vec{\mathbf{v}}_0 = A\vec{\mathbf{v}}.$$

2 Lecture outline

1. Constructing a finite topology
   Axioms for general topology
   – The empty set and the set X are both open.
   – The union of any finite or infinite collection number of open sets is open.
   – The intersection of two open sets is open.

   Suppose that we start with $X = \{123456\}$ and choose a "subbasis," consisting of $\{123\}$, $\{245\}$, and $\{456\}$.

   - Find all the other sets that must be open because of the intersection axiom and the empty-set axiom.

   - Find all the other sets that must be open because of the union axiom and the axiom that set X is open.

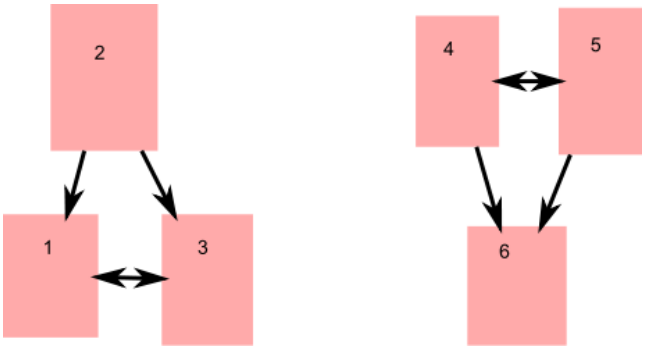
   - We now have the smallest collection of open sets that satisfies the axioms and includes the subbasis. A closed set is one whose complement is open. List all the closed sets.

   - What is the smallest legal collection of open sets in the general case?

   - What is the largest legal collection of open sets in the general case?

2. A Web-site model for finite topology

A model for a set of axioms is a set of real-world objects that satisfy the axioms. Consider a Web site X of six pages, linked together as follows:



In this model, an "open set" is defined by the property that no page in the set can be reached by a link from outside the set. We need to show that this definition is consistent with the axioms for open sets.
Use an "11-legged alligator" (if you cannot construct a counterexample, it is true) argument for the following:

The empty set is open.

The set X of all six pages is open.

Prove that if sets A and B are open, their union $A \cup B$ is also open.

Do this one by contraposition:
If sets A and B are open, so is their intersection $A \cap B$.

3. Topology in $\mathbb{R}$ and $\mathbb{R}^n$

   One way to introduce a topology for the set $\mathbb{R}$ is to decree that any open interval is an open set and so is the empty set.
   Prove that (a, b) is open if we decree that the set of points for which $|x - x_0| < \epsilon$, with $\epsilon > 0$, is an open set.

   Now the rule that only finite intersections of open sets have to be open becomes meaningful. Show that the infinite intersection of the open sets $(-1/n, 1/n)$ is not an open set!

   The usual way to introduce a topology for the set $\mathbb{R}^n$ is to decree that any "open ball," the set of points for which
   $|\mathbf{x} - \mathbf{x_0}| < \epsilon$, with $\epsilon > 0$, is an open set.

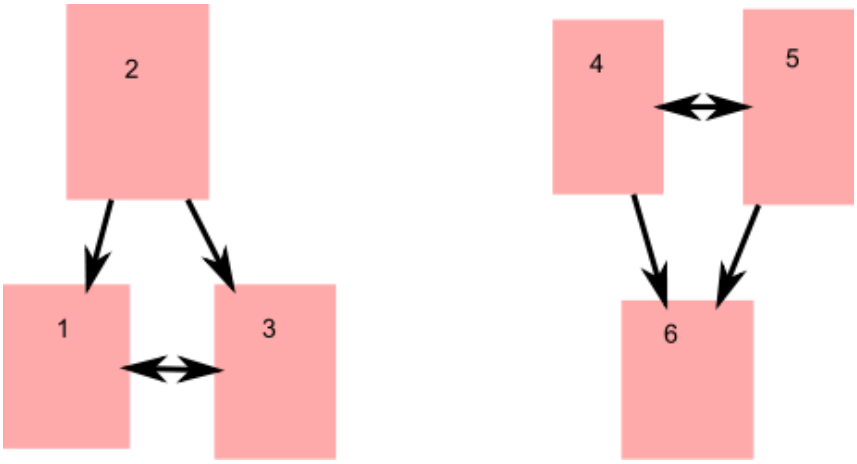
4. More concepts of general topology

   These definitions are intuitively reasonable for $\mathbb{R}$ and $\mathbb{R}^n$, but they also apply to the Web-site finite topology,

   - Closed sets

     A closed set A is one whose complement $A^c = X - A$ is open. Careful: this is different from "one that is not open." There are lots of sets that are neither open nor closed, and there are sets that are both open and closed.
   - A *neighborhood* of a point is any set that has as a subset an open set containing the point. A neighborhood does not have to be open.
   - The *closure* of set $A \subset \mathbb{R}^n$, denoted $\overline{A}$, is "the smallest closed set that contains A," i.e. the intersection of all the closed sets that contain A
   - The *interior* of a set $A \subset \mathbb{R}^n$, denoted $\mathring{A}$, is "the largest open set that is contained in A," i.e. the union of all the open subsets of A.
   - The *boundary* of A, denoted ∂A, is the set of all points $\mathbf{x}$ with the property that any neighborhood of $\mathbf{x}$ includes points of A and also includes points of the complement A^c.
     The boundary of A is the difference between the closure of A and its interior.

5. Applying these new definitions to the Web site topology



Open:$\{2\}, \{45\}, \{123\}, \{456\}, \{245\}, \{12345\}, \{2456\}$
Closed:$\{13456\}, \{1236\}, \{456\}, \{123\}, \{136\}, \{6\}, \{13\}$
Both: Empty set and $\{123456\}$

- Is $\{345\}$ a neighborhood of page 4?
- What is the closure of $\{23\}$?
- Of $\{26\}$?
- What is the interior of $\{23\}$?
- Of $\{23456\}$?
- What is the boundary of $\{23\}$?

6. The "open ball" definition of an open set satisfies the axioms of topology. A set $U \in \mathbb{R}^n$ is open if $\forall \mathbf{x} \in U, \exists r > 0$ such that the open ball $B_r(\mathbf{x}) \subset U$.

   - Prove that the empty set is open.

   - Prove that all of $\mathbb{R}^n$ is open.

   - Prove that the union of any collection of open sets is open.

   - Prove that the intersection of two open sets is open.

   - Prove that in $\mathbb{R}^2$, the boundary of the open disc $x^2 + y^2 < 1$ is the circle $x^2 + y^2 = 1$.

   - Find the infinite intersection of open balls of radius $\frac{1}{n}$ around the origin, for all positive integers. Is it open, closed, or neither?

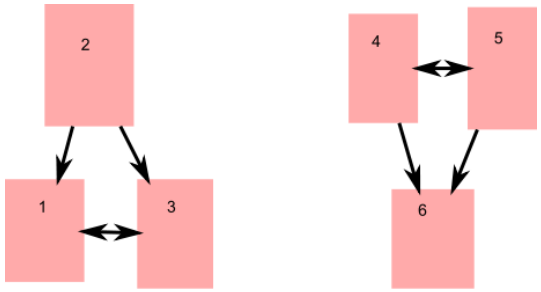
7. A topological definition of convergence

   Sequence s_n converges to a limit s if for every open set A containing s, $\exists N$ such that $\forall n > N$, $s_n \in A$. In other words, the points of the sequence eventually get inside A and stay there.
   Specialize to $\mathbb{R}$ and $\mathbb{R}^n$.

   A sequence s_n of real numbers converges to a limit s if $\forall \epsilon > 0, \exists N$ such that $\forall n > N$, $|s - s_n| < \epsilon$. (open sets defined as open intervals)

   A sequence $\mathbf{a_1}, \mathbf{a_2}, ...$ in $\mathbb{R}^n$ converges to the limit $\mathbf{a}$ if $\forall \epsilon > 0, \exists M$ such that if $m > M$, $|\mathbf{a_m} - \mathbf{a}| < \epsilon$. (open sets defined by open balls)
   We will prove that the sequence of points converges if and only if the sequences of coordinates all converge.

8. Something surprising about the open ball topology



For the Web diagram above , the sequence (6,5,4,6,5,4,5,4,5,4,...) converges both to 4 and to 5. Both {456} and {45} are open sets (no incoming links) but {4} {5}, {46}, and {56} are not.

This cannot happen in $\mathbb{R}^n$. If the sequence $\mathbf{a_1}, \mathbf{a_2}, ...$ in $\mathbb{R}^n$ converges to $\mathbf{a}$ and same sequence also converges to the limit $\mathbf{b}$,we can prove that $\mathbf{a} = \mathbf{b}$.

Why? The open ball topological space is *Hausdorff*. Given any two distinct points a and b, we can find open sets A and B with $a \in A$, $b \in B$, and $A \cap B = \emptyset$. In a Hausdorff space, the limit of a sequence is unique.

9. Proof 9.1

    - Define "Hausdorff space," and prove that in a Hausdorff space the limit of a sequence is unique.
    - Prove that $\mathbb{R}^n$, with the topology defined by open balls, is a Hausdorff space.

10. A closed subset contains all its limit points

    We defined a closed subset to be the complement of an open set.

    Using this definition, prove that if every element of the convergent sequence $(\mathbf{x}_n)$ is in the closed subset $C \subset \mathbb{R}^n$, then the limit x_0 of the sequence is also in C.

11. Convergent sequences in $\mathbb{R}^n$:

    A sequence $\mathbf{a_1}, \mathbf{a_2}, ...$ in $\mathbb{R}^n$ converges to the limit $\mathbf{a}$ if
    $\forall \epsilon > 0, \exists M$ such that if $m > M$, $|\mathbf{a_m} - \mathbf{a}| < \epsilon$.

    Prove that the sequence converges if and only if the sequences of coordinates all converge.

12. Proof 9.2
    Starting from the triangle inequality for two vectors, prove the triangle inequality for m vectors in $\mathbb{R}^n$, then prove the "infinite triangle inequality:"

$$|\sum_{i=1}^{\infty} \vec{\mathbf{a_i}}| \leq \sum_{i=1}^{\infty} |\vec{\mathbf{a_i}}|$$

    You may assume that the series $\sum_{i=1}^{\infty} \vec{\mathbf{a_i}}$ is "absolutely summable" (the infinite series of lengths on the right is convergent) but you must prove that this series is "summable" (infinite sum of vectors on the left is convergent.) As on page 100 of the textbook, you may use theorems 0.5.8 and 1.5.13.

13. Proof of inequalities involving matrix length

    The length of a matrix is calculated by treating it as a vector: take the square root of the sum of the squares of all the entries.

    Show that if matrix A consists of a single row, then $|A\vec{\mathbf{b}}| \leq |A||\vec{\mathbf{b}}|$ is just the Cauchy-Schwarz inequality.

    Prove the following:

    - $|A\vec{\mathbf{b}}| \leq |A||\vec{\mathbf{b}}|$ when A is an $m \times n$ matrix.
    - $|AB| \leq |A||B|$
    - $|I| = \sqrt{n}$ for the $n \times n$ identity matrix.

14. A geometric series of matrices

The geometric series formula for a square matrix A is

$$(I - A)^{-1} = I + A + A^2 + ...$$

Let $A = \begin{bmatrix} 0 & \frac{1}{2} \\ -\frac{1}{2} & 0 \end{bmatrix}$, $A^2 = \begin{bmatrix} -\frac{1}{4} & 0 \\ 0 & -\frac{1}{4} \end{bmatrix}$.

(a) Evaluate $I + A^2 + A^4.....$

(b) Evaluate $A + A^3 + A^5.... = A(I + A^2 + A^4....)$.

(c) Evaluate $I + A + A^2 + ...$

(d) Evaluate $(I - A)^{-1}$ and compare.

15. The exponential of a matrix

Let A be a square matrix, and define

$$\exp(At) = \sum_{r=0}^{\infty} \frac{(A)^r t^r}{r!}.$$

To show that this series converges, we use the infinite-sum version of the triangle inequality (proof 9.2), which applies to matrices if we treat them as vectors by using matrix length.

$$|\exp(At)| \leq \sum_{r=0}^{\infty} |\frac{(A)^r t^r}{r!}|.$$

Denoting the length of matrix A by $|A|$. we have

$$|\exp(At)| \leq \sum_{r=0}^{\infty} \frac{(|A|t)^r}{r!}.$$

or

$$|\exp(At)| \leq \exp(|A|t) + \sqrt{n} - 1.$$

It is easy to calculate the exponential of a diagonal matrix directly from this definition. If, for example, $D = \begin{bmatrix} b & 0 \\ 0 & c \end{bmatrix}$, then $Dt = \begin{bmatrix} bt & 0 \\ 0 & ct \end{bmatrix}$ and

$$\exp(Dt) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} bt & 0 \\ 0 & ct \end{bmatrix} + \frac{1}{2}\begin{bmatrix} (bt)^2 & 0 \\ 0 & (ct)^2 \end{bmatrix} + \cdots = \begin{bmatrix} e^{bt} & 0 \\ 0 & e^{ct} \end{bmatrix}$$

Now suppose that there is a basis of eigenvectors for A.

Then $A = PDP^{-1}$, where D is diagonal and P is the change of basis matrix whose columns are the eigenvectors.

Prove by induction that $A^r = PD^rP^{-1}$.

Prove that $\exp(At) = P\exp(Dt)P^{-1}$.

16. Calculating an exponential

We have already found that $A = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}$ has

eigenvector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ with eigenvalue 2 and

eigenvector $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ with eigenvalue 3.

Write A in the form $A = PDP^{-1}$.

Work out $\exp(At) = P \exp(Dt) P^{-1}$.

17. Solving systems of linear differential equations

We adopt the convention of putting a dot over a quantity to denote its time derivative.

The solution to the differential equation $\dot{x} = kx$ is $x = \exp(kt)x_0$, where x_0 can have any value.

Suppose that there is more than one variable, for example

$\dot{x} = x + y$

$\dot{y} = -2x + 4y$.

If we set $\vec{\mathbf{v}} = \begin{bmatrix} x \\ y \end{bmatrix}$ then this pair of equations becomes

$\dot{\vec{\mathbf{v}}} = A\vec{\mathbf{v}}$, where $A = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}$

The solution is the same as in the single-variable case:

$\vec{\mathbf{v}} = \exp(At)\vec{\mathbf{v}}_0$

Proof:

$$\exp At = \sum_{r=0}^{\infty} \frac{A^r t^r}{r!}.$$

$$\frac{d}{dt} \exp At = \sum_{r=1}^{\infty} \frac{r A^r t^{r-1}}{r!}.$$

Set $s = r - 1$.

$$\frac{d}{dt} \exp At = \sum_{s=0}^{\infty} \frac{A^{s+1} t^s}{s!} = A \sum_{s=0}^{\infty} \frac{A^s t^s}{s!} = A \exp At.$$

So

$$\dot{\vec{\mathbf{v}}} = \frac{d}{dt} \exp At\vec{\mathbf{v}}_0 = A \exp At\vec{\mathbf{v}}_0 = A\vec{\mathbf{v}}.$$

18. Checking the solution

The equation is

$\dot{\vec{\mathbf{v}}} = A\vec{\mathbf{v}}$, where $A = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}$

We earlier found that $A = PDP^{-1}$,

where $P = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$, $P^{-1} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$, $D = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$

Therefore $\exp At = P \begin{bmatrix} e^{2t} & 0 \\ 0 & e^{3t} \end{bmatrix} P^{-1}$

As "initial conditions," take $\vec{\mathbf{v}}_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

- Calculate $P^{-1}\vec{\mathbf{v}}_0$. This element of $\mathbb{R}^2$ expresses the initial conditions relative to the basis of eigenvectors.

- Calculate $\begin{bmatrix} e^{2t} & 0 \\ 0 & e^{3t} \end{bmatrix} P^{-1}\vec{\mathbf{v}}_0$. This element of $\mathbb{R}^2$ expresses the vector at time t, still relative to the basis of eigenvectors.

- Calculate $P \begin{bmatrix} e^{2t} & 0 \\ 0 & e^{3t} \end{bmatrix} P^{-1}\vec{\mathbf{v}}_0$. This element of $\mathbb{R}^2$ expresses the vector at time t, but now relative to the standard basis.

- Differentiate the answer with respect to t and check that
  $\dot{x} = x + y$
  $\dot{y} = -2x + 4y$.

19. Solving a differential equation when there is no eigenbasis.

The system of differential equations

$\dot{x} = 3x - y$

$\dot{y} = x + y$

can be written $\dot{\vec{\mathbf{v}}} = A\vec{\mathbf{v}}$, where $A = \begin{bmatrix} 3 & -1 \\ 1 & 1 \end{bmatrix}$.

Our standard technique leads to $p(t) = t^2 - 4t + 4 = (t-2)^2$, so there is one only eigenvalue.

Let $N = A - 2I = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$.

We have found that $p(A) = A^2 - 4A + 4I = (A - 2I)^2 = 0$, so $N^2 = 0$.

The addition formula for the exponential function,
$\exp(a + b) = \exp(a)\exp(b)$, which you proved on the homework by using Taylor series, is valid whenever $ab = ba$. It is therefore valid for commuting matrices. In particular, it is valid for the sum of a multiple of the identity matrix and any other matrix.

Since matrices $2I$ and N commute, $\exp(At) = \exp(2It)\exp(Nt)$

Show that $\exp At = e^{2t}(I + Nt)$,and confirm that $(\exp At)\vec{\mathbf{e}}_1$ is a solution to the differential equation.

20. Dealing with complex eigenvalues

Consider the 2×2 case where the eigenvalues of matrix A are complex. In this case we have learned how to express A in the form

$A = PCP^{-1}$, where C is the conformal matrix

$C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} + \begin{bmatrix} 0 & -b \\ b & 0 \end{bmatrix}$

Then $Ct = atI + btJ$, where $J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ and $J^2 = -I$.

Since the matrices atI and btJ commute, the addition formula for the exponential function holds, and

$\exp(Ct) = \exp(atI + btJ) = \exp(atI)\exp(btJ)$

Calculate $\exp(btJ)$ by substituting into the series for the exponential function and using $J^2 = -I$.

So $\exp(Ct) = \exp(atI)\exp(btJ) = \exp(atI)[(\cos bt)I + (\sin bt)J]$

and the solution to $\dot{\vec{\mathbf{v}}} = A\vec{\mathbf{v}}$

is $\vec{\mathbf{v}} = P\exp(Ct)P^{-1}\vec{\mathbf{v_0}}$.

21. Solving the "harmonic oscillator" differential equation

Applying Newton's second law of motion to a mass of 1 attached to a spring with "spring constant" 4 leads to the differential equation

$$\ddot{x} = -4x.$$

Solve this equation by using matrices for the case where $x(0) = 1, v(0) = 0$. The trick is to consider a vector

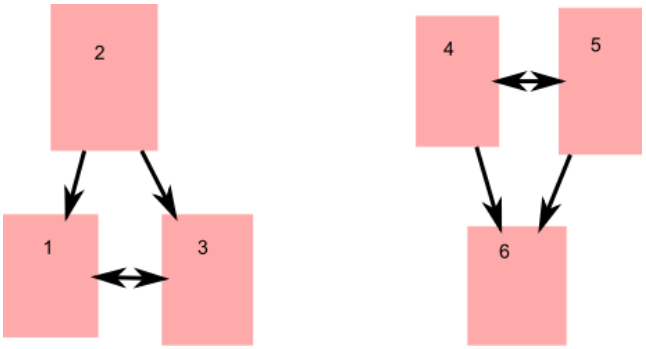
$$\vec{\mathbf{w}} = \begin{bmatrix} x(t) \\ v(t) \end{bmatrix}, \text{ where } v = \dot{x}.$$

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. State the three axioms for general topology, and show that they are satisfied in the topology of subsets of this set X of six linked Web pages:



   where an "open set" is a subset $S \subset X$ with the property that no page in S can be reached by a link from $X - S$.

   Define "interior," "closed set," and "closure," and show how you can apply these definitions to the set $A = \{1, 2\}$ to determine $\mathring{A}$ and $\overline{A}$.

   In the "open ball" topology for $\mathbb{R}$, any open interval is an open set and any closed interval is a closed set. Apply the definitions of interior and closure to find $\mathring{A}$ and $\overline{A}$ for the subset $A = (2, 3] \subset \mathbb{R}$.

2. (Proof 9.1)

   - Define "Hausdorff space," and prove that in a Hausdorff space the limit of a sequence is unique.
   - Prove that $\mathbb{R}^n$, with the topology defined by open balls, is a Hausdorff space.

3. A sequence $\mathbf{a_1}, \mathbf{a_2}, ...$ in $\mathbb{R}^n$ converges to the limit $\mathbf{a}$ if
$\forall \epsilon > 0, \exists M$ such that if $m > M$, $|\mathbf{a_m} - \mathbf{a}| < \epsilon$.

   Prove that the sequence converges if and only if, for all j, the sequence of jth coordinates $(a_m)_j$ converges to a_j.

4. (Proof 9.2) Starting from the triangle inequality for two vectors, prove the triangle inequality for m vectors in $\mathbb{R}^n$, then prove the "infinite triangle inequality:"

$$|\sum_{i=1}^{\infty} \vec{\mathbf{a}}_\mathbf{i}| \leq \sum_{i=1}^{\infty} |\vec{\mathbf{a}}_\mathbf{i}|$$

   You may assume that the series $\sum_{i=1}^{\infty} \vec{\mathbf{a}}_\mathbf{i}$ is "absolutely summable" (the infinite series of lengths on the right is convergent) but you must prove that this series is "summable" (infinite sum of vectors on the left is convergent.) As on page 100 of the textbook, you may use theorems 0.5.8 (if $\sum_{n=1}^{\infty} |a_n|$ converges ,then so does $\sum_{n=1}^{\infty} a_n$) and topic 4 (a sequence of vectors in $\mathbb{R}^n$ converges if and only if each component converges)

5. Let A be an $n \times n$ matrix, and let $\vec{\mathbf{v}}(t)$ be a vector in $\mathbb{R}^n$ whose components are functions of time. Define $\exp At$ as an infinite series, and prove that

$$\frac{d}{dt} \exp At = A \exp At.$$

   Thereby show that the vector $\vec{\mathbf{v}}(t) = (\exp At)\vec{\mathbf{v}}_0$ is the solution to the differential equation $\dot{\vec{\mathbf{v}}} = A\vec{\mathbf{v}}$ that satisfies the initial condition $\vec{\mathbf{v}}(0) = \vec{\mathbf{v}}_0$.

6. (Extra topic)
   Give a topological definition, one that uses only the concept of open set, for convergence of a sequence s_n. Show that, according to this definition, the sequence (6,5,4,6,5,4,5,4,5,4,...) in the Web site topology converges both to page 4 and to page 5. Then show, given that open intervals in $\mathbb{R}$ are open sets, that the sequence $(1, -\frac{1}{2}, \frac{1}{3}, -\frac{1}{4}, \frac{1}{5}, \cdots)$ converges to 0 according to your definition.

4 Workshop Problems

1. Topology

   (a) Properties of closed sets
   Recall the axioms of topology, which refer only to open sets:

   - The empty set and the set X are both open.
   - The union of any collection of open sets is open.
   - The intersection of two open sets is open.

   A closed set C is defined as a set whose complement C^c is open.

   You may use the following well-known properties of set complements, sometimes called "De Morgan's Laws":
   $(A \cup B)^c = A^c \cap B^c, (A \cap B)^c = A^c \cup B^c$.

   i. Prove directly from the axioms of topology that the union of two closed sets is closed.

   ii. In the Web site topology, a closed set of pages is one that has no outgoing links to other pages on the site. Prove that in this model, the union of two closed sets is closed.

   iii. Prove that if A and B are closed subsets of $\mathbb{R}^2$ (with the topology specified by open balls), their union is also closed.

   (b) Subsets of $\mathbb{R}$

   i. Let $A = \{0\} \cup (1, 2]$. Determine $A^c, \mathring{A}, \overline{A}$, and ∂A.

   ii. What interval is equal to $\bigcup_{n=2}^{\infty}[-1 + \frac{1}{n}, 1 - \frac{1}{n}]$? Is it a problem that this union of closed sets is not a closed set?

   iii. Let $\mathbb{Q}_1$ denote the set of rational numbers in the interval $(-1, 1)$. Determine the closure, interior, and boundary of this set.

2. Convergence in $\mathbb{R}^n$

   (a) Suppose that the sequence $\mathbf{a_1}, \mathbf{a_2}, ...$ in $\mathbb{R}^n$ converges to $\mathbf{0}$, and the sequence of real numbers $k_1, k_2, \cdots$, although not necessarily convergent, is bounded: $\exists K > 0$ such that $\forall n \in \mathbb{N}, |k_n| < K$.
   Prove that the sequence $k_1\mathbf{a_1}, k_2\mathbf{a_2}, ...$ in $\mathbb{R}^n$ converges to $\mathbf{0}$.

   (b) Prove that if $J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, then $\exp(Jt) = I\cos t + J\sin t$. Show that this is consistent with the Taylor series for e^{it}.

3. Differential equations

   (a) The original patriarchal differential equation problem
   Isaac has established large flocks of sheep for his sons Jacob and Esau. Anticipating sibling rivalry, he has arranged that the majority of the growth of each son's flock will come from lambs born to the other son. So, if $x(t)$ denotes the total weight of all of Jacob's sheep and $y(t)$ denotes the total weight of all of Esau's sheep, the time evolution of the weight of the flocks is given by the differential equations
   $\dot{x} = x + 2y$
   $\dot{y} = 2x + y$

      i. Calculate $\exp(At)$, where $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$.

      ii. Show that if the flocks are equal in size, they will remain that way. What has this got to do with the eigenvectors of A?

      iii. Suppose that when $t = 0$, the weight of Jacob's flock is S while the weight of Esau's flock is $2S$. Find formulas for the sizes as functions of time, and show that the flocks will become more nearly equal in weight as time passes.

   (b) Suppose that $\dot{\vec{\mathbf{v}}} = A\vec{\mathbf{v}}$, where $A = \begin{bmatrix} 3 & 1 \\ -1 & 1 \end{bmatrix}$. Since $p(t) = (t-2)^2$, there is no basis of eigenvectors. By writing A as the sum of a multiple of the identity matrix and a nilpotent matrix, calculate $\exp(At)$.

5 Homework

1. Suppose that you want to construct a Web site of six pages numbered 1 through 6, where the open sets of pages, defined as in lecture, include {126}, {124}, and {56}.

   (a) Prove that in the Web site model of finite topology, the intersection of two open sets is open.

   (b) What other sets must be open in order for the family of open sets to satisfy the intersection axiom?

   (c) What other sets must be open in order for the family of open sets to satisfy the union axiom?

   (d) List the smallest family of open sets that includes the three given sets and satisfies all three axioms. (You have already found all but one of these sets!)

   (e) Draw a diagram showing how six Web pages can be linked together so that only the sets in this family are open. This is tricky. First deal with 5 and 6. Then deal with 1 and 2. Then incorporate 4 into the network, and finally 3. There are many correct answers since, for example, if page 1 links to page 2 and page 2 links to page 3, then adding a direct link from page 1 to page 3 does not change the topology.

2. More theorems about limits of sequences

   The sequence $\vec{\mathbf{a_1}}, \vec{\mathbf{a_2}}, ...$ in $\mathbb{R}^n$ converges to $\vec{\mathbf{a}}$.

   The sequence $\vec{\mathbf{b_1}}, \vec{\mathbf{b_2}}, ...$ in $\mathbb{R}^n$ converges to $\vec{\mathbf{b}}$.

   (a) Prove that the sequence of lengths $|\vec{\mathbf{b_1}}|, |\vec{\mathbf{b_2}}|, ...$ in $\mathbb{R}$ is bounded: $\exists K$ such that $\forall n, |\vec{\mathbf{b_n}}| < K$. Hint: write $\vec{\mathbf{b_m}} = \vec{\mathbf{b_m}} - \vec{\mathbf{b}} + \vec{\mathbf{b}}$, then use the triangle inequality.

   (b) Define the sequence of dot products: $c_n = \vec{\mathbf{a_n}} \cdot \vec{\mathbf{b_n}}$.
   Prove that $c_1, c_2, \cdots$ converges to $\vec{\mathbf{a}} \cdot \vec{\mathbf{b}}$.
   Hint: Subtract and add $\vec{\mathbf{a}} \cdot \vec{\mathbf{b_n}}$, then use the triangle inequality and the Cauchy-Schwarz inequality.

3. Let $A = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix}$

   (a) By considering the length of A, show that

   $$\lim_{n\to\infty} A^n$$

   must be the zero matrix.

   (b) Find a formula for A^n when $n \geq 1$, and prove it by induction. Note that the formula is not valid for $n = 0$.

   (c) Verify the formula

   $$(I - A)^{-1} = I + A + A^2 + ...$$

   for this choice of A. As was the case in the example on page 20, you can evaluate the infinite sum on the right by summing a geometric series, but you should split off the first term and start the geometric series with the second term.

4. The differential equation $\ddot{x} = -3\dot{x} - 2x$ describes the motion of an "over-damped oscillator." The acceleration $\ddot{x}$ is the result of the sum of a force proportional to $\dot{x}$, supplied by a shock absorber, and a force proportional to x, supplied by a spring.

   (a) Introduce $v = \dot{x}$ as a new variable, and define the vector $\vec{\mathbf{w}} = \begin{bmatrix} x \\ v \end{bmatrix}$.

   Find a matrix A such that $\dot{\vec{\mathbf{w}}} = A\vec{\mathbf{w}}$.

   (b) Calculate the matrix $\exp(At)$.

   (c) Graph $x(t)$ for the following three sets of initial values that specify position and velocity when $t = 0$:

   Release from rest: $\vec{\mathbf{w_0}} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

   Quick shove: $\vec{\mathbf{w_0}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

   Push toward the origin: $\vec{\mathbf{w_0}} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$.

5. Suppose that A is a matrix of the form $S = \begin{bmatrix} a & b \\ b & a \end{bmatrix}$. Prove that

$$\exp(St) = \exp(at) \begin{bmatrix} \cosh(bt) & \sinh(bt) \\ \sinh(bt) & \cosh(bt) \end{bmatrix}.$$

Then use this result to solve

$\dot{x} = x + 2y$

$\dot{y} = 2x + y$

without having to diagonalize the matrix S.

6. Let $B = \begin{bmatrix} -1 & 9 \\ -1 & 5 \end{bmatrix}$. Show that there is only one eigenvalue λ and find an eigenvector for it. Then show that $N = B - \lambda I$ is nilpotent.

   (a) By writing $B = \lambda I + N$, calculate B^2.

   (b) By writing $B = \lambda I + N$, solve the system of equations

   $\dot{x} = -x + 9y$

   $\dot{y} = -x + 5y$

   for arbitrary initial conditions $\vec{\mathbf{v}}_0 = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$.

7. In Week 4, we wrote $A = \begin{bmatrix} 7 & -10 \\ 2 & -1 \end{bmatrix}$ in the form $A = PCP^{-1}$, where $C = \begin{bmatrix} 3 & -2 \\ 2 & 3 \end{bmatrix}$ is conformal and $P \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$

   Follow up on this analysis to solve the differential equation $\dot{\vec{\mathbf{v}}} = A\vec{\mathbf{v}}$ for initial conditions $\vec{\mathbf{v}}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

8. Let A be a 2×2 matrix which has two distinct real eigenvalues λ_1 and λ_2, with associated eigenvectors $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$.

   (a) Show that the matrix $P_1 = \frac{A - \lambda_2 I}{\lambda_1 - \lambda_2}$ is a projection onto the subspace spanned by eigenvector $\vec{\mathbf{v}}_1$. Find its image and kernel, and show that $P_1^2 = P_1$.

   (b) Similarly, the matrix $P_2 = \frac{A - \lambda_1 I}{\lambda_2 - \lambda_1}$ is a projection onto the subspace spanned by eigenvector $\vec{\mathbf{v}}_2$. Show that $P_1 P_2 = P_2 P_1 = 0$, that $P_1 + P_2 = I$, and that $\lambda_1 P_1 + \lambda_2 P_2 = A$.

   (c) Show that $\exp(t\lambda_1 P_1 + t\lambda_2 P_2) = \exp(\lambda_1 t) P_1 + \exp(\lambda_2 t) P_2$, and use this result to solve the equations

   $\dot{x} = -4x + 5y$

   $\dot{y} = -2x + 3y$

   for arbitrary initial conditions $\vec{\mathbf{v}}_0 = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$.

1. Road map of the module
2. Topology (in $\mathbb{R}$) true/false
   (a) A set S is closed iff it is not open
   (b) A finite set is always closed
   (c) All intersections of open sets are open
   (d) An open set that contains every rational number must contain all of $\mathbb{R}$
3. Prove that the intersection of two open sets in $\mathbb{R}^n$ is open (using the definitions of open sets in $\mathbb{R}^n$, not just general facts about open sets)
4. Prove that a singleton set (set of one element) in $\mathbb{R}^n$ is closed
5. Using integration, solve the differential equation (of one variable) $\dot{x} = kx$. How does this relate to the technique we used this week for solving (some) differential equations of multiple variables?
6. Using the Taylor-series definition, prove that if D is a diagonal matrix, then the exponentiation rule we learned back in Module 1 works.
7. If time: Suppose that $\dot{\vec{v}} = A\vec{v}$, where $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$, and that $v_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Given this, solve for $\vec{v}$. Notice that there is only one eigenvalue, so there is no basis of eigenvectors.

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)

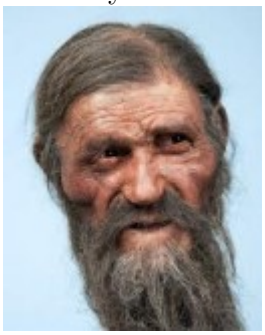
Proofs to present in section or to a classmate who has done them.

- 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 $X_k = (0, \frac{1}{k})$ (not 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

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

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

1. Road map of the module
2. Heine-Borel hip hip hooray
   (a) What is an open cover of a set S? What is a finite subcover?
   (b) Rather than hire a normal company to do your cake decoration, you decide to hire Heine-Borel Cake Decorators to frost your one-dimensional birthday cake, which extends across the interval [0,1]. HBCD offers you a few different cake-decorating plans, but you're worried that they might not finish frosting the cake in time.
      i. Plan 1: During hour -1, your cake will be frosted on the interval $\left(\frac{1}{3}, 1\right]$. During hour 0, your cake will be frosted on $[0, 0.00001)$. During hour i after that, your cake will be frosted on the open interval $\left(\frac{1}{2^{i+1}}, \frac{1}{2^i}\right)$. Does this cake-frosting scheme form an open cover of your birthday cake? (Will everything get frosted?)
      ii. Plan 2: During hour -1, your cake will be frosted on the interval $\left(\frac{1}{3}, 1\right]$. During hour 0, your cake will be frosted on $[0, 0.00001)$. During hour i after that, your cake will be frosted on the open interval $\left(\frac{1}{2^{i+2}}, \frac{1}{2^i}\right)$. The way that this plan is set up, you'd have to wait infinitely long for the frosting to finish. Why don't you have to wait that long?
      iii. HBCD convinces you that your cake would look better if they only frosted on (0,1). How could they change plan 2 in this case to make you wait an infinite length of time before your cake is ready?
3. Affine approximation: Consider the function $f\begin{pmatrix} x \\ y \end{pmatrix} = x^4y^2$. Calculate the derivative of this function, evaluate it at and $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$, and then use this to approximate $f\begin{pmatrix} 0.95 \\ 2.05 \end{pmatrix}$.
4. True/false
   (a) For any collection of sets S_n where $S_{n+1} \subseteq S_n$, $\cap_{i=1}^{\infty} S_n$ is nonempty
   (b) In $\mathbb{R}^n$, any convergent sequence is Cauchy.
   (c) In $\mathbb{R}^n$, any Cauchy sequence is convergent.
   (d) If a set $S \subseteq \mathbb{R}^n$ is not compact, then no open cover has a finite subcover.
   (e) In $\mathbb{R}^2$, if both $\lim_{x\to 0} f\begin{pmatrix} x \\ 0 \end{pmatrix}$ and $\lim_{y\to 0} f\begin{pmatrix} 0 \\ y \end{pmatrix}$ exist, then $\lim_{\vec{x}\to 0} f(\vec{x})$ exists.
   (f) If $\nabla_{\vec{e_1}} f(\vec{a}) = 1$ and $\nabla_{\vec{e_2}} f(\vec{a}) = 2$, then $\nabla_{\vec{e_1}+\vec{e_2}} f(\vec{a}) = 3$.
5. If time: Using the pigeonhole principle, prove that given any five points on a sphere, there is a closed hemisphere containing at least four of them.

MATHEMATICS 23a/E-23a, Fall 2018
Linear Algebra and Real Analysis I

Week 11 (Differentiability, Newton's method, inverse and implicit functions, manifolds)

Author: Paul Bamberg
R scripts by Paul Bamberg
Last modified: November 28, 2018 by Paul Bamberg (fixed reference in workshop problem 2b)

.

Reading

- Hubbard, section 1.7 (you have already read most of this)
- Hubbard, sections 1.8 and 1.9 (computing derivatives and differentaibility)
- Hubbard, section 2.8 page 233-235 and page 246. (Newton's method)
- Hubbard, section 2.10 up through page 264. (inverse function theorem)
- Hubbard, Section 3.1 (Implicit functions and manifolds)

Recorded Lectures

- Lecture 22 (Week 11, Class 1) (watch on November 20 or 21)
- Lecture 23 (Week 11, Class 2) (watch on November 27 or 28)
- Lecture 24 (Fortnight 12, Class 1) (watch on November 29 or 30)

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

- 11.1 Let $U \subset \mathbb{R}^n$ be an open set, and let f and g be functions from U to $\mathbb{R}$. Prove that if f and g are differentiable at $\mathbf{a}$ then so is fg, and that

$$[\mathbf{D}(fg)(\mathbf{a})] = f(\mathbf{a})[\mathbf{D}g(\mathbf{a})] + g(\mathbf{a})[\mathbf{D}f(\mathbf{a})].$$

- 11.2 Using the mean value theorem, prove that if a function $f : \mathbb{R}^2 \to \mathbb{R}$ has partial derivatives $D_1 f$ and $D_2 f$ that are continuous at $\mathbf{a}$, it is differentiable at $\mathbf{a}$ and its derivative is the Jacobian matrix $\begin{bmatrix} D_1 f(\mathbf{a}) & D_2 f(\mathbf{a}) \end{bmatrix}$.

- 11.3 The implicit function theorem

  Let W be an open subset of $\mathbb{R}^n$, and let $\mathbf{F} : W \to \mathbb{R}^{n-k}$ be a C^1 mapping such that $\mathbf{F}(\mathbf{c}) = \mathbf{0}$. Assume that $[\mathbf{DF}(\mathbf{c})]$ is onto.

  Prove that the n variables can be ordered so that the first $n-k$ columns of $[\mathbf{DF}(\mathbf{c})]$ are linearly independent, and that $[\mathbf{DF}(\mathbf{c})] = [A|B]$ where A is an invertible $(n-k) \times (n-k)$ matrix.

  Set $\mathbf{c} = \begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix}$, where $\mathbf{a}$ are the $n-k$ passive variables and $\mathbf{b}$ are the k active variables.

  Let $\mathbf{g}$ be the "implicit function" from a neighborhood of $\mathbf{b}$ to a neighborhood of $\mathbf{a}$ such that $\mathbf{g}(\mathbf{b}) = \mathbf{a}$ and $\mathbf{F}\begin{pmatrix} \mathbf{g}(\mathbf{y}) \\ \mathbf{y} \end{pmatrix} = 0$.

  Prove that $[\mathbf{Dg}(\mathbf{b})] = -A^{-1}B$.

R Scripts

- Script 3.3A-ComputingDerivatives.R
  Topic 1 - Testing for differentiability
  Topic 2 - Illustrating the derivative rules

- Script 3.3B-NewtonsMethod.R
  Topic 1 - Single variable
  Topic 2 - 2 equations, 2 unknowns
  Topic 3 - Three equations in three unknowns

- Script 3.3C-InverseFunction.R
  Topic 1 - A parametrization function and its inverse
  Topic 2 - Visualizing coordinates by means of a contour plot
  Topic 3 - An example that is economic, not geometric

1 Executive Summary

1.1 Definition of the derivative

- Converting the derivative to a matrix
  The linear function $f(h) = mh$ is represented by the 1×1 matrix $[m]$.
  When we say that $f'(a) = m$, what we mean is that the function
  $f(a+h) - f(a)$ is well approximated, for small h, by the linear function mh. The error made by using the approximation is a "remainder" $r(h) = f(a+h) - f(a) - mh$. If f is differentiable, this remainder approaches 0 faster than h, i.e.

$$\lim_{h \to 0} \frac{r(h)}{h} = \lim_{h \to 0} \frac{f(a+h) - f(a) - mh}{h} = 0.$$

  This definition leads to the standard rule for calculating the number m,

$$m = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}.$$

- Extending this definition to $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$
  A linear function $L(\vec{\mathbf{h}})$ is represented by an $m \times n$ matrix.
  When we say that $\mathbf{f}$ is differentiable at $\mathbf{a}$, we mean that the function
  $\mathbf{f}(\mathbf{a} + \vec{\mathbf{h}}) - \mathbf{f}(\mathbf{a})$ is well approximated, for any $\vec{\mathbf{h}}$ whose length is small, by a linear function L, called the derivative $[\mathbf{Df}(\mathbf{a})]$.
  The error made by using the approximation is a "remainder"
  $\mathbf{r}(\vec{\mathbf{h}}) = \mathbf{f}(\mathbf{a} + \vec{\mathbf{h}}) - \mathbf{f}(\mathbf{a}) - [\mathbf{Df}(\mathbf{a})](\vec{\mathbf{h}})$.
  $\mathbf{f}$ is called differentiable if this remainder approaches 0 faster than $|\vec{\mathbf{h}}|$, i.e.

$$\lim_{\vec{\mathbf{h}} \to \vec{\mathbf{0}}} \frac{1}{|\vec{\mathbf{h}}|} \mathbf{r}(\vec{\mathbf{h}}) = \lim_{\vec{\mathbf{h}} \to \vec{\mathbf{0}}} \frac{1}{|\vec{\mathbf{h}}|} (\mathbf{f}(\mathbf{a} + \vec{\mathbf{h}}) - \mathbf{f}(\mathbf{a}) - [\mathbf{Df}(\mathbf{a})](\vec{\mathbf{h}})) = \mathbf{0}.$$

  In that case, $[\mathbf{Df}(\mathbf{a})]$ is represented by the Jacobian matrix $[\mathbf{Jf}(\mathbf{a})]$.

  Proof: Since L exists and is linear, it is sufficient to consider its action on each standard basis vector. We choose $\vec{\mathbf{h}} = t\vec{\mathbf{e_i}}$ so that $|\vec{\mathbf{h}}| = t$. Knowing that the limit exists, we can use any sequence that converges to the origin to evaluate it, and so

$$\lim_{t \to 0} \frac{1}{t} (\mathbf{f}(\mathbf{a} + t\vec{\mathbf{e_i}}) - \mathbf{f}(\mathbf{a}) - tL\vec{\mathbf{e_i}})) = 0? \text{ and } L(\vec{\mathbf{e_i}}) = \lim_{t \to 0} \frac{1}{t} (\mathbf{f}(\mathbf{a} + t\vec{\mathbf{e_i}}) - \mathbf{f}(\mathbf{a}))$$

  What is hard is proving that f is differentiable – that L exists – since that requires evaluating a limit where $\vec{\mathbf{h}} \to \vec{\mathbf{0}}$. Eventually we will prove that f is differentiable at $\mathbf{a}$ if all its partial derivatives are continuous there.

1.2 Proving differentiability and calculating derivatives

In every case $\mathbf{f}$ is a function from U to $\mathbb{R}^m$, where U is an open subset of $\mathbb{R}^n$.

- $\mathbf{f}$ is constant: $\mathbf{f} = \mathbf{c}$. Then $[\mathbf{Df}(\mathbf{a})]$ is the zero linear transformation, since

$$\lim_{\vec{\mathbf{h}}\to\vec{\mathbf{0}}} \frac{1}{|\vec{\mathbf{h}}|}(\mathbf{f}(\mathbf{a}+\vec{\mathbf{h}}) - \mathbf{f}(\mathbf{a}) - [\mathbf{Df}(\mathbf{a})]\vec{\mathbf{h}}) = \lim_{\vec{\mathbf{h}}\to\vec{\mathbf{0}}} \frac{1}{|\vec{\mathbf{h}}|}(\mathbf{c} - \mathbf{c} - \vec{\mathbf{0}}) = \vec{\mathbf{0}}.$$

- $\mathbf{f}$ is affine: a constant plus a linear function, $\mathbf{f} = \mathbf{c} + L$. $[\mathbf{Df}(\mathbf{a})] = L$, since

$$\lim_{\vec{\mathbf{h}}\to\vec{\mathbf{0}}} \frac{1}{|\vec{\mathbf{h}}|}(\mathbf{f}(\mathbf{a}+\vec{\mathbf{h}})-\mathbf{f}(\mathbf{a})-[\mathbf{Df}(\mathbf{a})]\vec{\mathbf{h}}) = \lim_{\vec{\mathbf{h}}\to\vec{\mathbf{0}}} \frac{1}{|\vec{\mathbf{h}}|}(\mathbf{c}+L(\mathbf{a}+\vec{\mathbf{h}})-(\mathbf{c}+L(\mathbf{a}))-L(\vec{\mathbf{h}})) = 0.$$

$\mathbf{f}$ has differentiable components: if $\mathbf{f} = \begin{pmatrix} f_1 \\ \cdot \\ \cdot \\ \cdot \\ f_n \end{pmatrix}$: then $\mathbf{Df}(\mathbf{a}) = \begin{bmatrix} \mathbf{D}f_1(\mathbf{a}) \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{D}f_n(\mathbf{a}) \end{bmatrix}$

.

- $\mathbf{f} + \mathbf{g}$ is the sum of two functions $\mathbf{f}$ and $\mathbf{g}$, both differentiable at $\mathbf{a}$.
  The derivative of $\mathbf{f} + \mathbf{g}$ is the sum of the derivatives of $\mathbf{f}$ and $\mathbf{g}$. (easy to prove)

- $f\mathbf{g}$ is the product of scalar-valued function f and vector-valued $\mathbf{g}$, both differentiable. Then
  $[\mathbf{D}(f\mathbf{g})(\mathbf{a})]\vec{\mathbf{v}} = f(\mathbf{a})([\mathbf{Dg}(\mathbf{a})]\vec{\mathbf{v}}) + ([\mathbf{D}f(\mathbf{a})]\vec{\mathbf{v}})\mathbf{g}(\mathbf{a})$.

- $\mathbf{g}/f$ is the quotient of vector-valued function $\mathbf{g}$ and scalar-valued f, both differentiable, and $f(\mathbf{a}) \neq 0$. Then

$$[\mathbf{D}(\frac{\mathbf{g}}{f})(\mathbf{a})]\vec{\mathbf{v}} = \frac{[\mathbf{Dg}(\mathbf{a})]\vec{\mathbf{v}}}{f(\mathbf{a})} - \frac{([\mathbf{D}f(\mathbf{a})]\vec{\mathbf{v}})\mathbf{g}(\mathbf{a})}{(f(\mathbf{a})^2}.$$

- $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ are open sets, and $\mathbf{a}$ is a point in U at which we want to evaluate a derivative.

  $\mathbf{g} : U \to V$ is differentiable at $\mathbf{a}$, and $[\mathbf{Dg}(\mathbf{a})]$ is an $m \times n$ Jacobian matrix.

  $\mathbf{f} : V \to \mathbb{R}^p$ is differentiable at $\mathbf{g}(\mathbf{a})$, and $[\mathbf{Df}(\mathbf{g}(\mathbf{a}))]$ is a $p \times m$ Jacobian matrix.

  The chain rule states that $[\mathbf{D}(\mathbf{f} \circ \mathbf{g})(\mathbf{a}))] = [\mathbf{Df}(\mathbf{g}(\mathbf{a}))] \circ [\mathbf{Dg}(\mathbf{a})]$.

- The combined effect of all these rules is effectively that if a function is defined by well-behaved formulas (no division by zero), it if differentiable, and its derivative is represented by its Jacobian matrix.

1.3 Connection between Jacobian matrix and derivative

- If $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ is defined on an open set $U \in \mathbb{R}^n$, and

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}\begin{pmatrix} x_1 \\ ... \\ x_n \end{pmatrix} = \begin{pmatrix} f_1(\mathbf{x}) \\ ... \\ f_m(\mathbf{x}) \end{pmatrix}$$

  the Jacobian matrix $[\mathbf{Jf}(\mathbf{x})]$ is made up of all the partial derivatives of $\mathbf{f}$:

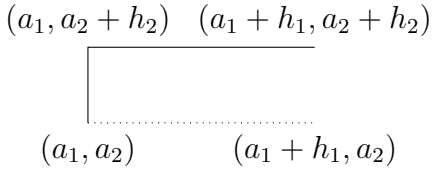
$$[\mathbf{Jf}(\mathbf{a})] = \begin{bmatrix} D_1 f_1(\mathbf{a}) .... D_n f_1(\mathbf{a}) \\ ... \\ D_1 f_m(\mathbf{a}) .... D_n f_m(\mathbf{a}) \end{bmatrix}$$

- We can invent pathological cases where the Jacobian matrix of f exists (because all the partial derivatives exist), but the function f is not differentiable. In such a case, using the formula

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

  generally gives the wrong answer for the directional derivative! You are trying to use a linear approximation where none exists.

- Using the Jacobian matrix of partial derivatives to get a good affine approximation for $f(\mathbf{a}+\vec{\mathbf{h}})$ is tantamount to assuming that you can reach the point $\mathbf{a}+\vec{\mathbf{h}}$ by moving along lines that are parallel to the coordinate axes and that the change in the function value along the solid horizontal line is well approximated by the change along the dotted horizontal line. With the aid of the mean value theorem, you can show that this is the case if (proof 11.2) the partial derivatives of f at $\mathbf{a}$ are continuous.



1.4 Newton's method – one variable

Newton's method is based on the tangent-line approximation. Function f is differentiable. We are trying to solve the equation $f(x) = 0$, and we have found a value a_0 that is close to the desired x. So we use the best affine approximation $f(x) \approx f(a_0) + f'(x_0)(x - a_0)$.
Then we find a value a_1 for which this tangent-line approximation equals zero. $f(a_0) + f'(x_0)(a_1 - a_0) = 0$, and $a_1 = a_0 - f(a_0)/f'(a_0)$.
When $f(a_0)$ is small, $f'(a_0)$ is large, and $f'(a_0)$ does not change too rapidly, a_1 is a much improved approximation to the desired solution x. Details, for which Kantorovich won the Nobel prize in economics, are in Hubbard.

1.5 Newton's method – more than one variable

Example: we are trying to solve a system of n nonlinear equations in n unknowns, e.g.
$x^2e^y - \sin(y) - 0.3 = 0$
$\tan x + x^2y^2 - 1 = 0.$
Ordinary algebra is no help – there is no nonlinear counterpart to row reduction.
U is an open subset of $\mathbb{R}^n$, and we have a differentiable function $\vec{\mathbf{f}}(\mathbf{x}) : U \to \mathbb{R}^n$.
In the example, $\vec{\mathbf{f}}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x^2e^y - \sin(y) - 0.3 \\ \tan x + x^2y^2 - 1 \end{pmatrix}$, which is differentiable.
We are trying to solve the equation $\vec{\mathbf{f}}(\mathbf{x}) = \vec{\mathbf{0}}$.
Suppose we have found a value $\mathbf{a_0}$ that is close to the desired $\mathbf{x}$.
Again we use the best affine approximation

$$\vec{\mathbf{f}}(\mathbf{x}) \approx \vec{\mathbf{f}}(\mathbf{a_0}) + [\mathbf{D}\tilde{\mathbf{f}}(\mathbf{a_0})](\mathbf{x} - \mathbf{a_0}).$$

We set out to find a value $\mathbf{a_1}$ for which this affine approximation equals zero.

$$\vec{\mathbf{f}}(\mathbf{a_0}) + [\mathbf{D}\tilde{\mathbf{f}}(\mathbf{a_0})](\mathbf{a_1} - \mathbf{a_0}) = \vec{\mathbf{0}}$$

This is a linear equation, which we know how to solve!
If $[\mathbf{D}\tilde{\mathbf{f}}(\mathbf{a_0})]$ is invertible (and if it is not, we look for a better $\mathbf{a_0}$), then

$$\mathbf{a}_1 = \mathbf{a}_0 - [\mathbf{D}\tilde{\mathbf{f}}(\mathbf{a_0})]^{-1}\vec{\mathbf{f}}(\mathbf{a}_0).$$

Iterating this procedure is the best known for solving systems of nonlinear equations. Hubbard has a detailed discussion (which you are free to ignore) of how to use Kantorovich's theorem to assess convergence.

1.6 The inverse function theorem – short version

For function $f : [a, b] \to [c, d]$, we know that if f is strictly increasing or strictly decreasing on interval $[a, b]$, there is an inverse function g for which $g \circ f$ and $f \circ g$ are both the identity function. We can find $g(y)$ for a specific y by solving $f(x) - y = 0$, perhaps by Newton's method. If $f(x_0) = y_0$ and $f'(x_0) \neq 0$, we can prove that g is differentiable at y_0 and that $g'(y_0) = 1/f'(x_0)$.
"Strictly monotone" does not generalize, but "nonzero $f'(x_0)$" generalizes to "invertible $[\mathbf{Df}(\mathbf{x}_0)]$." Start with a function $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ whose partial derivatives are all continuous, so that we know that it is differentiable everywhere. Choose a point $\mathbf{x}_0$ where the derivative $[\mathbf{Df}(\mathbf{x}_0)]$ is an invertible matrix. Set $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$. Then there is a differentiable local inverse function $\mathbf{g} = \mathbf{f}^{-1}$ such that

- $\mathbf{g}(\mathbf{y}_0) = \mathbf{x}_0$.
- $f(\mathbf{g}(\mathbf{y})) = \mathbf{y}$ if $\mathbf{y}$ is close enough to $\mathbf{y}_0$.
- $[\mathbf{Dg}(\mathbf{y})] = [\mathbf{Df}(\mathbf{g}(\mathbf{y}))]^{-1}$ (follows from the chain rule)

1.7 Implicit functions – review of the linear case.

We have n unknowns, $n-k$ equations, e.g for $n=3, k=1$
$2x+3y-z=0$, $4x-2y+3z=0$
Create an $(n-k)\times n$ matrix: $T=\begin{bmatrix}2 & 3 & -1\\ 4 & -2 & 3\end{bmatrix}$
If the matrix T is not onto, its rows (the equations) are linearly dependent. Otherwise, when we row reduce, we will find $n-k=2$ pivotal columns and $k=1$ nonpivotal columns. We assign values arbitrarily to the "active" variables that correspond to the nonpivotal columns, and then the values of the "passive" variables that corresponds to the pivotal column are determined.
Suppose that we reorder the unknowns so that the "active" variables come last. Then, after we row reduce the matrix, the first $n-k$ columns will be pivotal. So the first $n-k$ columns will be linearly independent, and they form an invertible square matrix. The matrix is now of the form $T=[A|B]$, where A is invertible.
The solution vector is of the form $\vec{\mathbf{v}}=\begin{bmatrix}\vec{\mathbf{x}}\\ \vec{\mathbf{y}}\end{bmatrix}$, where the passive variables $\vec{\mathbf{x}}$ come first, the active variables $\vec{\mathbf{y}}$ come second.
A solution to $T\vec{\mathbf{v}}=\vec{\mathbf{0}}$ is obtained by choosing $\vec{\mathbf{y}}$ arbitrarily and setting $\vec{\mathbf{x}}=-A^{-1}B\vec{\mathbf{y}}$. Our system of equations determines $\vec{\mathbf{x}}$ "implicitly" in terms of $\vec{\mathbf{y}}$.

1.8 Implicit function theorem – the nonlinear case.

We have a point $\mathbf{c}\in\mathbb{R}^n$, a neighborhood W of $\mathbf{c}$, and a function $\mathbf{F}: W\to\mathbb{R}^{n-k}$ for which $\mathbf{F}(\mathbf{c})=0$ and $[\mathbf{DF}(\mathbf{c})]$ is onto. $\mathbf{F}$ imposes *constraints*.
The variables are ordered so that the $n-k$ pivotal columns in the Jacobian matrix, which correspond to the passive variables, come first. Let $\mathbf{a}$ denote the passive variables at $\mathbf{c}$; let $\mathbf{b}$ denote the active variables at $\mathbf{c}$.

The implicit function $\mathbf{g}$ expresses the passive variables in terms of the active variables, and $\mathbf{g}(\mathbf{b})=\mathbf{a}$. For $\mathbf{y}$ near $\mathbf{b}$, $\mathbf{x}=\mathbf{g}(\mathbf{y})$ determines passive variables such that $\mathbf{F}\begin{pmatrix}\mathbf{a}\\ \mathbf{b}\end{pmatrix}=\mathbf{0}$. Tweak $\mathbf{y}$, and $\mathbf{g}$ specifies how to tweak $\mathbf{x}$ so that the constraints are still satisfied.

Although we usually cannot find a formula for $\mathbf{g}$, we can find its derivative at $\mathbf{b}$ by the same recipe that worked in simple cases.

Evaluate the Jacobian matrix $[\mathbf{DF}(\mathbf{c})]$.

Extract the first $n-k$ columns to get an invertible square matrix A.

Let the inverse of this matrix act on the remaining k columns (matrix B) and change the sign to get the $(n-k)\times k$ Jacobian matrix for $\mathbf{g}$.

That is, $[\mathbf{Dg}(\mathbf{b})]=-A^{-1}B$.

1.9 Curves, Surfaces, Graphs, and Manifolds

Manifolds are a generalization of smooth curves and surfaces.

The simplest sort of manifold is a flat one, described by linear equations. An example is the line of slope 2 that passes through the point $x = 0, y = -2$: a one-dimensional submanifold of $\mathbb{R}^2$
There are three equivalent ways to describe such a manifold.

- (The definition) As the *graph* of a function that expresses the passive variables in terms of the active variables: either $y = f(x) = -2 + 2x$ or $x = g(y) = \frac{1}{2}(y + 2)$.
- As a "locus" defined by a constraint equation $F\begin{pmatrix} x \\ y \end{pmatrix} = 2x - y - 2 = 0$.
- By a parametrization function $g(t) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + t\begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

Definition: A subset $M \subset \mathbb{R}^n$ is a smooth manifold if locally it is the graph of a C^1 function (the partial derivatives are continuous). "Locally" means that for any point $\mathbf{x} \in M$ we can find a neighborhood U of $\mathbf{x}$ such that within $M \cap U$, there is a C^1 function that expresses $n - k$ passive variables in terms of the other k active variables. The number k is the dimension of the manifold. In $\mathbb{R}^3$ there are four possibilities:

- $k = 3$. Any open subset $M \subset \mathbb{R}^3$ is a smooth 3-dimensional manifold. In this case $k = 3$, and the manifold is the graph of a function $f : \mathbb{R}^3 \to \{\vec{\mathbf{0}}\}$, whose codomain is the trivial vector space $\{\vec{\mathbf{0}}\}$ that contains just a single point. Such a function is necessarily constant, and its derivative is zero.
- $k = 2$. The graph of $z = f\begin{pmatrix} x \\ y \end{pmatrix} = x^2 + y^2$ is a paraboloid.
- $k = 1$. The graph of the function $\begin{pmatrix} x \\ y \end{pmatrix} = \vec{\mathbf{f}}(z) = \begin{pmatrix} \cos 2\pi z \\ \sin 2\pi z \end{pmatrix}$ is a helix.
- $k = 0$. In this case the manifold consists of one of more isolated points. Near any of these points $\mathbf{x}_0$, it is the graph of a function $\vec{\mathbf{f}} : \{\vec{\mathbf{0}}\} \to \mathbb{R}^3$ whose domain is a zero-dimensional vector space and whose image is the point $\mathbf{x}_0 \in \mathbb{R}^3$.

There is no requirement that a manifold be the graph of a single function, or that the "active" variables be the same at every point on the manifold. The unit circle, the locus of $x^2 + y^2 - 1 = 0$, is the union of four function graphs, two of which have x as the active variable, two of which have y. By using a parameter t that is not one of the variables, we can represent it by the parametrization $\begin{pmatrix} x \\ y \end{pmatrix} = g(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$

2 Lecture outline

1. The derivative as a linear transformation

   When we say that function f is differentiable at $x = a$ and that $f'(a) = m$, what we mean is that the function $f(a + h) - f(a)$ is well approximated, for small h, by a linear function $L(h) = mh$, where $m = f'(a)$.

   Show how this idea can be viewed as a "tangent-line approximation" to $f(x)$ for x near to a.

   In the single-variable case, we usually think of the derivative $f'(a)$ as just a number, not a linear function of an increment h, but that view will not generalize to derivatives in $\mathbb{R}^n$. Here is a view of single-variable calculus that generalizes correctly.

   Any linear function $L(h) = mh$ is represented by the 1×1 matrix $[m]$, which in turn is represented by the real number m.

   The error made by using the tangent-line approximation
   $f(a + h) - f(a) = f'(a)h$ is a "remainder"

$$r(h) = f(a + h) - f(a) - f'(a)h.$$

   If f is differentiable, this remainder approaches 0 faster than h, i.e.

$$\lim_{h \to 0} \frac{r(h)}{h} = \lim_{h \to 0} \frac{f(a + h) - f(a) - f'(a)h}{h} = 0.$$

   This definition leads to the standard rule for calculating the number $f'(a)$,

$$f'(a) = \lim_{h \to 0} \frac{f(a + h) - f(a)}{h}.$$

   What mathematical object represents a linear transformation
   $L : \mathbb{R}^n \to \mathbb{R}^m$?

2. Extending this definition to $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$
A linear function $L(\vec{\mathbf{h}})$ is represented by an $m \times n$ matrix.

What matrix represents the linear function

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = L \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 + x_2 \\ 3x_1 - x_2 \end{pmatrix}?$$

When we say that $\mathbf{f}$ is *differentiable* at $\mathbf{a}$, we mean that the function $\mathbf{f}(\mathbf{a} + \vec{\mathbf{h}}) - \mathbf{f}(\mathbf{a})$ is well approximated, for any $\vec{\mathbf{h}}$ whose length is small, by a linear function L, called the derivative $[\mathbf{Df}(\mathbf{a})]$.

The error made by using the approximation is a "remainder"
$\mathbf{r}(\vec{\mathbf{h}}) = \mathbf{f}(\mathbf{a} + \vec{\mathbf{h}}) - \mathbf{f}(\mathbf{a}) - [\mathbf{Df}(\mathbf{a})](\vec{\mathbf{h}})$.
$\mathbf{f}$ is called differentiable if this remainder approaches 0 faster than $|\vec{\mathbf{h}}|$, i.e.

$$\lim_{\vec{\mathbf{h}} \to \vec{\mathbf{0}}} \frac{1}{|\vec{\mathbf{h}}|} \mathbf{r}(\vec{\mathbf{h}}) = \lim_{\vec{\mathbf{h}} \to \vec{\mathbf{0}}} \frac{1}{|\vec{\mathbf{h}}|} (\mathbf{f}(\mathbf{a} + \vec{\mathbf{h}}) - \mathbf{f}(\mathbf{a}) - [\mathbf{Df}(\mathbf{a})](\vec{\mathbf{h}})) = \mathbf{0}.$$

In that case, $[\mathbf{Df}(\mathbf{a})]$ is represented by the Jacobian matrix $[\mathbf{Jf}(\mathbf{a})]$.

Proof: Since L exists and is linear, it is sufficient to consider its action on each standard basis vector. We choose $\vec{\mathbf{h}} = t\vec{\mathbf{e_i}}$ so that $|\vec{\mathbf{h}}| = t$. Knowing that the limit exists, we can use any sequence that converges to the origin to evaluate it, and so

$$\lim_{t \to 0} \frac{1}{t} (\mathbf{f}(\mathbf{a} + t\vec{\mathbf{e_i}}) - \mathbf{f}(\mathbf{a}) - tL\vec{\mathbf{e_i}})) = 0? \text{ and } L(\vec{\mathbf{e_i}}) = \lim_{t \to 0} \frac{1}{t} (\mathbf{f}(\mathbf{a} + t\vec{\mathbf{e_i}}) - \mathbf{f}(\mathbf{a}))$$

What is hard is proving that f is differentiable – that L exists – since that requires evaluating a limit where $\vec{\mathbf{h}} \to \vec{\mathbf{0}}$. Such a limit exists only if every sequence $\vec{\mathbf{h}}_1, \vec{\mathbf{h}}_2, \cdots$ that converges to $\vec{\mathbf{0}}$ leads to the conclusion that

$$\lim \frac{1}{|\vec{\mathbf{h}}_n|} \mathbf{r}(\vec{\mathbf{h}}_n) = 0$$

Mere existence of partial derivatives of $\mathbf{f}$ at $\mathbf{a}$ does not guarantee that $\mathbf{f}$ is differentiable at $\mathbf{a}$. Eventually we will prove (proof 11.2) that f is differentiable at $\mathbf{a}$ if all its partial derivatives are *continuous* there.

3. Proving differentiability and calculating derivatives

   In every case $\mathbf{f}$ is a function from U to $\mathbb{R}^m$, where U is an open subset of $\mathbb{R}^n$.

   - $\mathbf{f}$ is constant: $\mathbf{f} = \mathbf{c}$. Then $[\mathbf{Df}(\mathbf{a})]$ is the zero linear transformation, since

     $$\lim_{\vec{\mathbf{h}}\to\vec{\mathbf{0}}} \frac{1}{|\vec{\mathbf{h}}|}(\mathbf{f}(\mathbf{a}+\vec{\mathbf{h}}) - \mathbf{f}(\mathbf{a}) - [\mathbf{Df}(\mathbf{a})]\vec{\mathbf{h}}) = \lim_{\vec{\mathbf{h}}\to\vec{\mathbf{0}}} \frac{1}{|\vec{\mathbf{h}}|}(\mathbf{c} - \mathbf{c} - \vec{\mathbf{0}}) = \vec{\mathbf{0}}.$$

   - $\mathbf{f}$ is affine: a constant plus a linear function, $\mathbf{f} = \mathbf{c} + L$. $[\mathbf{Df}(\mathbf{a})] = L$, since

     $$\lim_{\vec{\mathbf{h}}\to\vec{\mathbf{0}}} \frac{1}{|\vec{\mathbf{h}}|}(\mathbf{f}(\mathbf{a}+\vec{\mathbf{h}})-\mathbf{f}(\mathbf{a})-[\mathbf{Df}(\mathbf{a})]\vec{\mathbf{h}}) = \lim_{\vec{\mathbf{h}}\to\vec{\mathbf{0}}} \frac{1}{|\vec{\mathbf{h}}|}(\mathbf{c}+L(\mathbf{a}+\vec{\mathbf{h}})-(\mathbf{c}+L(\mathbf{a}))-L(\vec{\mathbf{h}})) = 0.$$

     $\mathbf{f}$ has differentiable components: if $\mathbf{f} = \begin{pmatrix} f_1 \\ \cdot \\ \cdot \\ \cdot \\ f_n \end{pmatrix}$: then $\mathbf{Df}(\mathbf{a}) = \begin{bmatrix} \mathbf{D}f_1(\mathbf{a}) \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{D}f_n(\mathbf{a}) \end{bmatrix}$

     .

   - $\mathbf{f} + \mathbf{g}$ is the sum of two functions $\mathbf{f}$ and $\mathbf{g}$, both differentiable at $\mathbf{a}$. The derivative of $\mathbf{f} + \mathbf{g}$ is the sum of the derivatives of $\mathbf{f}$ and $\mathbf{g}$. (proof on next page)

4. Derivative of a sum

   $\mathbf{f}$ and $\mathbf{g}$ are differentiable at $\mathbf{a}$. The obvious guess is that the derivative of $\mathbf{f}+\mathbf{g}$ is the sum of their derivatives.

   Since $\mathbf{f}$ and $\mathbf{g}$ are differentiable at $\mathbf{a}$, we know that

$$\lim_{\vec{\mathbf{h}}\to\vec{\mathbf{0}}}\frac{1}{|\vec{\mathbf{h}}|}(\mathbf{f}(\mathbf{a}+\vec{\mathbf{h}})-\mathbf{f}(\mathbf{a})-([\mathbf{Df}(\mathbf{a})]\vec{\mathbf{h}})=0.$$

$$\lim_{\vec{\mathbf{h}}\to\vec{\mathbf{0}}}\frac{1}{|\vec{\mathbf{h}}|}(\mathbf{g}(\mathbf{a}+\vec{\mathbf{h}})-\mathbf{g}(\mathbf{a})-([\mathbf{Dg}(\mathbf{a})]\vec{\mathbf{h}})=0.$$

   To prove the obvious guess, show that

$$\lim_{\vec{\mathbf{h}}\to\vec{\mathbf{0}}}\frac{1}{|\vec{\mathbf{h}}|}((\mathbf{f}+\mathbf{g})(\mathbf{a}+\vec{\mathbf{h}})-(\mathbf{f}+\mathbf{g})(\mathbf{a})-([\mathbf{Df}(\mathbf{a})]+[\mathbf{Dg}(\mathbf{a})])\vec{\mathbf{h}})=0.$$

5. Product rule (your proof 11.1):

   Now comes something harder: the product rule for two scalar-valued functions f and g. It is easy to guess what the derivative of fg must be, since in single variable calculus, $(fg)'=fg'+gf'$.

Product rule: $[\mathbf{D}(fg)(\mathbf{a})] = f(\mathbf{a})[\mathbf{D}g(\mathbf{a})] + g(\mathbf{a})[\mathbf{D}f(\mathbf{a})]$.

- Step 1: Write the "remainder" $r(\vec{\mathbf{h}})$ that must have the property

$$\lim_{\vec{\mathbf{h}}\to\vec{\mathbf{0}}} \frac{r(\vec{\mathbf{h}})}{|\vec{\mathbf{h}}|} = 0.$$

  $r(\vec{\mathbf{h}}) = f(\mathbf{a}+\vec{\mathbf{h}})g(\mathbf{a}+\vec{\mathbf{h}}) - f(\mathbf{a})g(\mathbf{a}) - f(\mathbf{a})[\mathbf{D}g(\mathbf{a})]\vec{\mathbf{h}} - g(\mathbf{a})[\mathbf{D}f(\mathbf{a})]\vec{\mathbf{h}}$

- Step 2 – a trick that must be memorized: Subtract and add $f(\mathbf{a})g(\mathbf{a}+\vec{\mathbf{h}})$, and subtract and add $g(\mathbf{a}+\vec{\mathbf{h}})[\mathbf{D}f(\mathbf{a})]\vec{\mathbf{h}}$.
  $r(\vec{\mathbf{h}}) = f(\mathbf{a}+\vec{\mathbf{h}})g(\mathbf{a}+\vec{\mathbf{h}}) - f(\mathbf{a})g(\mathbf{a}+\vec{\mathbf{h}}) + f(\mathbf{a})g(\mathbf{a}+\vec{\mathbf{h}}) - f(\mathbf{a})g(\mathbf{a})$
  $-f(\mathbf{a})[\mathbf{D}g(\mathbf{a})]\vec{\mathbf{h}} - g(\mathbf{a}+\vec{\mathbf{h}})[\mathbf{D}f(\mathbf{a})]\vec{\mathbf{h}} + g(\mathbf{a}+\vec{\mathbf{h}})[\mathbf{D}f(\mathbf{a})]\vec{\mathbf{h}} - g(\mathbf{a})[\mathbf{D}f(\mathbf{a})]\vec{\mathbf{h}}.$

- Step 3: split into three terms, one involving the remainder for f, one involving the remainder for g, and one involving $[\mathbf{D}f(\mathbf{a})]$.
  $r(\vec{\mathbf{h}}) = r_1(\vec{\mathbf{h}}) + r_2(\vec{\mathbf{h}}) + r_3(\vec{\mathbf{h}})$, where
  $r_1(\vec{\mathbf{h}}) = f(\mathbf{a}+\vec{\mathbf{h}})g(\mathbf{a}+\vec{\mathbf{h}}) - f(\mathbf{a})g(\mathbf{a}+\vec{\mathbf{h}}) - g(\mathbf{a}+\vec{\mathbf{h}})[\mathbf{D}f(\mathbf{a})]\vec{\mathbf{h}}$
  $r_2(\vec{\mathbf{h}}) = f(\mathbf{a})g(\mathbf{a}+\vec{\mathbf{h}}) - f(\mathbf{a})g(\mathbf{a}) - f(\mathbf{a})[\mathbf{D}g(\mathbf{a})]\vec{\mathbf{h}}.$
  $r_3(\vec{\mathbf{h}}) = g(\mathbf{a}+\mathbf{h})[\mathbf{D}f(\mathbf{a})]\vec{\mathbf{h}} - g(\mathbf{a})[\mathbf{D}f(\mathbf{a})]\vec{\mathbf{h}}.$

- Step 4: Divide each term by $|\vec{\mathbf{h}}|$, and use the differentiability of f and g to prove that the limit $\lim_{\vec{\mathbf{h}}\to\vec{\mathbf{0}}} \frac{r(\vec{\mathbf{h}})}{|\vec{\mathbf{h}}|}$ is zero. For each term you have the product of two factors: one approaches zero, while the other is bounded.
  $r_1(\vec{\mathbf{h}}) = (f(\mathbf{a}+\vec{\mathbf{h}}) - f(\mathbf{a}) - [\mathbf{D}f(\mathbf{a})]\vec{\mathbf{h}})g(\mathbf{a}+\vec{\mathbf{h}}).$

$$\lim_{\vec{\mathbf{h}}\to\vec{\mathbf{0}}} \frac{r_1(\vec{\mathbf{h}})}{|\vec{\mathbf{h}}|} = 0,$$

  since the first factor over $|\vec{\mathbf{h}}|$ goes to zero and the second is bounded.
  $r_2(\vec{\mathbf{h}}) = f(\mathbf{a})(g(\mathbf{a}+\vec{\mathbf{h}}) - g(\mathbf{a}) - [\mathbf{D}g(\mathbf{a})]\vec{\mathbf{h}}).$

$$\lim_{\vec{\mathbf{h}}\to\vec{\mathbf{0}}} \frac{r_2(\vec{\mathbf{h}})}{|\vec{\mathbf{h}}|} = 0,$$

  since the second factor over $|\vec{\mathbf{h}}|$ goes to zero and the first is constant.
  $r_3(\vec{\mathbf{h}}) = [g(\mathbf{a}+\vec{\mathbf{h}}) - g(\mathbf{a})][\mathbf{D}f(\mathbf{a})]\vec{\mathbf{h}}.$

$$\lim_{\vec{\mathbf{h}}\to\vec{\mathbf{0}}} \frac{r_3(\vec{\mathbf{h}})}{|\vec{\mathbf{h}}|} = 0,$$

  since the first factor goes to zero by continuity and the second factor over $|\vec{\mathbf{h}}|$ is bounded.

6. (Proof 11.1)
   Let $U \subset \mathbb{R}^n$ be an open set, and let f and g be functions from U to $\mathbb{R}$. Prove that if f and g are differentiable at $\mathbf{a}$ then so is fg, and that

$$[\mathbf{D}(fg)(\mathbf{a})] = f(\mathbf{a})[\mathbf{D}g(\mathbf{a})] + g(\mathbf{a})[\mathbf{D}f(\mathbf{a})].$$

   (Simpler than in Hubbard because both f and g are scalar-valued functions)

7. Chain rule in $\mathbb{R}^n$ – not a proof, but still pretty convincing

   The chain rule for differentiating composition of functions in general is as simple as you could hope for on the basis of single-variable calculus. Remember the single-variable version:

   $$(f \circ g)'(a) = f'(g(a))g'(a).$$

   This can be rewritten

   $$[D(f \circ g)(a)] = [Df(g(a))][Dg(a)],$$

   where the square brackets convert the old-style derivatives (numbers) into 1×1 Jacobian matrices.

   This says that the derivative of the composition of f and g is the composition of the linear function "multiply by $g'(a)$" and the linear function "multiply by $f'(g(a))$" Notice that f has to be differentiable at $g(a)$.

   Here is the generalization:

   $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ are open sets, and $\mathbf{a}$ is a point in U at which we want to evaluate a derivative.

   $\mathbf{g} : U \to V$ is differentiable at $\mathbf{a}$, and $[\mathbf{Dg}(\mathbf{a})]$ is a $m \times n$ Jacobian matrix.

   $\mathbf{f} : V \to \mathbb{R}^p$ is differentiable at $\mathbf{g}(\mathbf{a})$, and $[\mathbf{Df}(\mathbf{g}(\mathbf{a}))]$ is a $p \times m$ Jacobian matrix.

   The chain rule states that $[\mathbf{D}(\mathbf{f} \circ \mathbf{g})(\mathbf{a}))] = [\mathbf{Df}(\mathbf{g}(\mathbf{a}))] \circ [\mathbf{Dg}(\mathbf{a})]$.

   Draw a diagram to illustrate what happens in the case $n = m = p = 2$ when you use derivatives to find a linear approximation to
   $(\mathbf{f} \circ \mathbf{g})(\mathbf{a} + \tilde{\mathbf{h}}) - (\mathbf{f} \circ \mathbf{g})(\mathbf{a})$.
   This approximation can be done in a single step or in two steps.

8. Two easy chain rule examples

(a) $\mathbf{g} : \mathbb{R} \to \mathbb{R}^2$ maps time into the position of a particle moving around the unit circle:

$$\mathbf{g}(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}.$$

$f : \mathbb{R}^2 \to \mathbb{R}$ maps a point into the temperature at that point.

$$f \begin{pmatrix} x \\ y \end{pmatrix} = x^2 - y^2$$

The composition $f \circ \mathbf{g}$ maps time directly into temperature .
Confirm that $[D(f \circ g)(t)] = [\mathbf{D}f(\mathbf{g}(t))] \circ [\mathbf{D}\mathbf{g}(t)]$.

(b) Let $\phi : \mathbb{R} \to \mathbb{R}$ be any differentiable function. You can make a function $f : \mathbb{R}^2 \to \mathbb{R}$ that is constant on any circle centered at the origin by forming the composition $f \begin{pmatrix} x \\ y \end{pmatrix} = \phi(x^2 + y^2)$.
Show that f satisfies the partial differential equation $yD_1f - xD_2f = 0$.

9. Connection between Jacobian matrix and derivative

   - If $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ is defined on an open set $U \in \mathbb{R}^n$, and

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}\begin{pmatrix} x_1 \\ ... \\ x_n \end{pmatrix} = \begin{pmatrix} f_1(\mathbf{x}) \\ ... \\ f_m(\mathbf{x}) \end{pmatrix}$$

     the Jacobian matrix $[\mathbf{Jf}(\mathbf{x})]$ is made up of all the partial derivatives of $\mathbf{f}$:

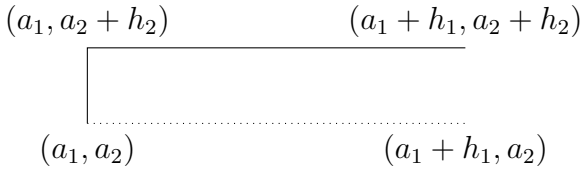
$$[\mathbf{Jf}(\mathbf{a})] = \begin{bmatrix} D_1 f_1(\mathbf{a}) .... D_n f_1(\mathbf{a}) \\ ... \\ D_1 f_m(\mathbf{a}) .... D_n f_m(\mathbf{a}) \end{bmatrix}$$

   - We can invent pathological cases where the Jacobian matrix of f exists (because all the partial derivatives exist), but the function f is not differentiable. In such a case, using the formula

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

     generally gives the wrong answer for the directional derivative! You are trying to use a linear approximation where none exists.

   - Using the Jacobian matrix of partial derivatives to get a good affine approximation for $f(\mathbf{a} + \vec{\mathbf{h}})$ is tantamount to assuming that you can reach the point $\mathbf{a} + \vec{\mathbf{h}}$ by moving along lines that are parallel to the coordinate axes and that the change in the function value along the solid horizontal line is well approximated by the change along the dotted horizontal line. With the aid of the mean value theorem, you can show that this is the case if (proof 11.2) the partial derivatives of f at $\mathbf{a}$ are continuous.



10. Jacobian matrix for a parametrization function gives a good affine approximation

    Here is the function that converts the latitude u and longitude v of a point on the unit sphere to the Cartesian coordinates of that point.

$$f\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \cos u \cos v \\ \cos u \sin v \\ \sin u \end{pmatrix}$$

    Work out the Cartesian coordinates of the point with $\sin u = \frac{3}{5}$ (37 degrees North latitude) and $\sin v = 1$(90 degrees East longitude), and calculate the Jacobian matrix at that point. Then find the best affine approximation to the Cartesian coordinates of the nearby point where u is 0.01 radians less (going south) and v is 0.02 radians greater (going east).

11. A non-differentiable function

Consider a surface where the height z is given by the function

$$f\begin{pmatrix}x\\y\end{pmatrix} = \frac{3x^2y - y^3}{x^2+y^2}; f\begin{pmatrix}0\\0\end{pmatrix} = 0.$$

This function is not differentiable at the origin, and so you cannot calculate its directional derivatives there by using the Jacobian matrix!

(a) Along the first standard basis vector, the directional derivative at the origin is zero. Find two vectors along other directions that also have this property.

(b) Along the second standard basis vector, the directional derivative at the origin is -1.
Find two vectors along other directions that also have this property. (This surface is sometimes called a "monkey saddle," because a monkey could sit comfortably on it with its two legs and its tail placed along these three downward-sloping directions.)

(c) Calculate the directional derivative along an arbitrary unit vector $\vec{\mathbf{e}}_\theta = \begin{pmatrix}\cos\theta\\\sin\theta\end{pmatrix}$. Using the trig identity $\sin 3\theta = 3\sin\theta\cos^2\theta - \sin^3\theta$, quickly rederive the special cases of parts (a) and (b).

(d) Using the definition of the derivative, give a convincing argument that this function is not differentiable at the origin.

12. The mean-value theorem in $\mathbb{R}^n$

For functions of one variable, this is an old friend.

If $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b), then $\exists c \in [a, b]$ such that

$$f(b) - f(a) = f'(c)(b - a)$$

The generalization uses the segment from $\mathbf{a}$ to $\mathbf{b}$ like the closed interval $[a, b]$.

The function f (takes values in $\mathbb{R}$) must be differentiable on an open set U that includes this entire segment.

The conclusion is that

$$f(\mathbf{b}) - f(\mathbf{a}) = [\mathbf{D}f(\mathbf{c})](\mathbf{b} \vec{-} \mathbf{a}).$$

The proof (Hubbard p. 148) is easy. Define a function $\mathbf{h}(t)$ that maps the interval $0 \leq t \leq 1$ uniformly into the segment from $\mathbf{a}$ to $\mathbf{b}$. The formula is

$$\mathbf{h}(t) = (1 - t)\mathbf{a} + t\mathbf{b}$$

Now $g = f \circ \mathbf{h}$ satisfies all the hypotheses of the single-variable mean-value theorem. So there exists t_0 in (0,1) for which

$g(1) - g(0) = g'(t_0)(1 - 0)$.

By the chain rule, $g'(t_0) = [\mathbf{D}f(\mathbf{h}(t_0))]D\mathbf{h}(t_0)$

Set $\mathbf{c} = \mathbf{h}(t_0)$ and this becomes

$$f(\mathbf{b}) - f(\mathbf{a}) = [\mathbf{D}f(\mathbf{c})](\mathbf{b} \vec{-} \mathbf{a}).$$

If points $\mathbf{b}$ and $\mathbf{a}$ differ only in their ith component, so that $\mathbf{b} = \mathbf{a} + \vec{\mathbf{e}}_i$, then

$$f(\mathbf{a} + \vec{\mathbf{e}}_i) - f(\mathbf{a}) = D_i f(\mathbf{a} + t_0\vec{\mathbf{e}}_i), 0 < t_0 < 1.$$

13. (Proof 11.2) Using the mean value theorem, prove that if a function $f : \mathbb{R}^2 \to \mathbb{R}$ has partial derivatives $D_1 f$ and $D_2 f$ that are continuous at $\mathbf{a}$, it is differentiable at $\mathbf{a}$ and its derivative is the Jacobian matrix $\begin{bmatrix} D_1 f(\mathbf{a}) & D_2 f(\mathbf{a}) \end{bmatrix}$.

14. Derivative of a function of a matrix (Example 1.7.17 in Hubbard):

    A matrix is also a vector. When we square an $n \times n$ matrix A, the entries of $S(A) = A^2$ are functions of all the entries of A. If we change A by adding to it a matrix H of small length, we will make a change in the function value A^2 that is a linear function of H plus a small "remainder."

    We could in principle represent A by a column vector with n^2 components and the derivative of S by a very large matrix, but it is more efficient to leave H in matrix form and use matrix multiplication to find the effect of the derivative an a small increment matrix H. The derivative is still a linear function, but it is represented by matrix multiplication in a different way.

    (a) Using the definition of the derivative, show that the linear function that we want is $DS(H) = AH + HA$.

    (b) Confirm that DS is a linear function of H

    (c) Check that $DS(H)$ is a good approximation to $S(A+H)-S(A)$ for the following simple case, where the matrices A and H do not commute.
    $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $H = \begin{bmatrix} 0 & h \\ k & 0 \end{bmatrix}$

15. Derivative of the matrix-inverse function

    Define the function $T(A) = A^{-1}$. We expect that $T(A+H) - T(A)$ can be well approximated by an expression that is linear in H.

    Proof strategy: use a geometric series

    If we were dealing with numbers, we could write

$$\frac{1}{a+h} = \frac{1}{a}\frac{1}{1+\frac{h}{a}} = \frac{1}{a}(1 - \frac{h}{a} + \frac{h^2}{a^2} - ...) =$$

    This approach works with matrices, too, but we must be careful with the order of factors and remember that $(BC)^{-1} = C^{-1}B^{-1}$.

    - Prove that $(A+H)^{-1} = (I + A^{-1}H)^{-1}A^{-1}$.

    - Expand in a geometric series.

    - Evaluate $T(A+H) - T(A)$ and identify the term that is linear in H. Now we have our guess for the derivative.

- The "remainder" is $T(A+H)-T(A)+A^{-1}HA^{-1}$, and we have found that

  $T(A+H) = (A+H)^{-1} =$
  $A^{-1} - A^{-1}HA^{-1} + A^{-1}HA^{-1}HA^{-1} - A^{-1}HA^{-1}HA^{-1}HA^{-1} + \cdots$.

  Get a formula for this remainder that includes two factors of H. Then take its length. Use two strategies:

  Length of product $\leq$ product of lengths.

  Length of sum $\leq$ sum of lengths (generalized triangle inequality.)

- Prove that

$$\lim_{H \to 0} \frac{|Remainder|}{|H|} = 0.$$

16. Chain rule for functions of matrices

    We have shown that the derivative of the squaring function $S(A) = A^2$ is $DS(H) = AH + HA$.

    We also showed that for $T(A) = A^{-1}$, the derivative is $DT(H) = -A^{-1}HA^{-1}$

    Now the function $U(A) = A^{-2}$ can be expressed as the composition $U = S \circ T$.

    Find the derivative $DU(H)$ by using the chain rule.

    The chain rule says "the derivative of a composition is the composition of the derivatives," even in a case like this where composition is not represented by matrix multiplication.

17. Newton's method

(a) One variable: Function f is differentiable. You are trying to solve the equation $f(x) = 0$, and you have found a value a_0, close to the desired x, for which $f(a_0)$ is small. Derive the formula $a_1 = a_0 - f(a_0)/f'(a_0)$ for an improved estimate.

(b) Use Newton's method to find an approximate value for the cube root of 8.1.

(c) n variables: U is an open subset of $\mathbb{R}^n$, and function $\vec{\mathbf{f}}(\mathbf{x}) : U \to \mathbb{R}^n$ is differentiable. You are trying to solve the equation $\vec{\mathbf{f}}(\mathbf{x}) = \vec{\mathbf{0}}$,
and you have found a value $\mathbf{a_0}$, close to the desired $\mathbf{x}$, for which $\vec{\mathbf{f}}(\mathbf{a_0})$ is small. Derive the formula

$$\mathbf{a}_1 = \mathbf{a}_0 - [\mathbf{D}\tilde{\mathbf{f}}(\mathbf{a_0})]^{-1}\vec{\mathbf{f}}(\mathbf{a}_0).$$

for an improved estimate.

18. Newton's method – an example with two variables
We want an approximate solution to the equations

$\log x + \log y = 3$

$x^2 - y = 1$

$$\text{i.e. } f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \log x + \log y - 3 \\ x^2 - y - 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Knowing that $\log 3 \approx 1.1$, show that $\mathbf{x}_0 = \begin{pmatrix} 3 \\ 9 \end{pmatrix}$ is an approximate solution to this equation, then use Newton's method to improve the approximation. Here is a check:

$\log 2.81 + \log 6.87 = 2.98$

$2.81^2 - 6.87 = 1.02$

19. Derivative of inverse function

Suppose that $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ is a continuously differentiable function. Choose a point $\mathbf{x}_0$ where the derivative $[\mathbf{Df}(\mathbf{x}_0)]$ is an invertible matrix. Set $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$. Let $\mathbf{g}$ be the differentiable local inverse function $\mathbf{g} = \mathbf{f}^{-1}$ such that $\mathbf{g}(\mathbf{y}_0) = \mathbf{x}_0$ and $f(\mathbf{g}(\mathbf{y})) = \mathbf{y}$ if $\mathbf{y}$ is close enough to $\mathbf{y}_0$.

Prove that $[\mathbf{Dg}(\mathbf{y}_0)] = [\mathbf{Df}(\mathbf{x}_0)]^{-1}$

20. An economic example of the inverse-function theorem:

Your model: Providing x in health benefits and y in educational benefits leads to happiness H and cost C according the the equation

$$\begin{pmatrix} H \\ C \end{pmatrix} = \mathbf{f}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + x^{0.5}y \\ x^{1.5} + y^{0.5} \end{pmatrix}.$$

Currently, $x = 4, y = 9, H = 22, C = 11$. Your budget is cut, and you are told to adjust x and y to reduce C to 10 and H to 19. Find an approximate solution by using the inverse-function theorem.

We cannot find formulas for the inverse function $\mathbf{g}\begin{pmatrix} H \\ C \end{pmatrix}$ that would solve the problem exactly, but we can calculate the derivative of $\mathbf{g}$.

(a) Check that $[\mathbf{Df}] = \begin{bmatrix} 1 + \frac{y}{2\sqrt{x}} & \sqrt{x} \\ \frac{3}{2}\sqrt{x} & \frac{1}{2\sqrt{y}} \end{bmatrix} = \begin{bmatrix} \frac{13}{4} & 2 \\ 3 & \frac{1}{6} \end{bmatrix}$ is invertible.

(b) Use the derivative $[\mathbf{Dg}] = \begin{bmatrix} -0.03 & 0.36 \\ 0.55 & -0.6 \end{bmatrix}$ to approximate $\mathbf{g}\begin{pmatrix} 19 \\ 10 \end{pmatrix}$

21. Implicit functions – review of the linear case.

From linear algebra we are already familiar with the situation where a system of linear equations has more than one solution. This situation arises when there are fewer equations than unknowns.

We have n unknowns, $n-k$ equations, e.g for $n=3, k=1$
$2x+3y-4z=0$
$x+2y-3z=0$
Create an $(n-k)\times n$ matrix: $T=\begin{bmatrix}2 & 3 & -4\\ 1 & 2 & -3\end{bmatrix}$
If the matrix T is not onto, its rows (the equations) are linearly dependent. Otherwise, when we row reduce, we will find $n-k=2$ pivotal columns and $k=1$ nonpivotal columns. We assign values arbitrarily to the "active" variables that correspond to the nonpivotal columns, and then the values of the "passive" variables that corresponds to the pivotal column are determined.
Suppose that we reorder the unknowns so that the "active" variables come last. Then, after we row reduce the matrix, the first $n-k$ columns will be pivotal. So the first $n-k$ columns will be linearly independent, and they form an invertible square matrix. The matrix is now of the form $T=[A|B]$, where A is invertible.
The solution vector is of the form $\vec{\mathbf{v}}=\begin{bmatrix}\vec{\mathbf{x}}\\ \vec{\mathbf{z}}\end{bmatrix}$, where the passive variables $\vec{\mathbf{x}}$ come first, the active variables $\vec{\mathbf{z}}$ come second.
Substitute this form of $\vec{\mathbf{v}}$ into $T\vec{\mathbf{v}}=\vec{\mathbf{0}}$ and show that if we choose $\vec{\mathbf{z}}$ arbitrarily and set $\vec{\mathbf{x}}=-A^{-1}B\vec{\mathbf{z}}$, we have a solution.

Apply this technique to the given matrix and thereby get a solution where "passive" x and y are expressed in terms of "active" z.

Our system of equations determines $\vec{\mathbf{x}}$ "implicitly" in terms of $\vec{\mathbf{z}}$.

22. The implicit function theorem

Let W be an open subset of $\mathbb{R}^n$, and let $\mathbf{F} : W \to \mathbb{R}^{n-k}$ be a C^1 mapping such that $\mathbf{F}(\mathbf{c}) = \mathbf{0}$. Assume that $[\mathbf{DF}(\mathbf{c})]$ is onto.

Prove that the n variables can be ordered so that the first $n-k$ columns of $[\mathbf{DF}(\mathbf{c})]$ are linearly independent, and that $[\mathbf{DF}(\mathbf{c})] = [A|B]$ where A is an invertible $(n-k) \times (n-k)$ matrix.

Set $\mathbf{c} = \begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix}$, where $\mathbf{a}$ are the $n-k$ passive variables and $\mathbf{b}$ are the k active variables.

Let $\mathbf{g}$ be the "implicit function" from a neighborhood of $\mathbf{b}$ to a neighborhood of $\mathbf{a}$ such that $\mathbf{g}(\mathbf{b}) = \mathbf{a}$ and $\mathbf{F}\begin{pmatrix} \mathbf{g}(\mathbf{y}) \\ \mathbf{y} \end{pmatrix} = 0$.

Prove that $[\mathbf{Dg}(\mathbf{b})] = -A^{-1}B$.

23. Implicit functions – three variables, one equation

The nonlinear equation $F\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x^2 - 4z^2 - 4y^2 - 1 = 0$ implicitly determines x as a function of y and z.
If x is positive, $x = \sqrt{1 + 4y^2 + 4z^2}$; if x is negative, $x = -\sqrt{1 + 4y^2 + 4z^2}$.
Near the point $x = 3, y = 1, z = 1$, which is the appropriate explicit function $x = g\begin{pmatrix} y \\ z \end{pmatrix}$?

There are three ways, in this simple example, to find the derivative of the implicitly defined function g. Two of them take advantage of the fact that there is only one constraint equation and only one active variable.

(a) Solve algebraically to find a formula for the function g that expresses the passive variable x in terms of active variables y and z.

Then calculate $[Dg\begin{pmatrix} 1 \\ 1 \end{pmatrix}]$ for the appropriate function.

(b) Use the implicit function theorem to calculate the partial derivatives of g from the partial derivatives of F:
Calculate the Jacobian matrix $[DF]$ at $x = 3, y = 1, z = 1$.

Split off a square matrix A on the left, so that $[DF] = [A|B]$, and confirm that $[Dg\begin{pmatrix} 1 \\ 1 \end{pmatrix}] = -A^{-1}B$.

(c) Since there is only one passive variable x, just replace x by $g\begin{pmatrix} y \\ z \end{pmatrix}$ to get $h\begin{pmatrix} y \\ z \end{pmatrix} = g\begin{pmatrix} y \\ z \end{pmatrix}^2 - 4z^2 - 4y^2 - 1 = 0$.

Take the partial derivative $D_1 h\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, using the chain rule, and solve for $D_1 g\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. In this context the notation $\frac{\partial g}{\partial y}$ is standard.

24. Finding the derivative of an implicitly defined function

The plane $x + 2y - 3z + 4 = 0$ and the cone $x^2 + y^2 - z^2 = 0$ intersect in a curve that includes the point $\mathbf{c} = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix}$. Near that point this curve is the graph of a function $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{g}(z)$.

Use the implicit function theorem to determine $\mathbf{g}'(5)$, then find the approximate coordinates of a point on the curve with $z = 5.1$.

Check: 2.9+2(4.2) - 3(5.1)+4=0; $2.9^2 + 4.2^2 - 5.1^4 = 0.04$.

25. Curves, Surfaces, Graphs, and Manifolds

Manifolds are a generalization of smooth curves and surfaces to an arbitrary number of dimensions.

The simplest sort of manifold is a flat one. Consider an affine subset $X \subset \mathbb{R}^n$. This subset could be a subspace, or it could be a subspace plus a constant, like a line or plane that does not include the origin.

There are three equivalent ways to describe such a manifold. The first will generalize to the definition of a manifold, but the second and third are more common and useful. Try them out on the line of slope 2 that passes through the point $x = 0, y = -2$.

(a) As the **graph** of a function. There are two ways to do this!

(b) As the **locus** defined by an equation $F\begin{pmatrix} x \\ y \end{pmatrix} = 0$.

(c) By a **parametrization** in term of a point and a direction vector, using a parameter t.

Graphs and manifolds in $\mathbb{R}^3$ and $\mathbb{R}^n$

Example 1: Consider the graph of $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{bmatrix} 2 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} - 4$

This graph is the plane $2x + 3y - z = 4$. It is a subset of the space $\mathbb{R}^3$ whose dimension, $n = 3$, is the sum of the dimension of the domain of f, $k = 2$, and the dimension of the codomain of f, $n - k = 1$.

Describe this manifold as the locus of a function $F : \mathbb{R}^3 \to \mathbb{R}$.

Invent a parametric description of this manifold.

Example 2: Consider the graph of $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{f}(z) = \begin{bmatrix} 3 + 2z \\ 3 - z \end{bmatrix}$

This graph is a line.

Invent a parametric description of this manifold.

Invent a description of this manifold as a locus.

Again, the manifold is a subset of the space $\mathbb{R}^3$ whose dimension, $n = 3$, is the sum of the dimension of the domain of f, $k = 1$, and the dimension of the codomain of f, $n - k = 2$.

Make the function $\mathbf{f} : \mathbb{R}^k \to \mathbb{R}^{n-k}$ arbitrary and we have a graph Γ:

the set of pairs $(\mathbf{x}, \mathbf{y})$ such that $\mathbf{y} = \mathbf{f}(\mathbf{x})$. This is a subset of $\mathbb{R}^n$.

Require that the function be C^1 (continuously differentiable) and we almost have a manifold. But it is asking too much for a single function always to do the job.

26. Definition, and examples in $\mathbb{R}^3$

A subset $M \subset \mathbb{R}^n$ is a smooth manifold if locally it is the graph of a C^1 function (the partial derivatives are continuous). "Locally" means that for any point $\mathbf{x} \in M$ we can find a neighborhood U of $\mathbf{x}$ such that within $M \cap U$, there is a C^1 function that expresses $n-k$ variables in terms of the other k. The number k is the dimension of the manifold. Finding such a function is precisely the job of the implicit function theorem.

- $k = 3$. Any open subset $M \subset \mathbb{R}^3$ is a smooth 3-dimensional manifold. In this case $k = 3$, and the manifold is the graph of a function $f : \mathbb{R}^3 \to \{\vec{\mathbf{0}}\}$, whose codomain is the trivial vector space $\{\vec{\mathbf{0}}\}$ that contains just a single point. Such a function is necessarily constant, and its derivative is zero.

  Why does M have to be open?

- $k = 2$. Consider the graph of $f\begin{pmatrix} x \\ y \end{pmatrix} = x^2 + y^2$.

  What is the name of this surface?

- $k = 1$. Consider the graph of the function $\vec{\mathbf{f}}(z) = \begin{pmatrix} \cos 2\pi z \\ \sin 2\pi z \end{pmatrix}$ What is the name of this curve?

- $k = 0$. In this case the manifold consists of one of more isolated points. Near any of these points $\mathbf{x}_0$, it is the graph of a function $\vec{\mathbf{f}} : \{\vec{\mathbf{0}}\} \to \mathbb{R}^3$ whose domain is the trivial vector space and whose image is the point $\mathbf{x}_0 \in \mathbb{R}^3$.

  This function is differentiable. Its derivative is the zero function. Whatever $\epsilon > 0$ you choose, $\vec{\mathbf{f}}(\vec{\mathbf{0}} + \vec{\mathbf{h}}) - \vec{\mathbf{f}}(\vec{\mathbf{0}}) < \epsilon$ for all $\vec{\mathbf{h}}$, since the only possible $\vec{\mathbf{h}}$ is $\vec{\mathbf{0}}$.

27. Manifolds defined by geometry

For the unit circle, we need four functions to implement the preceding definition.

Show how to break up the circle into four pieces, each of which is a function graph.

What single equation can characterize the unit circle as a locus?

For the unit sphere, we need six functions in place of the single equation $x^2 + y^2 + z^2 = 1$.

How would you slice up an apple to illustrate this?

28. Other examples of manifolds

- An "indifference" surface. Let x = days of vacation per year, y = fraction of health insurance paid by the company, z = number of shares of stock options available for purchase annually. The manifold M consists of all points $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ that employees regard as neither better nor worse than $\begin{pmatrix} 15 \\ 0.5 \\ 100 \end{pmatrix}$
- We are given as $\mathbf{F}$ three equations involving five variables, so that $n = 5, k = 2$. In this case $M \cap U$ might be part of a two-dimensional manifold, a surface in $\mathbb{R}^5$. In economics this is easy to picture: choose five variables that specify how the Federal budget is divided up among Cabinet departments, and invent three equations that impose constraints specified by the Pentagon, the White House, and the Congressional budget office. Three of the budget variables become "passive": they are functions of the other two "active" variables.
- The set of equilibrium states of one mole of helium gas confined to a cylinder fitted with a movable piston. There are many pairs of variables that specify the state (e.g. pressure and volume, or temperature and entropy). This is a 2-dimensional manifold, but it is not an obvious subset of some $\mathbb{R}^n$.
- The set of positions for a pair of particles joined by a rigid rod of length l. Specify x_1, y_1 ,and x_2, with $|x_2 - x_1| \leq l$, and you can calculate y_2, but you need to know whether $y_2 - y_1$ is positive or negative. Make a sketch to illustrate the problem.

- The configuration of a set of four particles joined into a quadrilateral by four rigid rods. This instructive example is discussed in detail on pp. 295-296 of Hubbard, and it provides the basis for some interesting homework problems. Sketch an example to show how the location of one joint depends on the others, but only locally.

29. Identifying a smooth manifold by the implicit function theorem

This is Hubbard Theorem 3.1.10.

We have a subset $M \subset \mathbb{R}^n$ defined as a "locus" and would like to show that M is a smooth k-dimensional manifold.

Suppose there is an open subset $U \subset \mathbb{R}^n$ and a C^1 function $\mathbf{F} : U \to \mathbb{R}^{n-k}$ such that the "locus," the set of solutions of the equation $\mathbf{F}(\mathbf{z}) = 0$, is $M \cap U$.

If $[\mathbf{DF}(\mathbf{z})]$ is onto (surjective) for every $\mathbf{z} \in M \cap U$, then $M \cap U$ is a smooth k-dimensional manifold embedded in $\mathbb{R}^n$.

Make a sketch to illustrate $M \cap U$.

To say that M itself is a manifold, we have to find such a U for every point $\mathbf{z}$ in the manifold, perhaps with different functions at different points.

Suppose that M consists of a circle of radius 1 whose center is the origin and a circle of radius 1 centered at $x = 3, y = 0$. What pair of choices for F and U will define the entire manifold?

Proof: the implicit function theorem says precisely this. The statement that $[\mathbf{DF}(\mathbf{z})]$ is onto guarantees the differentiability of the implicitly defined function.

30. Testing that a locus is a manifold (Example 3.1.11)

$F\begin{pmatrix} x \\ y \end{pmatrix} = x^8 + 2x^3 + y + y^5 - c = 0$. The locus is a smooth manifold, which looks surprisingly normal when plotted for various values of c.

Prove that $[\mathbf{D}F]$ is onto for any value of c.

If you want to represent the manifold as a graph, which should be the active variable?

Is there a function $y = f(x)$ that describes the manifold? Can you find a formula for it?

31. Testing that a locus is a manifold (Example 3.1.12)

$F\begin{pmatrix} x \\ y \end{pmatrix} = x^4 + y^4 + x^2 - y^2 = c$. Construct $[\mathbf{D}F]$, and find the three points where it fails to be onto because both partial derivatives vanish. For each point, choose c so that the point is on the locus.

Here is a plot, generated in R, for various positive values of c.

Here is a plot, generated in R, for various negative values of c.

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 11.1) Let $U \subset \mathbb{R}^n$ be an open set, and let f and g be functions from U to $\mathbb{R}$. Prove that if f and g are differentiable at $\mathbf{a}$ then so is fg, and that

$$[\mathbf{D}(fg)(\mathbf{a})] = f(\mathbf{a})[\mathbf{D}g(\mathbf{a})] + g(\mathbf{a})[\mathbf{D}f(\mathbf{a})].$$

2. Draw a diagram like the one on page 13 to illustrate how the chain rule works for functions from $\mathbb{R}^2$ to $\mathbb{R}^2$.

   $U \subset \mathbb{R}^2$(on the left), $V \subset \mathbb{R}^2$ (in the middle), and $W \subset \mathbb{R}^2$(on the right) are open sets, $\mathbf{a}$ is a point in U, and $\vec{\mathbf{h}}$ is a small increment vector attached to $\mathbf{a}$.

   $\mathbf{g} : U \to V$ is differentiable at $\mathbf{a}$, and its derivative $[\mathbf{D}\mathbf{g}(\mathbf{a})]$ is a 2×2 Jacobian matrix. You can use this Jacobian to find a good approximation to the increment vector that takes you in V from $\mathbf{g}(\mathbf{a})$ to $\mathbf{g}(\mathbf{a} + \vec{\mathbf{h}})$

   $\mathbf{f} : V \to W$ is differentiable at $\mathbf{g}(\mathbf{a})$, and its derivative there, $[\mathbf{D}\mathbf{f}(\mathbf{g}(\mathbf{a}))]$, is a 2×2 Jacobian matrix. You can apply this Jacobian to the increment vector in V and thereby find a good approximation to the increment vector in W, $(\mathbf{f} \circ \mathbf{g})(\mathbf{a} + \vec{\mathbf{h}}) - (\mathbf{f} \circ \mathbf{g})(\mathbf{a})$,

   You could alternatively skip over V and apply the Jacobian $[\mathbf{D}(\mathbf{f} \circ \mathbf{g})(\mathbf{a}))]$ to $\vec{\mathbf{h}}$ to get a good approximation to the increment in W.

   Show that equating these two approximations leads to the chain rule. What makes the proof messy is showing that the difference between the approximations goes rapidly to zero when $|\vec{\mathbf{h}}|$ is small – do not try!.

3. (Proof 11.2) Using the mean value theorem, prove that if a function $f : \mathbb{R}^2 \to \mathbb{R}$ has partial derivatives $D_1 f$ and $D_2 f$ that are continuous at $\mathbf{a}$, it is differentiable at $\mathbf{a}$ and its derivative is the Jacobian matrix $\begin{bmatrix} D_1 f(\mathbf{a}) & D_2 f(\mathbf{a}) \end{bmatrix}$.

4. (The inverse function theorem)
   Review: suppose that $f : \mathbb{R} \to \mathbb{R}$ is differentiable at x_0, $g : \mathbb{R} \to \mathbb{R}$ is its differentiable inverse function, and $y_0 = f(x_0)$. Apply the chain rule to show that $g'(y_0) = 1/f'(x_0)$.

   Now do the same thing in two dimensions. Suppose that $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ is a continuously differentiable function. Choose a point $\mathbf{x}_0$ where the derivative $[\mathbf{Df}(\mathbf{x}_0)]$ is an invertible matrix. Set $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$. Let $\mathbf{g}$ be the differentiable local inverse function $\mathbf{g} = \mathbf{f}^{-1}$ such that

   $\mathbf{g}(\mathbf{y}_0) = \mathbf{x}_0$ and $f(\mathbf{g}(\mathbf{y})) = \mathbf{y}$ if $\mathbf{y}$ is close enough to $\mathbf{y}_0$.

   Again use the chain rule to prove that $[\mathbf{Dg}(\mathbf{y}_0)] = [\mathbf{Df}(\mathbf{x}_0)]^{-1}$

5. (Proof 11.3)The implicit function theorem

   Let W be an open subset of $\mathbb{R}^n$, and let $\mathbf{F} : W \to \mathbb{R}^{n-k}$ be a C^1 mapping such that $\mathbf{F}(\mathbf{c}) = \mathbf{0}$. Assume that $[\mathbf{DF}(\mathbf{c})]$ is onto.

   Prove that the n variables can be ordered so that the first $n-k$ columns of $[\mathbf{DF}(\mathbf{c})]$ are linearly independent, and that $[\mathbf{DF}(\mathbf{c})] = [A|B]$ where A is an invertible $(n-k) \times (n-k)$ matrix.

   Set $\mathbf{c} = \begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix}$, where $\mathbf{a}$ are the $n-k$ passive variables and $\mathbf{b}$ are the k active variables.

   Let $\mathbf{g}$ be the "implicit function" from a neighborhood of $\mathbf{b}$ to a neighborhood of $\mathbf{a}$ such that $\mathbf{g}(\mathbf{b}) = \mathbf{a}$ and $\mathbf{F}\begin{pmatrix} \mathbf{g}(\mathbf{y}) \\ \mathbf{y} \end{pmatrix} = 0$.

   Prove that $[\mathbf{Dg}(\mathbf{b})] = -A^{-1}B$.

6. (Extra topic) Here is a simple example of a non-differentiable function:

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \frac{x^2 y}{x^2 + y^2}; f\begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0.$$

   Prove that this function is not differentiable at $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ by using the definition of the directional derivative,

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

   to calculate the directional derivative at $\mathbf{a} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ for the three unit vectors

$$\vec{\mathbf{v}}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \vec{\mathbf{v}}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \vec{\mathbf{v}}_3 = \begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{pmatrix}$$

and concluding that the directional derivative is not a linear function of $\vec{\mathbf{v}}$.

4 Workshop Problems

1. Chain rule

   (a) Chain rule for matrix functions

   On page 23 of the lecture outline, we obtained the differentiation formula for $U(A) = A^{-2}$ by writing $U = S \circ T$ with $S(A) = A^2, T(A) = A^{-1}$. Prove the same formula from the chain rule in a different way, by writing $U = T \circ S$. You may reuse the formulas for the derivatives of S and T:

   If $S(A) = A^2$ then $[DS(A)](H) = AH + HA$.

   If $T(A) = A^{-1}$ then $[DT(A)](H) = -A^{-1}HA^{-1}$.

   (b) Chain rule with 2×2 matrices

   Start with a pair of polar coordinates $\begin{pmatrix} r \\ \theta \end{pmatrix}$.

   Function $\mathbf{g}$ converts them to Cartesian $\begin{pmatrix} x \\ y \end{pmatrix}$.

   Function $\mathbf{f}$ then converts $\begin{pmatrix} x \\ y \end{pmatrix}$ to $\begin{pmatrix} 2xy \\ x^2 - y^2 \end{pmatrix}$.

   Confirm that $[\mathbf{D}(\mathbf{f} \circ \mathbf{g})(\begin{pmatrix} r \\ \theta \end{pmatrix}))] = [\mathbf{Df}(\mathbf{g}\begin{pmatrix} r \\ \theta \end{pmatrix})] \circ [\mathbf{Dg}\begin{pmatrix} r \\ \theta \end{pmatrix}]$

2. Issues of differentiability

   (a) Suppose that A is a matrix and S is the cubing function given by the formula $S(A) = A^3$. Prove that S is differentiable and that its derivative is the linear function of the matrix H given by the formula $[DS(A)](H) = A^2H + AHA + HA^2$.

   The proof consists in showing that the length of the "remainder" goes to zero faster than the length of the matrix H.

   (b) A continuous but non-differentiable function

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

      i. As in seminar topic 6, show that both partial derivatives vanish at the origin, so that the Jacobian matrix at the origin is the zero matrix $[0\ 0]$, but that the directional derivative along $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is not zero. How does this calculation show that the function is not differentiable at the origin?

      ii. For all points except the origin, the partial derivatives are given by the formulas

$$D_1 f\begin{pmatrix} x \\ y \end{pmatrix} = \frac{2xy^3}{(x^2 + y^2)^2}, D_2 f\begin{pmatrix} x \\ y \end{pmatrix} = \frac{x^4 - x^2 y^2}{(x^2 + y^2)^2}$$

      Construct a "bad sequence" of points approaching the origin to show that $D_1 f$ is discontinuous at the origin.

3. Inverse functions and Newton's method

(a) An approximate solution to the equations

$$x^3 + y^2 - xy = 1.08$$

$$x^2 y + y^2 = 2.04$$

is $x_0 = 1$, $y_0 = 1$.

Use one step of Newton's method to improve this approximation.

(b) You are in charge of building the parking lots for a new airport. You have ordered from amazon.com enough asphalt to pave 1 square kilometer, plus 5.6 kilometers of chain-link fencing. Your plan is to build two square, fenced lots. The long-term lot is a square of side x=0.8 kilometers; the short-term lot is a square of side y=0.6 kilometers. The amount of asphalt A and the amount C of chain-link fencing required are then specified by the function

$$\begin{pmatrix} A \\ C \end{pmatrix} = F \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x^2 + y^2 \\ 4x + 4y \end{pmatrix},$$

Amazon decides to be generous. They deliver enough asphalt to pave 1.0032 square kilometers and 5.616 kilometers of fence.

i. Use the inverse-function theorem to find approximate new values for x and y that use what was shipped to you.

ii. Find a case where $A = 1$ but the value of C is such that this approach will fail because $[DF]$ is not onto. (This case corresponds to the maximum amount of fencing.)

4. Problems to be solved using R (if your group has experience with R, solve one of these as your third problem.)

   Both problems can by done by modifying R script 3.3B. Since Newton's method is the only generally applicable technique for solving nonlinear equations, it is really useful to know how to do it on a computer.

   (a) Saving Delos

   The ancient citizens of Delos, threatened with a plague, consulted the oracle of Delphi, who told them to construct a new cubical altar to Apollo whose volume was double the size of the original cubical altar. (For details, look up "Doubling the cube" on Wikipedia.)

   If the side of the original altar was 1, the side of the new altar had to be the real solution to $f(x) = x^3 - 2 = 0$.

   Numerous solutions to this problem have been invented. One uses a "marked ruler" or "neusis"; another uses origami.

   Your job is to use multiple iterations of Newton's method to find an approximate solution for which $x^3 - 2$ is less than 10^{-8} in magnitude.

   (b) An approximate solution to the system of nonlinear equations

$$x + y^2 + z^3 = 9$$
$$xy + xz + yz = 12$$
$$xyz = 7$$

   is $x = 3, y = 2, z = 1$.

   Use two iterations of Newton's method to find a good approximate solution to these equations.

5 Homework(due Dec. 4, but do problems 1-4 by Nov. 27)

1. (a) Hubbard, Exercise 1.7.21 (derivative of the determinant function). This is really easy if you work directly from the definition of the derivative.

   (b) Generalize this result to the 3×3 case. Hint: consider a matrix whose columns are $\vec{\mathbf{e}}_1 + h\vec{\mathbf{a}}_1$, $\vec{\mathbf{e}}_2 + h\vec{\mathbf{a}}_2$, $\vec{\mathbf{e}}_3 + h\vec{\mathbf{a}}_3$, and use the definition of the determinant as a triple product.

2. Chain rule: an example with 2×2 matrices
   A similar example with a 3×3 matrix is on page 151 of Hubbard.

   The function
   $\mathbf{f}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{2}(x+y) \\ \sqrt{xy} \end{pmatrix}$ was invented by Gauss about 200 years ago to deal with integrals of the form

$$\int_{-\infty}^{\infty} \frac{dt}{\sqrt{(t^2+x^2)(t^2+y^2)}}.$$

   It was revived in the late 20th century as the basis of the AGM (arithmetic-geometric mean) method for calculating π. You can get 1 million digits with a dozen or so iterations.

   The function is meant to be composed with itself; so it will be appropriate to compute the derivative of $\mathbf{f} \circ \mathbf{f}$ by the chain rule.

   (a) $\mathbf{f}$ is differentiable whenever x and y are positive; so its derivative is given by its Jacobian matrix. Calculate this matrix.

   We choose to evaluate the derivative of $\mathbf{f} \circ \mathbf{f}$ at the point $\begin{pmatrix} 8 \\ 2 \end{pmatrix}$.

   Conveniently, $\mathbf{f}\begin{pmatrix} 8 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \end{pmatrix}$. The chain rule says that

$$[\mathbf{D}(\mathbf{f} \circ \mathbf{f})]\begin{pmatrix} 8 \\ 2 \end{pmatrix} = [\mathbf{Df}\begin{pmatrix} 5 \\ 4 \end{pmatrix}][\mathbf{Df}\begin{pmatrix} 8 \\ 2 \end{pmatrix}].$$

   Evaluate the two numerical Jacobian matrices. Because the derivative of f is evaluated at two different points, they will not be the same.

   (b) Write the formula for $\mathbf{f} \circ \mathbf{f}$, compute and evaluate the lower left-hand entry in its Jacobian matrix, and check that it agrees with the value given by the chain rule.

3. Hubbard, Exercise 1.8.6, part (b) only. In the case where $\mathbf{f}$ and $\mathbf{g}$ are functions of time t, this formula finds frequent use in physics. You can either do the proof as suggested in part (a) or model your proof on the one for the dot product on page 143.

4. (similar to the second example on page 14)

   Hubbard, Exercise 1.8.9. The equation that you prove can be called a "first-order partial differential equation."

5. (similar to workshop problem 1a)

   We know the derivatives of the matrix-squaring function S and the matrix-inversion function T:

   If $S(A) = A^2$ then $[DS(A)](H) = AH + HA$.

   If $T(A) = A^{-1}$ then $[DT(A)](H) = -A^{-1}HA^{-1}$.

   (a) Use the chain rule to find a formula for the derivative of the function $U(a) = A^4$.

   (b) Use the chain rule to find a formula for the derivative of the function $W(a) = A^{-4}$.

6. (a) Hubbard, problem 2.10.2. Make a sketch to show how this mapping defines an alternative coordinate system for the plane, in which a point is defined by the intersection of two hyperbolas.

   (b) The point $x = 3, y = 2$ is specified in this new coordinate system by the coordinates $u = 6, v = 5$. Use the derivative of the inverse function to find approximate values of x and y for a nearby point where $u = 6.5, v = 4.5$. (This is essentially one iteration of Newton's method.)

   (c) Find h such that the point $u = 6 + h, v = 5.1$ has nearly the same x-coordinate as $u = 6, v = 5$.

   (d) Find k such that the point $x = 3 + k, y = 2.1$ has nearly the same u-coordinate as $x = 3, y = 2$.

   (e) For this mapping, you can actually find a formula for the inverse function that works in the region of the plane where x, y, u, and v are all positive. Find the rather messy formulas for x and y as functions of u and v, and use them to answer the earlier questions. Once you calculate the Jacobian matrix and plug in appropriate numerical values, you will be back on familiar ground.

7. (similar to workshop problem 2b)

   As a summer intern, you are given the job of reconciling the Democratic and Republican proposals for tax reform. Both parties agree on the following model:

   - x is the change in the tax rate for the middle class.
   - y is the change in the tax rate for the well-off.
   - The net impact on revenue is given by the function

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

   The Republican proposal is $y = -x$, while the Democratic proposal is $y = x$.

   (a) Show that f is continuous at the origin.

   (b) Show that both proposals are revenue neutral by calculating two appropriate directional derivatives. You will have to use the definition of the directional derivative, not the Jacobian matrix.

   (c) At the request of the White House, you investigate a 50-50 mix of the two proposals, the compromise case where $y = 0$, and you discover that it is not revenue neutral! Confirm this surprising conclusion by showing that the directional derivatives at the origin cannot be given by a linear function; i.e. that f is not differentiable.

   (d) Your final task is to explain the issue in terms that legislators can understand: the function is not differentiable because its partial derivatives are not continuous. Demonstrate that one of the partial derivatives of f is discontinuous at the origin. ($D_2 f$ is less messy.)

8. The CEO of a chain of retail stores will get a big bonus if she hits her volume and profit targets for December exactly. Her microeconomics consultant, fresh out of Harvard, tells her that both her target figures are functions of two variables, investment x in Internet advertising and investment y in television advertising. The former attracts savvier customers and so tends to contribute to volume more than to profit.

   The function that determines volume V and profit P is

$$\begin{pmatrix} V \\ P \end{pmatrix} = \begin{pmatrix} x^{\frac{3}{4}} y^{\frac{1}{3}} + x \\ x^{\frac{1}{4}} y^{\frac{2}{3}} + y \end{pmatrix}.$$

   With $x = 16, y = 8, V = 32, P = 16$, our CEO figures she is set for a big bonus. Suddenly, the board of directors, feeling that Wall Street is looking as much for profit as for volume this year, changes her targets to $V = 24, P = 24$. She needs to modify x and y to meet these new targets.

   Near $V = 32, P = 16$, there is an inverse function such that $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{g}\begin{pmatrix} V \\ P \end{pmatrix}$. Find its derivative [$\mathbf{Dg}$], and use the derivative to find values of x and y that are an approximate solution to the problem. Because the increments to V and P are large, you should not expect the approximate solution to be very good, but it will be better than doing nothing.

9. (Hubbard, exercise 3.12)

   Let $X \subset \mathbb{R}^3$ be the set of midpoints of segments joining a point of the curve C_1 of equation $y = x^2, z = 0$ to a point of the curve C_2 of equation $z = y^2, x = 0$.

   (a) Parametrize C_1 and C_2.

   (b) Parametrize X.

   (c) Find an equation for X (i.e. describe X as a locus)

   (d) Show that X is a smooth surface.

10. Manifold M is known by the equation $F\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x^2 + y^4 - 2z^2 - 2 = 0$ near the point $\mathbf{c} = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}$.

    (a) Locally, near $\mathbf{c}$, M is the graph of a function $x = g\begin{pmatrix} y \\ z \end{pmatrix}$. Determine $[\mathbf{D}g(\mathbf{c})]$ by using the implicit function theorem.

    (b) Use $[\mathbf{D}g(\mathbf{c})]$ to find the approximate value of x for a point of M near $\mathbf{c}$ for which $y = 1.1, z = 1.8$.

    (c) Check your answers by finding an explicit formula for g and taking its derivative.

Optional extra problems to be solved using R

11. In problem 8, use multiple iterations of Newton's method in R to find accurate values of x and y that meet the revised targets. Feel free to modify Script 3.3C.

12. (Related to group problem 4a)

    The quintic equation $x(x^2-1)(x^2-4)=0$ clearly has five real roots that are all integers. So does the equation $x(x^2-1)(x^2-4)-1=0$, but you have to find them numerically. Get all five roots using Newton's method, carrying out enough iterations to get an error of less than .001. Use R to do Newton's method and to check your answers. If you have R plot a graph, it will be easy to find an initial guess for each of the five roots.

1. Affine approximation and Newton's Method

   (a) Beginning with the formula for affine approximation (from Week 9), derive the formula for Newton's Method (want to approximate the zeroes of a function)

   (b) After a magical Boston snowfall, you and your pset buddies decide to have a snowball fight. Your team wants to practice throwing snowballs for x hours and construct a snowball-fight strategy plan for y hours. You are subject to the following constraints: First, you can prepare only for a total of 8 hours (so $x + y = 8$). Second, you decide that practicing throwing is more important than strategizing, so you create the constraint $x + y^3 = 27$ so that you'll spend a maximum of 3 hours strategizing. You guess that an approximate solution to these equations is $x = 6$, $y = 2$. Using **one iteration** of Newton's Method, improve your estimate.

2. Non-differentiable functions. Consider the function f given by:

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{cases} 0 & \text{if } \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ \frac{xy^2}{x^2+y^2} & \text{otherwise} \end{cases}$$

   (a) Using the definition of the directional derivative (from last week), calculate $\nabla_{\vec{e_1}} f\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and $\nabla_{\vec{e_2}} f\begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

   (b) If the derivative exists, what should the directional derivative along $\vec{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ be? Why?

   (c) Calculate the directional derivative from (b). What can you conclude and why?

3. What is a manifold? A highly non-rigorous way to describe it is that "a manifold is subset of $\mathbb{R}^n$ where, if you zoom in far enough, it looks flat." Let's try it out!

4. Ways to describe manifolds (k-dimensional manifold in $\mathbb{R}^n$)

   (a) **Graph-making function:** the *graph* $\tilde{g}$ of a graph-making function g that takes in k active variables and outputs $n - k$ passive variables

   (b) **Parametrization:** a map $\gamma : \mathbb{R}^k \to \mathbb{R}^n$ that takes in k parameters and returns a point on the manifold in n-dimensional space

   (c) **Locus function:** a map $F : \mathbb{R}^n \to \mathbb{R}^{n-k}$ that imposes constraints (lowering your degrees of freedom). A manifold is specified by the set of points hwere these functions are 0

5. True/False

   (a) Using enough iterations of Newton's mthod, any initial guess $\vec{x_0}$ will let you approximate some zero of a function f.

   (b) Any finite collection of points in $\mathbb{R}^2$ is a smooth manifold.

   (c) For any functions $f, g : \mathbb{R}^m \to \mathbb{R}^n$, $[D(f \circ g)(t)] = [D(f(g(t)))]$.

   (d) If all partials of a function $f : \mathbb{R}^m \to \mathbb{R}^n$ are continuous at a point $\vec{x}$, then f is differentiable at $\vec{x}$.

MATHEMATICS 23a/E-23a, Fall 2018
Linear Algebra and Real Analysis I
Fortnight 12 (Tangent spaces, Critical points, Lagrange multipliers)

Author: Paul Bamberg
R scripts by Paul Bamberg
Last modified: August 13, 2018 by Paul Bamberg
The seminar and workshop happen during Reading Period.
Fortunately, the final exam is not until Dec. 17

Reading

- Hubbard, Section 3.2 (Tangent spaces)
- Hubbard, Section 3.6 (Critical points)
- Hubbard, Section 3.7 through page 354 (constrained critical points)

Recorded Lectures

- Lecture 25 (Fortnight 12, Class 2) (watch on December 4 or 5)
- Lecture 26 (Fortnight 12, Class 3) (watch on December 6 or 7)

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

- 12.1(Hubbard, theorems 3.6.3 and 3.7.1) Let $U \in \mathbb{R}^n$ be an open subset and let $f : U \to \mathbb{R}$ be a C^1 (continuously differentiable) function.
  First prove, using a familiar theorem from single-variable calculus, that if $\mathbf{x}_0 \in U$ is an extremum, then $[\mathbf{D}f(\mathbf{x}_0)] = [0]$.
  Then prove that if $M \subset \mathbb{R}^n$ is a k-dimensional manifold, and $\mathbf{c} \in M \cap U$ is a local extremum of f restricted to M, then $T_{\mathbf{c}}M \subset \ker[\mathbf{D}f(\mathbf{c})]$.

- 12.2(Special case of Hubbard, theorem 3.7.5) Let M be a manifold known by a real-valued C^1 function $F(\mathbf{x}) = 0$, where F goes from an open subset U of $\mathbb{R}^n$ to $\mathbb{R}$ and $[\mathbf{D}F(\mathbf{x})]$ is nowhere zero. Let $f : U \to \mathbb{R}$ be a C^1 function.

  Prove that $\mathbf{c} \in M$ is a critical point of f restricted to M if and only if there exists a Lagrange multiplier λ such that $[\mathbf{D}f(\mathbf{c})] = \lambda[\mathbf{D}F(\mathbf{c})]$.

R Scripts

- Script 3.4A-ImplicitFunction.R
  Topic 1 - Three variables, one constraint
  Topic 2 - Three variables, two constraints

- Script 3.4B-Manifolds2D.R
  Topic 1 - A one-dimensional submanifold of $\mathbb{R}^2$ – the unit circle
  Topic 2 - Interesting examples from the textbook
  Topic 3 - Parametrized curves in $\mathbb{R}^2$
  Topic 4 - A two-dimensional manifold in $\mathbb{R}^2$
  Topic 5 - A zero-dimensional manifold in$\mathbb{R}^2$

- Script 3.4C-Manifolds3D.R
  Topic 1 - A manifold as a function graph
  Topic 2 - Graphing a parametrized manifold
  Topic 3 - Graphing a manifold that is specified as a locus

- Script 3.4D-CriticalPoints
  Topic 1 - Behavior near a maximum or minimum
  Topic 2 - Behavior near a saddle point

- Script 3.5A-LagrangeMultiplier.R
  Topic 1 - Constrained critical points in $\mathbb{R}^2$

1 Executive Summary

1.1 Using the implicit function theorem

Start with an open subset $U \subset \mathbb{R}^n$ and a C^1 function $\mathbf{F} : U \to \mathbb{R}^{n-k}$. Consider the "locus," $M \cap U$, the set of solutions of the equation $\mathbf{F}(\mathbf{z}) = 0$.
If $[\mathbf{DF}(\mathbf{z})]$ is onto (surjective) for every $\mathbf{z} \in M \cap U$, then $M \cap U$ is a smooth k-dimensional manifold embedded in $\mathbb{R}^n$.
Proof: the implicit function theorem says precisely this. The statement that $[\mathbf{DF}(\mathbf{z})]$ is onto guarantees the differentiability of the implicitly defined function. If $[\mathbf{DF}(\mathbf{z})]$ does not exist or fails to be onto, perhaps even just at a single point, the locus is not a manifold. We use the notation $M \cap U$ because $\mathbf{F}$ may define just part of a larger manifold M that cannot be described as the locus as a single function. To say that M itself is a manifold, we have to find an appropriate U and $\mathbf{F}$ for every point $\mathbf{z}$ in the manifold.

1.2 Parametrizing a manifold

For a k-dimensional submanifold of $\mathbb{R}^n$, the parametrization function is $\gamma : U \to M$, where $U \subset \mathbb{R}^k$ is an open set. The variables in $\mathbb{R}^k$ are called "parameters." The parametrization function γ must be C^1 , one-to-one, and onto M. In other words, we want γ to give us the entire manifold. Finding a local parametrization that gives part of the manifold is of no particular interest, because there is, by definition, a function graph that does that.

An additional requirement: The derivative of the parametrization function is one-to-one for all parameter values. This requirement guarantees that the columns of the the Jacobian matrix $[\mathbf{D}\gamma]$ are linearly independent.

1.3 Tangent space as graph, kernel, or image

Locally, a k-dimensional submanifold M of $\mathbb{R}^n$ is the graph of a function
$\mathbf{g} : \mathbb{R}^k \to \mathbb{R}^{n-k}$. The derivative of $\mathbf{g}$, $[\mathbf{Dg}(\mathbf{b})]$, is an $(n-k) \times k$ matrix that converts a vector of increments to the k active variables, $\dot{\mathbf{y}}$, into a vector of increments to the $n-k$ passive variables, $\dot{\mathbf{x}}$. That is, $\dot{\mathbf{x}} = [\mathbf{Dg}(\mathbf{b})](\dot{\mathbf{y}})$
A point $\mathbf{c}$ of M is specified by the active variables $\mathbf{b}$ and the accompanying passive variables $\mathbf{a}$. The tangent space $T_M(\mathbf{c})$ is the *graph* of this derivative. It is a k-dimensional subspace of $\mathbb{R}^n$.
The k-dimensional manifold M can also be specified as the locus of the equation $\mathbf{F}(\mathbf{z}) = 0$, for $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^{n-k}$ The tangent space $T_{\mathbf{c}}M$ is the *kernel* of the linear transformation $[\mathbf{DF}(\mathbf{c})]$.
Finally, the manifold M can also be described as the image of a parametrization function $\gamma : U \subset \mathbb{R}^k \to \mathbb{R}^n$,

In this case any point of M is the image of some point $\mathbf{u}$ in the parameter space, and the tangent space is $T_{\gamma(\mathbf{u})}M = \text{Img}\,[\mathbf{D}\gamma(\mathbf{u})]$. Whether specified as graph, kernel, or image, the tangent space $T_{\mathbf{c}}M$ is the same! It contains the increment vectors that lead from $\mathbf{c}$ to nearby points that are "almost on the manifold."

1.4 Critical points

Suppose that function $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at point $\mathbf{x}_0$ and that the derivative $[Df(\mathbf{x}_0)]$ is not zero. Then there exists a vector $\vec{\mathbf{v}}$ for which the directional derivative is not zero, the function $g(t) = f(\mathbf{x}_0 + t\vec{\mathbf{v}} - f(\mathbf{x}_0)$ has a nonzero derivative at $t = 0$, and, even if we just consider points that lie on a line through $\mathbf{x}_0$ with direction vector $\vec{\mathbf{v}}$, the function f cannot have a maximum or minimum at $\mathbf{x}_0$. So in searching for a a maximum or minimum of f at points where it is differentiable, we need to consider only "critical points" where $[Df(\mathbf{x}_0] = 0$.

A critical point is not necessarily a maximum or minimum, but for $f : \mathbb{R}^n \to \mathbb{R}$ there is a useful test that generalizes the second-derivative test of single-variable calculus. The proof relies on sections 3.3-3.5 of Hubbard, which we are skipping.

Form the "Hessian matrix" of second partial derivatives (Hubbard, p. 348), evaluated at the critical point x of interest.

$$H_{i,j}(\mathbf{x}) = D_i D_j f(\mathbf{x}).$$

H is a symmetric matrix. If it has a basis of eigenvectors and none of the eigenvalues are zero, we can classify the critical point.

If H has a basis of eigenvectors, all with positive eigenvalues, the critical point is a minimum.

If H has a basis of eigenvectors, all with negative eigenvalues, the critical point is a maximum.

If H has a basis of eigenvectors, some with positive eigenvalues and some with negative eigenvalues, the critical point is a saddle: it is neither a maximum or a minimum.

1.5 Constrained critical points

These are of great important in physics, economics, and other areas to which mathematics is applied.

Consider a point $\mathbf{c}$ on manifold M where the function $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable. Perhaps f has a maximum or minimum at c when its value is compared to the value at nearby points on M, even though there are points not on M where f is larger or smaller. . In that case we should not consider all increment vectors, but only those increment vectors $\vec{\mathbf{v}}$ that lie in the tangent space to the manifold. The derivative $[\mathbf{D}f(\mathbf{c})]$ does not have to be the zero linear transformation, but it has to give zero when applied to any increment that lies in the tangent space $T_{\mathbf{c}}M$, or

$T_{\mathbf{c}}M \subset \text{Ker}[\mathbf{D}f(\mathbf{c})]$.

When manifold M is specified as the locus where some function $\mathbf{F} = 0$, there is an ingenious way of finding constrained critical points by using "Lagrange multipliers,"but not this week!

1.6 Constrained critical points - three approaches

We have proved the following:
If $M \subset \mathbb{R}^n$ is a k-dimensional manifold, and $\mathbf{c} \in M \cap U$ is a local extremum of f restricted to M, then $T_{\mathbf{c}}M \subset \ker[\mathbf{D}f(\mathbf{c})]$.

Corresponding to each of the three ways that we can "know" the manifold M, there is a technique for finding the critical points of f restricted to M.

- Manifold as a graph
  Near the critical point, the passive variables $\mathbf{x}$ are a function $\mathbf{g}(\mathbf{y}$ of the active variables $\mathbf{y}$. Define the graph-making function

  $$\tilde{\mathbf{g}}(\mathbf{y}) = \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix}$$

  Now $f(\mathbf{g}(\mathbf{y})$ specifies values of f only at points on the manifold. Just search for unconstrained critical points of this function by setting $[Df \circ \tilde{\mathbf{g}}(\mathbf{y})] = 0$. This approach works well if you can represent the entire manifold as a single function graph.

- Parametrized manifold
  Points on the manifold are specified by a parametrization $\gamma(\mathbf{u})$.
  Now $f(\gamma(\mathbf{u}))$ specifies values of f only at points on the manifold. Just search for unconstrained critical points of this function by setting $[Df \circ \gamma(\mathbf{u})] = 0$. This approach works well if you can parametrize the entire manifold.

- Manifold specified by constraints
  Points on the manifold all satisfy the constraints $\mathbf{F}(\mathbf{x}) = 0$.
  In this case we know that
  $T_{\mathbf{c}}M = \text{Ker}[\mathbf{DF}(\mathbf{c})]$, so the rule for a critical point becomes
  $\text{Ker}[\mathbf{DF}(\mathbf{c})] \subset \text{Ker}[\mathbf{D}f(\mathbf{c})]$.

  If there is just a single constraint $F(\mathbf{x}) = 0$, both derivative matrices consist of just a single row, and we can represent the condition for a critical point as Ker $\alpha \subset$ Ker β.
  Suppose that $\vec{\mathbf{v}} \in$ ker α and that $\beta = \lambda\alpha$. The quantity λ is called a Lagrange multiplier. Then by linearity, $[\mathbf{D}f(\mathbf{c})]\vec{\mathbf{v}} = \beta\vec{\mathbf{v}} = \lambda\alpha\vec{\mathbf{v}} = 0$.

  So $[\mathbf{D}f(\mathbf{c})]\vec{\mathbf{v}} = 0$ for any vector in the tangent space of $F = 0$, and we have a constrained critical point.

  It is not quite so obvious that the condition $\beta = \lambda\alpha$ is necessary as well as sufficient. We will need to do a proof by contradiction (proof 12.3).

2 Lecture Notes

1. Parametrizing a manifold

   It the spring term we will need to integrate over manifolds in order to evaluate the line integrals and surface integrals that are crucial in physics. The only practical way to do this is by parametrizing the manifold as the image of an open set in $\mathbb{R}^k$. The variables in $\mathbb{R}^k$ are called "parameters," and the functions on the manifold that give the parameter values are called "coordinate functions."

   What parameters do geographers use for the surface of the Earth?

   Here are the strict requirements for the parametrization of a k-dimensional manifold M in $\mathbb{R}^n$, as given in Hubbard, Definition 3.1.18.

   - The parametrization function is $\gamma : U \to M$, where $U \subset \mathbb{R}^k$ is an open set.

     What problem with differentiability would arise if U were not open?

   - The parametrization function γ is C^1 , one-to-one, and onto M. In other words, we want γ to give us the entire manifold. Finding a local parametrization that gives part of the manifold is of no particular interest, because there is, by definition, a function graph that does that.
   - The derivative of the parametrization function is one-to-one for all parameter values.

     If $k < n$, does $[\mathbf{D}\gamma]$ have more rows or more columns? Could it be onto in this case?

     What does this requirement say about the columns of $[\mathbf{D}\gamma]$?

     If $k = 1$ (a smooth curve), how can $[\mathbf{D}\gamma]$ fail to be one-to-one?

2. Examples of parametrizations, with some problems

   - A circle, traced out by a particle that goes around in 2π seconds, parametrized by time. The notation used is standard in physics texts.

     $$\vec{\mathbf{r}}(t) = \begin{bmatrix} \cos t \\ \sin t \end{bmatrix}, 0 \leq t < 2\pi$$

     What requirement is violated, and why is there no way to fix the problem?

   - Polar coordinates to parametrize the manifold $\mathbb{R}^2$.

     $$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}, 0 \leq \theta < 2\pi, r \geq 0$$

     Where is the requirement that $[\mathbf{D}\gamma]$ is one-to-one not met? What is the related problem with the coordinate function that assigns θ to each point in the plane?

   - Part of the unit sphere (a surface), parametrized by longitude θ and latitude ϕ. The choice of ϕ is standard in geography but not in physics.

     $$\vec{\mathbf{r}}\begin{pmatrix} \theta \\ \phi \end{pmatrix} = \begin{pmatrix} \cos\theta\cos\phi \\ \sin\theta\cos\phi \\ \sin\phi \end{pmatrix}, 0 < \theta < \pi, 0 < \phi < \frac{\pi}{2}$$

     What part of the sphere is the manifold given by this parametrization?

     What problems occur if we try to use this parametrization for the entire sphere?

3. Graph of a derivative

The concept of the tangent line to the graph of a function $f : \mathbb{R} \to \mathbb{R}$ is familiar from single-variable calculus. In general, a tangent line is just an affine subset, not a vector space. Similarly, the tangent plane to the graph of a function $f : \mathbb{R}^2 \to \mathbb{R}$ is an affine subset but not a vector space.

In either case, if the graph happened to pass through the origin, the tangent line or plane would be a vector space.

Locally, a k-dimensional submanifold M of $\mathbb{R}^n$ is the graph of a function $\mathbf{g} : \mathbb{R}^k \to \mathbb{R}^{n-k}$. We can denote the vector of k active variables by $\mathbf{b}$ and the vector of $n-k$ passive variables by $\mathbf{a} = \mathbf{g}(\mathbf{b})$.

For consistency with the implicit function theorem, we put the active variables last, so a point of the manifold is

$$\mathbf{c} = \begin{pmatrix} \mathbf{g}(\mathbf{b}) \\ \mathbf{b} \end{pmatrix}$$

The derivative of the function $\mathbf{g}$ for active variables $\mathbf{b}$, $[\mathbf{Dg}(\mathbf{b})]$,
is an $(n-k) \times k$ matrix. It converts a vector of increments to the k active variables, $\dot{\mathbf{y}}$, into a vector of increments to the $n-k$ passive variables, $\dot{\mathbf{x}}$. That is, $\dot{\mathbf{x}} = [\mathbf{Dg}(\mathbf{b})](\dot{\mathbf{y}})$

Previously we used $\vec{\mathbf{h}}$ to represent the vector of increments on which a derivative acts, but now we have adopted Hubbard's convention of using $\dot{\mathbf{y}}$ for the active "input" to the derivative and $\dot{\mathbf{x}}$ for the passive "output."

Prove that the graph of the derivative is a k-dimensional subspace of $\mathbb{R}^n$. This subspace is called the **tangent space**.

4. Tangent space as a graph (Hubbard Definition 3.2.1)

   Given a point on a manifold, represent the manifold locally as a function graph. Then the tangent space is the graph of the derivative of this function.

   It is true, but not immediately obvious from the definition, that the tangent space comes out the same if we make a different choice of active and passive variables. Try a special case: for example, the circle of radius 5 at the point $x = 3, y = 4$.

   What happens if you express x as a function of y?

   What happens if you express y as a function of x?

   Consider point $\mathbf{c}$ on a manifold M. Call the active variables $\mathbf{b}$ and the passive variables $\mathbf{a}$. The manifold is locally the graph of $\mathbf{x} = \mathbf{g}(\mathbf{y})$.

   Distinguish carefully:

   $\dot{\mathbf{x}} = [\mathbf{Dg}(\mathbf{b})]\dot{\mathbf{y}}$ is the equation whose graph is the tangent **space**.

   $\mathbf{x} - \mathbf{a} = [\mathbf{Dg}(\mathbf{b})](\mathbf{y} - \mathbf{b})$is the equation of the tangent **plane**.

   Example: The graph of $x = y^2 + z^2$ is a paraboloid, centered on the x axis.

   Find the equation of the tangent space at $x = 5, y = 2, z = 1$.

   Find the equation of the tangent plane at $x = 5, y = 2, z = 1$.

5. The equal-dimension lemma

   If subspaces V and W both have dimension k and $V \subset W$, then $V = W$.

   Proof: Choose k basis vectors for V. Since $V \subset W$, these vectors are also k independent vectors in W and therefore also form a basis for W. Thus any vector in W is a linear combination of these vectors and is also a vector in V, i.e., $W \subset V$. But $V \subset W$ and $W \subset V$ means $V = W$.

6. Tangent space as a kernel

   Suppose that a k-dimensional manifold M is specified as the locus of the equation $\mathbf{F}(\mathbf{z}) = 0$ for a function $\mathbf{F}$ whose domain is an open subset of $\mathbb{R}^n$ and whose codomain is $\mathbb{R}^{n-k}$. The implicit function theorem says that such a function is a satisfactory way to describe a manifold at any point $\mathbf{c}$ where $[\mathbf{DF}(\mathbf{c})]$ is onto.

   To prove: the tangent space $T_{\mathbf{c}}M$ is the kernel of the linear transformation $[\mathbf{DF}(\mathbf{c})]$.

   The proof (your proof 12.1) that this method of finding the tangent space is equivalent to the definition is on pp. 310-311 in Hubbard.

   - Write the matrix $[\mathbf{DF}(\mathbf{c})]$ as $[A|B]$, with the columns that correspond to passive variables coming first. Given that $[\mathbf{DF}(\mathbf{c})]$ is onto, the square matrix A will be invertible.
   - The dimension of the image of $[\mathbf{DF}(\mathbf{c})]$ is $n - k$. So the dimension of Ker $[\mathbf{DF}(\mathbf{c})]$, by the rank-nullity theorem, is $n - (n - k) = k$.
   - The derivative of the function $\mathbf{g}(\mathbf{b})$ that expresses the passive variables $\mathbf{a}$ locally in terms of the active variables $\mathbf{b}$ is given by the implicit function theorem:
     $[\mathbf{Dg}(\mathbf{b})] = -A^{-1}B$.

   - A vector in the tangent space can be written as $\dot{\mathbf{z}} = \begin{bmatrix} \dot{\mathbf{x}} \\ \dot{\mathbf{y}} \end{bmatrix}$, where $\dot{\mathbf{x}}$ and $\dot{\mathbf{y}}$ are increments to passive and active variables respectively. Suppose that $\begin{pmatrix} \dot{\mathbf{x}} \\ \dot{\mathbf{y}} \end{pmatrix}$ is an element of Ker $[\mathbf{DF}(\mathbf{c})]$.

     Thus $[A|B] \begin{pmatrix} \dot{\mathbf{x}} \\ \dot{\mathbf{y}} \end{pmatrix} = A\dot{\mathbf{x}} + B\dot{\mathbf{y}} = 0$.

     So $\dot{\mathbf{x}} = -A^{-1}B\dot{\mathbf{y}} = [\mathbf{Dg}(\mathbf{b})]\dot{\mathbf{y}}$.

     We have established that Ker $[\mathbf{DF}(\mathbf{c})]$ is a subspace $T_{\mathbf{c}}M$. But both these spaces have dimension k, so they are equal.

7. Tangent space as a kernel (Proof 12.1 - Hubbard Theorem 3.2.4)

   Suppose that $U \subset \mathbb{R}^n$ is an open subset, $\mathbf{F} : U \to \mathbb{R}^{n-k}$ is a C^1 mapping, and manifold M can be described as the set of points that satisfy $\mathbf{F}(\mathbf{z}) = 0$. Use the implicit function theorem to show that if $[\mathbf{DF}(\mathbf{c})]$ is onto for $\mathbf{c} \in M$, then the tangent space $T_{\mathbf{c}}M$ is the kernel of $[\mathbf{DF}(\mathbf{c})]$. You may assume that the variables have been numbered so that when you row-reduce $[\mathbf{DF}(\mathbf{c})]$, the first $n - k$ columns are pivotal.

8. Tangent space as an image (Hubbard, Proposition 3.2.7)

   Let $U \subset \mathbb{R}^k$ be open, and let $\gamma : U \to \mathbb{R}^n$ be a parametrization of manifold M. Show that

$$T_{\gamma(\mathbf{u})}M = \text{Img}[\mathbf{D}\gamma(\mathbf{u})].$$

9. Summary - three ways to characterize the tangent space of manifold M

   - If M is represented as a function graph $\mathbf{x} = \mathbf{g}(\mathbf{y})$, the tangent space $T_{\mathbf{c}}M$ at $\mathbf{c} = \begin{pmatrix} \mathbf{g}(\mathbf{b}) \\ \mathbf{b} \end{pmatrix}$ is the **graph** of its derivative and $\dot{\mathbf{x}} = [\mathbf{Dg}(\mathbf{b})]\dot{\mathbf{y}}$.
   - If M is represented as the locus of $\mathbf{F}(\mathbf{z}) = \mathbf{0}$, the tangent space at $\mathbf{z} = \mathbf{c}$ is the **kernel** of the derivative of $\mathbf{F}$, $T_{\mathbf{c}}M$= Ker $[\mathbf{DF}(\mathbf{c})]$, and $[\mathbf{DF}(\mathbf{c})] \begin{bmatrix} \dot{\mathbf{x}} \\ \dot{\mathbf{y}} \end{bmatrix} = \mathbf{0}$.
   - If M is represented by a parametrization function $\gamma : U \to \mathbb{R}^n$, the the tangent space at $\mathbf{c} = \gamma(\mathbf{u})$ is the **image** of the derivative of γ: $T_{\gamma(\mathbf{u})}M = \text{Img}[\mathbf{D}\gamma(\mathbf{u})]$.

10. Exploring a manifold

    A cometary-exploration robot is fortunate enough to land on an ellipsoidal comet whose surface is described by the equation

$$x^2 + \frac{y^2}{4} + \frac{z^2}{9} = 9.$$

    Its landing point is $x = 2, y = 4, z = 3$.

    - Prove that the surface of the comet is a smooth manifold.
    - The controllers of the robot want it to move to a nearby point on the surface where $y = 4.02, z = 3.06$. Use the implicit function theorem to determine the approximate x coordinate of this point.
      (Check: $1.98^2 + 4.02^2/4 + 3.06^2/9 = 9.0009$.)
    - Find a basis for the tangent space at the landing point.
    - Find the equation of the tangent plane at the landing point.
      (Check: $4(1.98) + 2(4.02) + (2/3)(3.06) = 18$.)

11. Manifolds – keeping track of dimension

Assume that, at the top level, there are nine categories $x_1, x_2, \cdots, x_9$ in the Federal budget. They must satisfy four constraints:

- One simply fixes the total dollar amount.
- One comes from your political advisors – it makes the budget looks good to likely voters in swing states.
- One comes form Congress - it guarantees that everyone can have his or her earmarks.
- One comes from the Justice Department – it guarantees compliance with all laws.

These four constraints together define a function $\mathbf{F}$ whose derivative is onto for budgets that satisfy the constraints. The acceptable budgets, for which $\mathbf{F}(\mathbf{x}) = 0$, form a k-dimensional submanifold M of $\mathbb{R}^n$.
Specify the dimension of the domain and codomain for

(a) A function $\mathbf{g}$ that specifies that passive variables in terms of the active variables.

(b) The function $\mathbf{F}$ that specifies the constraints.

(c) A parametrization function γ that generates a valid budget from a set of parameters.

For each alternative, specify the shape of the matrix that represents the derivative of the relevant function and explain how, given a valid budget $\mathbf{c}$, it could be used to find a basis for the tangent space $T_{\mathbf{c}}M$.

12. Necessary condition for an extremum (maximum or minimum)

    Recall the rule for a function of one variable:

    Let U be an open interval and $f : U \to \mathbb{R}$ a differentiable function. If x_0 is an extremum, then $f'(x_0) = 0$.

    In $\mathbb{R}^n$ the necessary condition for a differentiable function to have an extremum is the same.

    If $U \subset \mathbb{R}^n$ is open and $f : U \to \mathbb{R}$ is a differentiable function, then at an extremum $\mathbf{x}_0$ of f, $[\mathbf{D}f(\mathbf{x}_0)] = \mathbf{0}$.

    Define the function $g(t) = f(\mathbf{x}_0 + t\mathbf{e}_i)$.

    By the chain rule, $g'(0) = [\mathbf{D}f(\mathbf{x}_0)]\mathbf{e}_i = D_i f(\mathbf{x}_0)$.

    Since $g(t)$ has an extremum at $t = 0$, $g'(0) = D_i f(\mathbf{x}_0) = 0$. Illustrate this argument with a diagram for $\mathbb{R}^2$.

    Therefore each partial derivative of f is zero. Since f is differentiable, its derivative is zero at the extremum.

13. Finding critical points and extrema

    If $U \subset \mathbb{R}^n$ is open and $f : U \to \mathbb{R}$ is a differentiable function,
    a *critical point* of f is any point where its derivative $[\mathbf{D}f]$ vanishes.

    We just proved that an extremum of f must be a critical point. The converse is not true: it is common to have a critical point that is not an extremum.

    Show that for $f(x) = x^3$, $x = 0$ is a critical point but not an extremum.

    Example: $f\begin{pmatrix} x \\ y \end{pmatrix} = xy$.

    Evaluate $[\mathbf{D}f]$ at the origin.

    Find points near the origin where $f > 0$ and where $f < 0$.

14. Finding and classifying critical points

For a function defined on an open subset of $\mathbb{R}^n$, finding and classifying critical points is just an exercise in differentiation and algebra: evaluate all the partial derivatives, set them equal to zero, and try to solve the resulting set of (generally nonlinear) equations.

Here is a test for a critical point of a C^2 function to be a maximum or a minimum that generalizes what you learned in single-variable calculus. The proof relies on sections 3.3-3.5 of Hubbard, which we are skipping.

Form the "Hessian matrix" of second partial derivatives (Hubbard, p. 348), evaluated at the critical point x of interest.

$$H_{i,j}(\mathbf{x}) = D_i D_j f(\mathbf{x}).$$

H is a symmetric matrix. If it has a basis of eigenvectors and none of the eigenvalues are zero, we can classify the critical point.

If H has a basis of eigenvectors, all with positive eigenvalues, the critical point is a minimum.

If H has a basis of eigenvectors, all with negative eigenvalues, the critical point is a maximum.

If H has a basis of eigenvectors, some with positive eigenvalues and some with negative eigenvalues, the critical point is a saddle: it is neither a maximum or a minimum.

Why this test works (not quite a proof):

Suppose first that $f(\vec{\mathbf{x}}) = f(\vec{\mathbf{0}}) + \frac{1}{2}\lambda_1 x_1^2 + \frac{1}{2}\lambda_2 x_2^2 + \cdots + \frac{1}{2}\lambda_n x_n^2$.

Calculate $[\mathbf{D}f(\vec{\mathbf{0}})]$ and H.

Show that $\vec{\mathbf{0}}$ is a critical point and that the test is correct in this case.

Now suppose that all the eigenvalues are real and distinct.

Why is H symmetric?

Prove that any two eigenvectors $\vec{\mathbf{v}}_i$ and $\vec{\mathbf{v}}_j$ are orthogonal.

So we can make an orthonormal basis of eigenvectors. Relative to this basis, the situation is the same as on the preceding page.

Shortcomings of this argument:

- It assumes that the critical point is at the origin.
- It does not deal with the case where there are fewer than n eigenvalues.
- It applies only to quadratic functions, and we have not shown that any C^2 function can be approximated by such a function.
- It ignores the case where an eigenvalue is zero.

Quick test in $\mathbb{R}^2$.

Here $\det H = \lambda_1 \lambda_2$. What can we conclude if

- $\det H < 0$?
- $\det H > 0$?
- $\det H = 0$?

15. Critical points

$f\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{2}x^2 + \frac{1}{3}y^3 - xy$

Calculate the partial derivatives as functions of x and y, and show that the only critical points are $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$

Calculate the Hessian matrix H and evaluate it numerically at each critical point to get matrices H_0 and H_1.

Find the eigenvalues of H_0 and classify the critical point at $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

Find the eigenvalues of H_1 and classify the critical point at $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

16. (Proof 12.2 – Hubbard, theorems 3.6.3 and 3.7.1)
Let $U \in \mathbb{R}^n$ be an open subset and let $f : U \to \mathbb{R}$ be a C^1 (continuously differentiable) function.
First prove, using a familiar theorem from single-variable calculus, that if $\mathbf{x}_0 \in U$ is an extremum, then $[\mathbf{D}f(\mathbf{x}_0)] = [0]$.
Then prove that if $M \subset \mathbb{R}^n$ is a k-dimensional manifold, and $\mathbf{c} \in M \cap U$ is a local extremum of f restricted to M, then $T_{\mathbf{c}}M \subset \ker[\mathbf{D}f(\mathbf{c})]$.

17. Constrained critical points - three approaches

We have proved the following:
If $M \subset \mathbb{R}^n$ is a k-dimensional manifold, and $\mathbf{c} \in M \cap U$ is a local extremum of f restricted to M, then $T_{\mathbf{c}}M \subset \text{Ker}[\mathbf{D}f(\mathbf{c})]$.

Corresponding to each of the three ways that we can "know" the manifold M, there is a technique for finding the critical points of f restricted to M.

- Manifold as a graph
  Near the critical point, the passive variables $\mathbf{x}$ are a function $\mathbf{g}(\mathbf{y})$ of the active variables $\mathbf{y}$. Define the graph-making function

$$\tilde{\mathbf{g}}(\mathbf{y}) = \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix}$$

  Now $f(\tilde{\mathbf{g}}(\mathbf{y}))$ specifies values of f only at points on the manifold. Just search for unconstrained critical points of this function by setting $[Df \circ \tilde{\mathbf{g}}(\mathbf{y})] = 0$. This approach works well if you can represent the entire manifold as a single function graph.

- Parametrized manifold
  Points on the manifold are specified by a parametrization $\gamma(\mathbf{u})$.
  Now $f(\gamma(\mathbf{u}))$ specifies values of f only at points on the manifold. Just search for unconstrained critical points of this function by setting $[Df \circ \gamma(\mathbf{u})] = 0$. This approach works well if you can parametrize the entire manifold.

- Manifold specified by constraints
  Points on the manifold all satisfy the constraints $\mathbf{F}(\mathbf{x}) = 0$.
  In this case we know that
  $T_{\mathbf{c}}M = \text{Ker}[\mathbf{DF}(\mathbf{c})]$, so the rule for a critical point becomes
  $\text{Ker}[\mathbf{DF}(\mathbf{c})] \subset \text{Ker}[\mathbf{D}f(\mathbf{c})]$.

  If there is just a single constraint $F(\mathbf{x}) = 0$, both derivative matrices consist of just a single row, and we can represent the condition for a critical point as $\text{Ker } \alpha \subset \text{Ker } \beta$.
  Suppose that $\vec{\mathbf{v}} \in \text{ker } \alpha$ and that $\beta = \lambda\alpha$. The quantity λ is called a Lagrange multiplier. Then by linearity, $[\mathbf{D}f(\mathbf{c})]\vec{\mathbf{v}} = \beta\vec{\mathbf{v}} = \lambda\alpha\vec{\mathbf{v}} = 0$.
  So $[\mathbf{D}f(\mathbf{c})]\vec{\mathbf{v}} = 0$ for any vector in the tangent space of $F = 0$, and we have a constrained critical point.

  It is not quite so obvious that the condition $\beta = \lambda\alpha$ is necessary as well as sufficient. We will need to do a proof by contradiction (proof 12.3).

18. (Hubbard, theorem 3.7.5 - proof 12.3 is the special case where $m = 1$)
Let M be a manifold known by a real-valued C^1 function $\vec{\mathbf{F}}(\mathbf{x}) = \mathbf{0}$, where $\vec{\mathbf{F}}$ goes from an open subset U of $\mathbb{R}^n$ to $\mathbb{R}^m$ and $[\mathbf{DF}(\mathbf{x})]$ is onto.
Let $f : U \to \mathbb{R}$ be a C^1 function.

Prove that $\mathbf{c} \in M$ is a critical point of f restricted to M if and only if there exist m Lagrange multipliers $\lambda_1, \cdots, \lambda_m$ such that
$[\mathbf{D}f(\mathbf{c})] = \lambda_1[\mathbf{D}F_1(\mathbf{c})] + \cdots + \lambda_m[\mathbf{D}F_m(\mathbf{c})]$.

19. Two approaches to an elementary maximization problem

    Farmer Brown wants to build a rectangular pigpen using fencing of total length 20 meters. One side of the pigpen is his barn, so the width x (side parallel to the barn) and depth y (other two sides) are constrained to lie on the manifold $x + 2y = 20$.

    What choice of x maximizes the area $f\begin{pmatrix} x \\ y \end{pmatrix} = xy$?

    - Solve the problem by elementary methods, using y as the active variable.
    - Solve the problem by using Lagrange multipliers.

20. Using a parametrization

What rectangle inscribed in the ellipse $x^2+4y^2 = 4$ has the greatest perimeter $4(x + y)$?

Solve the problem by using the parametrization

$$\begin{pmatrix} x \\ y \end{pmatrix} = \vec{\gamma}(t) = \begin{pmatrix} 2\cos t \\ \sin t \end{pmatrix},$$

then get the same solution by using Lagrange multipliers.

21. (This problem is equivalent to the derivation of the "Boltzmann factor" $e^{-\frac{E}{kT}}$ that you may have heard of in a chemistry or physics course, but I have rewritten it so that there is no mention of entropy or probability.)

You have taken over from the Postal Service the task of sorting mail in Cambridge. You have 7 million pieces of mail to sort, and you must decide how to divide it among your three "sortation centers." Center 1, an abandoned post office, is rent-free. Center 2, rented from UPS, charges 1 kilobuck for every million pieces of mail that you sort there, Center 3, rented from Harvard, charges 2 kilobucks for every million pieces of mail that you sort there. You are willing to pay 4 kilobucks of rent.

So your constraints are $x_1 + x_2 + x_3 = 7$ and $x_2 + 2x_3 = 4$.

The total effort required to do the sorting is

$$f\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1 \log x_1 + x_2 \log x_2 + x_3 \log x_3 \text{ (natural logarithms).}$$

Using two Lagrange multipliers, find the values of x_1, x_2, x_3 that minimize f while satisfying the two constraints. Hint: if you let $t = \frac{x_2}{x_1}$ you can reduce the problem to solving a quadratic equation for t.

3 Seminar Topics (Dec. 6 or 7)

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. The unit circle is a simple example of a smooth 1-dimensional manifold in $\mathbb{R}^2$. Explain how to describe in it three different ways:

   - As a union of function graphs (the definition).
   - As the locus defined by an equation $F\begin{pmatrix} x \\ y \end{pmatrix} = 0$ for which $[DF]$ is onto at every point on the manifold.
   - As the image of a parametrization function where the parameter is an angle θ.

2. Let M be a k-dimensional submanifold of $\mathbb{R}^n$, and let $\mathbf{z}$ be a point of the manifold, some of whose components are "active" variables $\mathbf{y}$ and the rest of which are "passive" variables $\mathbf{x}$.

   Specify the dimension of the domain and codomain for

   (a) A function $\mathbf{g}$ that specifies the passive variables in terms of the active variables.

   (b) A function $\mathbf{F}$ that specifies $n-k$ constraints that are satisfied by points on the manifold.

   (c) A parametrization function γ that generates points on the manifold from a set of parameters.

   For each alternative, specify the shape of the matrix that represents the derivative of the relevant function and explain how, given a point $\mathbf{c}$ on the manifold, it could be used to find a basis for the tangent space $T_{\mathbf{c}}M$.

3. Suppose that $U \subset \mathbb{R}^n$ is an open subset, $\mathbf{F} : U \to \mathbb{R}^{n-k}$ is a C^1 mapping, and manifold M can be described as the set of points that satisfy $\mathbf{F}(\mathbf{z}) = 0$. Use the implicit function theorem to show that if $[\mathbf{DF}(\mathbf{c})]$ is onto for $\mathbf{c} \in M$, then the tangent space $T_{\mathbf{c}}M$ is the kernel of $[\mathbf{DF}(\mathbf{c})]$. You may assume that the variables have been numbered so that when you row-reduce $[\mathbf{DF}(\mathbf{c})]$, the first $n - k$ columns are pivotal.

4. (Proof 12.1) Let $U \in \mathbb{R}^n$ be an open subset and let $f : U \to \mathbb{R}$ be a C^1 (continuously differentiable) function.
   First prove, using a familiar theorem from single-variable calculus, that if $\mathbf{x}_0 \in U$ is an extremum, then $[\mathbf{D}f(\mathbf{x}_0)] = [0]$.
   Then prove that if $M \subset \mathbb{R}^n$ is a k-dimensional manifold, and $\mathbf{c} \in M \cap U$ is a local extremum of f restricted to M, then $T_{\mathbf{c}}M \subset \ker[\mathbf{D}f(\mathbf{c})]$.

5. (Proof 12.2) Let M be a manifold known by a real-valued C^1 function $F(\mathbf{x}) = 0$, where F goes from an open subset U of $\mathbb{R}^n$ to $\mathbb{R}$ and $[\mathbf{D}F(\mathbf{x})]$ is nowhere zero. Let $f : U \to \mathbb{R}$ be a C^1 function.

   Prove that $\mathbf{c} \in M$ is a critical point of f restricted to M if and only if there exists a Lagrange multiplier λ such that $[\mathbf{D}f(\mathbf{c})] = \lambda[\mathbf{D}F(\mathbf{c})]$.

4 Workshop Problems (Dec. 6 or 7)

If your group has R skills, choose 4a or 4b as your third problem

1. Implicitly defined functions

   (a) The nonlinear equation $\mathbf{F}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x^2+y^2+z^2-3 \\ x^2+z^2-2 \end{pmatrix} = 0$ implicitly determines x and y as a function of z. The first equation describes a sphere of radius 3, the second describes a cylinder of radius 2 whose axis is the y-axis. The intersection is a circle in the plane $y = 1$.

   Near the point $x = 1, y = 1, z = 1$, there is a function that expresses the two passive variables x and y in terms of the active variable z. $\mathbf{g}(z) = \begin{pmatrix} \sqrt{2-z^2} \\ 1 \end{pmatrix}$.

   Calculate $\mathbf{g}'(z)$ and determine the numerical value of $\mathbf{g}'(1)$

   Then get the same answer without using the function $\mathbf{g}$ by forming the Jacobian matrix $[\mathbf{DF}]$ evaluating it at $x = y = z = 1$, and using the implicit function theorem to determine $\mathbf{g}'(z) = -A^{-1}[B]$.

   (b) Dean Smith is working on a budget in which he will allocate x to the library, y to pay raises, and z to the Houses. He is constrained.

   The Library Committee, happy to see anyone get more funds as long as the library does even better, insists that $x^2 - y^2 - z^2 = 1$.

   The Faculty Council, content to see the Houses do well as long as other areas benefit equally, recommends that $x + y - 2z = 1$.

   To comply with these constraints, the dean tries $x = 3, y = 2, z = 2$.

   Given the constraints, x and y are determined by an implicitly defined function $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{g}(z)$.

   Use the implicit function theorem to calculate $\mathbf{g}'(2)$, and use it to find approximate values of x and y if z increased to 2.1.

2. Critical points

   (a) i. Find the one and only critical point of $f\begin{pmatrix} x \\ y \end{pmatrix} = 4x^2 + \frac{1}{2}y^2 + \frac{8}{x^2 y}$ on the square $\frac{1}{4} \leq x \leq 4, \frac{1}{4} \leq y \leq 4$.

   ii. Use second derivatives (the Hessian matrix) to determine whether this critical point is a maximum, minimum, or neither.

   (b) The function $F\begin{pmatrix} x \\ y \end{pmatrix} = x^2 y - 3xy + \frac{1}{2}x^2 + y^2$ has three critical points, two of which lie on the line $x = y$. Find each and use the Hessian matrix to classify it as maximum, minimum, or saddle point.

3. Lagrange Multipliers

   (a) Example with a two-dimensional manifold.

   At what point on the sphere $x^2 + y^2 + z^2 = 7$ does the function xy^2z^4 have a maximum?

   A useful trick with this sort of function is again to take the logarithm – a monotone function, so it has the same critical points.

   $f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \log x + 2\log y + 4\log z$ is the function to be maximized.

   $F\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x^2 + y^2 + z^2 - 7$ is the constraint.

   Set $[\mathbf{D}f] = \lambda[\mathbf{D}F]$ and solve.

   (b) Reversing the function to be optimized and the constraint
   You are building office space and have contracted to supply 10 units of space, half at the end of year 1 and half at the end of year 2. Because large building projects are inefficient, the cost of building x units of space is x^2. However, if you produce more than 5 units in the first year, you can rent out the excess space during the second year at 8 units of money per unit of space.It might be optimal to do something like producing $x = 6$ in the first year, renting out 1 unit for a year, and producing only $y = 4$ in the second year.

   i. Write down the function of x and y to be minimized and the constraint, then use a Lagrange multiplier to find the optimal amount to produce during the first year.

   ii. When you use Lagrange multipliers, the cost function and the constraint are almost interchangeable. Invent a problem that involves maximizing a linear function and that has the same solution as the original problem.

4. Manifolds and tangent spaces, investigated with help from R

   (a) Manifold M is known by the equation
   $F\begin{pmatrix} x \\ y \\ z \end{pmatrix} = xz - y^2 = 0$ near the point $\mathbf{c} = \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix}$.
   It can also be described parametrically by
   $\gamma\begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} s^2 \\ st^2 \\ t^4 \end{pmatrix}$ near $s = 2, t = 1$.

      i. Use the parametrization to find a basis for the tangent space $T_{\mathbf{c}}M$.
      ii. Use the function F to confirm that your basis vectors are indeed in the tangent space $T_{\mathbf{c}}M$.
      iii. Use the parametrization to do a wireframe plot of the parametrized manifold near $s = 2, t = 1$. See script 3.4C, topic 2.

   (b) (Hubbard, Example 3.1.14) $\mathbf{F}\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} z_3 \\ z_3 - z_1 z_2 \end{pmatrix}$
   Construct [**DF**]. It has two rows.
   Find the point for which [**DF**] is not onto. Use R to find points on the manifold near this point, and try to figure out what is going on. See the end of script 3.4C for an example of how to find points on a 1-dimensional manifold in $\mathbb{R}^3$.

5 Homework - due on Tuesday, December 11

Although all of these problems except the last one were designed so that they could be done with pencil and paper, it makes sense to do a lot of them in R, and the Week 12 scripts provide good models. For each problem that you choose to do in R, include a "see my script" reference in the paper version. Put all your R solutions into a single script, and upload it to the homework dropbox on the week 12 page.

When you use R, you will probably want to include some graphs that are not required by the statement of the problem.

1. Pat and Terry are in charge of properties for the world premiere of the student-written opera "Goldfinger" at Dunster House. In the climactic scene the anti-hero takes the large gold brick that he has made by melting down chalices that he stole from the Vatican Museum and places it in a safety deposit box in a Swiss bank while singing the aria "Papal gold, now rest in peace."

   The gold brick is supposed to have length $x = 8$, height $y = 2$, and width $z = 4$. With these dimensions in mind, Pat and Terry have spent their entire budget on 112 square inches of gold foil and 64 cubic inches of an alloy that melts at 70 degrees Celsius. They plan to fabricate the brick by melting the alloy in a microwave oven and casting it in a sand mold.

   Alas, the student mailboxes that they have borrowed to simulate safety-deposit boxes turn out to be not quite 4 inches wide. Fortunately, the equation

$$\mathbf{F}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} xyz - 64 \\ xy + xz + yz - 56 \end{pmatrix} = 0$$

   specifies x and y implicitly in terms of z.

   (a) Use the implicit function theorem to find $[D\mathbf{g}(4)]$, where $\mathbf{g}$ is the function that specifies $\begin{pmatrix} x \\ y \end{pmatrix}$ in terms of z, and find the approximate dimensions of a brick with the same volume and surface area as the original but with a width of only 3.9 inches.

   (b) Show that if the original dimensions had been $x = 2, y = 2, z = 16$, then the constraints of volume 64, surface area 136 specify y and z in terms of x but fail to specify x and y in terms of z.

   (c) Show that if the original brick had been a cube with $x = y = z = 4$, then, with the constraints of volume 64, surface area 96, we cannot show the existence of any implicit function. In fact there is no implicit function, but our theorem does not prove that fact. This happens because this cube has minimum surface area for the given volume.

2. (Physics version)In four-dimensional spacetime, a surface is specified as the intersection of the hypersphere $x^2 + y^2 + z^2 = t^2 - 2$ and the hyperplane $3x + 2y + z - 2t = 2$.

   (Economics version)A resource is consumed at rate t to manufacture goods at rates x, y, and z, and production is constrained by the equation $x^2 + y^2 + z^2 = t^2 - 2$.

   Furthermore, the expense of extracting the resource is met by selling the goods, so that $2t = 3x + 2y + z - 2$.

   In either case, we have a manifold that is the locus of

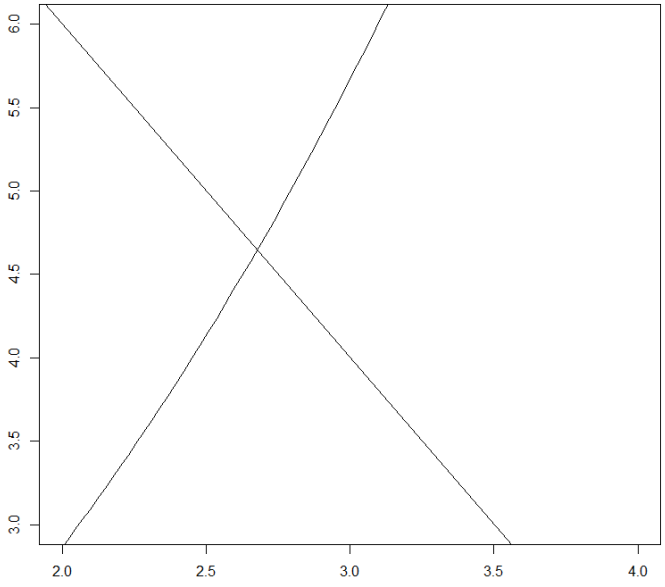
$$\mathbf{F}\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} x^2 + y^2 + z^2 - t^2 + 2 \\ 3x + 2y + z - 2t - 2 \end{pmatrix} = 0.$$

   (a) Show that this surface is a smooth 2-dimensional manifold.

   (b) One point on the manifold is $x = 1, y = 2, z = 3, t = 4$. Near this point the manifold is the graph of a function $\mathbf{g}$ that expresses x and y as functions of z and t. Using the implicit function theorem, determine $[\mathbf{Dg}]$ at the point $z = 3, t = 4$.

3. Consider the manifold specified by the parametrization
   $\mathbf{g}(t) = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} t + e^t \\ t + e^{2t} \end{pmatrix}, -\infty < t < \infty.$

   Find where it intersects the line $2x+y = 10$. You can get an initial estimate by using the graph below (generated in R), then use Newton's method to improve the estimate.



4. Manifold X, a hyperboloid, can be parametrized as

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \gamma \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \sec u \\ \tan u \cos v \\ \tan u \sin v \end{pmatrix}$$

If you use R, you can do a wireframe plot the same way that the sphere was plotted in script 3.4C, topic 2.

   (a) Find the coordinates of the point $\mathbf{c}$ on this manifold for which $u = \frac{\pi}{4}, v = \frac{\pi}{2}$.

   (b) Find the equation of the tangent space $T_{\mathbf{c}}X$ as the image of $[\mathbf{D}\gamma \begin{pmatrix} \frac{\pi}{4} \\ \frac{\pi}{2} \end{pmatrix}]$.

   (c) Find an equation $F \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0$ that describes the same manifold near $\mathbf{c}$, and find the equation of the tangent space $T_{\mathbf{c}}X$ as the kernel of $[\mathbf{DF}(\mathbf{c})]$.

   (d) Find an equation $x = g \begin{pmatrix} y \\ z \end{pmatrix}$ that describes the same manifold near $\mathbf{c}$, and find the equation of the tangent space $T_{\mathbf{c}}X$ as the graph of $[\mathbf{Dg} \begin{pmatrix} 0 \\ 1 \end{pmatrix}]$.

5. Hubbard, Exercise 3.6.2. This is the only problem of this genre on the homework that can be done with pencil and paper, but you must be prepared to do one like it on the final exam!

   The final problem, which requires R, is only for graduate-credit students.

6. Here is another function that has one maximum, one minimum, and two saddle points, for all of which x and y are less than 3 in magnitude.

$$f \begin{pmatrix} x \\ y \end{pmatrix} = x^3 - y^3 + 2xy - 5x + 6y.$$

   Locate and classify all four critical points using R, in the manner of script 3.4D. A good first step is to plot contour lines with x and y ranging from -3 to 3. If you do

```
contour(x,y,z, nlevels = 20)
```

   you will learn enough to start zooming in on all four critical points.

   An alternative, more traditional, approach is to take advantage of the fact that the function f is a polynomial. If you set both partial derivatives equal to zero, you can eliminate either x or y from the resulting equations, then find approximate solutions by plotting a graph of the resulting fourth-degree polynomial in x or y.

1. Given a manifold described in various ways, how can you determine a basis for the tangent space?
   (a) A parametrization?
   (b) A graph-making function?
   (c) A locus function?
   (d) What is the difference between a tangent space and a tangent plane?

2. Tangent spaces and tangent planes
   (a) The equation $F\begin{pmatrix} x \\ y \end{pmatrix} = x^2 + y - 1 = 0$ specifies a one-dimensional manifold in $\mathbb{R}^2$. From this, find a basis for the tangent space to the manifold at the point $(1, 0)$.
   (b) This manifold could also be described by a parametrization
   $$\gamma(t) = \begin{pmatrix} t \\ 1 - t^2 \end{pmatrix}$$
   Using this, find a basis for the tangent space to the manifold at the point $(1, 0)$.
   (c) Using the locus function from part **(a)**, find an equation for the tangent **plane** at this point like Paul did in lecture.
   (d) Do the same using your answer from part **(b)**.

3. Constrained critical points and Lagrange multipliers
   (a) Consider a one-dimensional manifold in $\mathbb{R}^2$ described as the locus of a function $F\begin{pmatrix} x \\ y \end{pmatrix} = 0$. By considering gradients, explain why the constrained critical points of a function $f\begin{pmatrix} x \\ y \end{pmatrix}$ restricted to this manifold will occur at a point where $[Df] = \lambda[DF]$.
   (b) Using Lagrange multipliers, find the constrained critical points of the function $f\begin{pmatrix} x \\ y \end{pmatrix} = y^2 - x$ subject to the constraint $F\begin{pmatrix} x \\ y \end{pmatrix} = x^2 + y^2 - 4 = 0$.
   (c) The manifold in (b) is a circle with radius 2, and it could be equivalently described by the following parametrization, where $-2\pi \leq \theta \leq 0$:
   $$\gamma(\theta) = \begin{pmatrix} 2\cos(\theta) \\ 2\sin(\theta) \end{pmatrix}$$
   Use this parametrization to solve again for the constrained critical points of f.
   As a fun fact to help you here, $4\sin(x)\cos(x) + \sin(x) = 0$ at $x = \arccos\left(\frac{-1}{4}\right)$
   (d) True/False: A constrained critical point of f on a manifold M is always an unconstrained critical point of f as well.
   (e) True/False: If an unconstrained critical point of f occurs at some point c that happens to lie on on a manifold M, then c will be a constrained critical point of f restricted to M as well.