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 $\vec{\mathbf{v}}_1, \cdots \vec{\mathbf{v}}_{\mathbf{k}}$.

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 $\vec{\mathbf{w}}_2 = T(\vec{\mathbf{v}}_2)$
  $T(a\vec{\mathbf{v}}_1 + b\vec{\mathbf{v}}_2) = aT(\vec{\mathbf{v}}_1) + bT(\vec{\mathbf{v}}_\mathbf{2}) = a\vec{\mathbf{w}}_\mathbf{1} + b\vec{\mathbf{w}}_2$.
  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, $\dim(\text{Ker } T) + \dim(\text{Img } T) = n$.

    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

$$\vec{\mathbf{v}}_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \vec{\mathbf{v}}_2 = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}, \text{ and } \vec{\mathbf{v}}_3 = \begin{bmatrix} 2 \\ 4 \\ 0 \end{bmatrix}$$

   (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.