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."

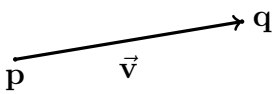
If an element of F^n is a point, we represent it by a bold letter like $\mathbf{p}$ and write it as a column of elements enclosed in parentheses.

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

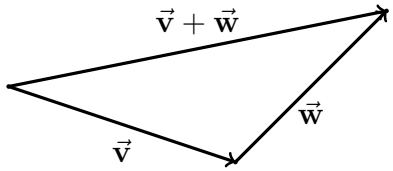
If an element of F^n is a vector, we represent it by a bold letter with an arrow like $\vec{\mathbf{v}}$ and write it as a column of elements enclosed in square brackets.

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

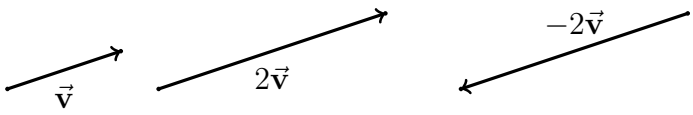
To add a vector to a point, we add the components in identical positions together. The result is a point: $\mathbf{q} = \mathbf{p} + \vec{\mathbf{v}}$. Geometrically we represent this by anchoring the vector at the initial point $\mathbf{p}$. The location of the arrowhead of the vector is the point $\mathbf{q}$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

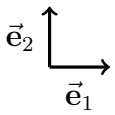


To form a scalar multiple of a vector, we multiply each component by the scalar. In $\mathbb{R}^n$, the geometrical effect is to multiply the length of the vector by the scalar. If the scalar is a negative number, we switch the position of the arrow to the other end of the vector.



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:

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

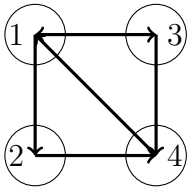
A similar rule holds for matrix inverses:

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

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*

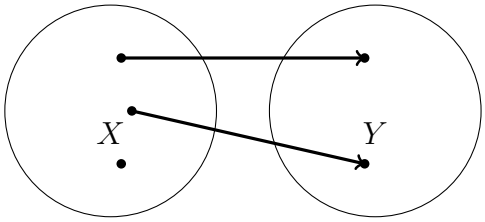
   - The *objects* of the category are finite sets.
   - The *arrows* of the category are functions from one finite set to another. The definition of a function involves quantifiers.

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

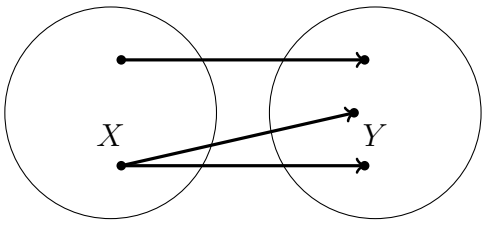
     $\forall x \in X, \exists! y \in Y$ such that $f(x) = y$.

     (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.

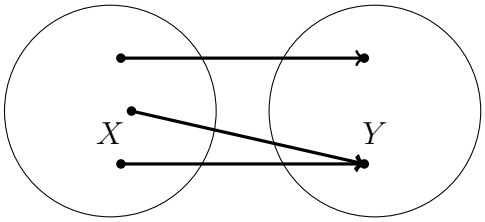
6. Invertible functions - an example of invertible arrows

   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!

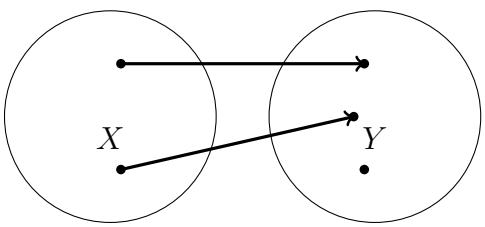
   This function is not invertible because it is not injective, but it is surjective.



   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.

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

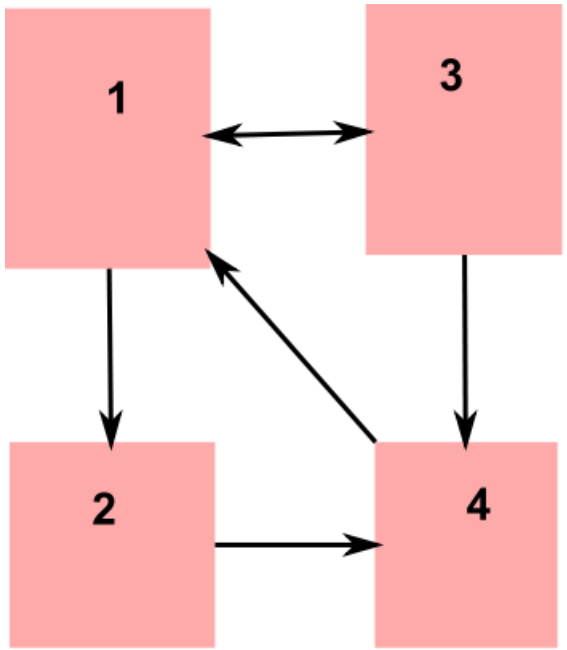
32. Application: graph theory

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.

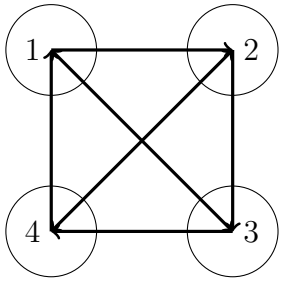
   (b) Suppose that $T : (Z_5)^2 \to (Z_5)^2$ is a linear transformation for which $T\begin{bmatrix}1\\-1\end{bmatrix} = \begin{bmatrix}1\\2\end{bmatrix}, T\begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}3\\0\end{bmatrix}$.

   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:

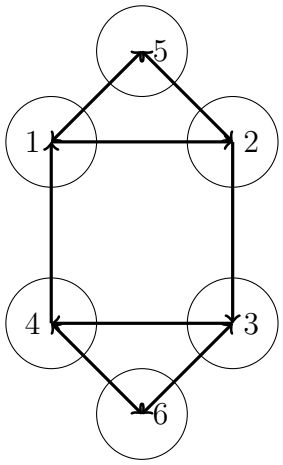
   - If the exam in year n had an Iliad question, the exam in year $n+1$ will have an Iliad question if the die roll is even, an Odyssey question if the die roll is odd.
   - If the exam in year n had an Odyssey question, the exam in year $n+1$ will have an Odyssey question if the die roll is 1, an Iliad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