Contents

Ground rules: On any homework problem, you may use the result of any previous homework problem, any result you know to be true from arithmetic and calculus, and any result which has been proven in class.

1 Induction proofs

1. Prove that for every positive integer n,

$$\sum_{j=0}^{n} j^2 = \frac{n(n+1)(2n+1)}{6}.$$

2. (All variables in this problem represent real numbers.) Assuming only that multiplication distributes over the sum of two real numbers, i.e. $a(b+c) = ab + ac$, prove by induction that multiplication distributes over the sum of any number of real numbers, i.e.

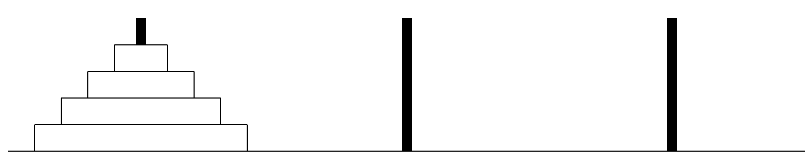
$$a(x_1 + x_2 + ... + x_n) = ax_1 + ... + ax_n.$$

Remark: This illustrates one of the most frequent type of induction proofs, namely extending results known about 2 variables to analogous results about n variables.

3. Prove that every set of n elements has exactly 2^n different subsets.

4. Let f_n be the Fibonacci sequence; that is, let $f_1 = 1$, $f_2 = 1$ and for all $n \geq 3$, $f_n = f_{n-1} + f_{n-2}$. (Thus the first few terms of this sequence are $1, 1, 2, 3, 5, 8, 13, 21, 34, ...$) Find and prove a formula for the sum of the first n Fibonacci numbers.

5. This is the "Tower of Hanoi" problem: suppose you have a collection of n disks, all of different sizes. Each disk has a hole in the middle which allows it to rest on one of three pegs. Initially all of the disks are stacked on top of each other on the first peg, arranged in order by size, with the largest on the bottom and the smallest on the top (see the picture below for $n = 4$). The object is to move the disks to the third peg, where a "move" consists of lifting the top disk from any peg and placing it on a different peg. The catch is that at no time is any disk allowed to rest on top of any smaller disk. Prove that the disks can be moved to the third peg in $2^n - 1$ moves.



6. Let $x_1, ..., x_n$ be real numbers. Prove

$$\left| \sin\left(\sum_{j=1}^{n} x_j \right) \right| \leq \sum_{j=1}^{n} \left| \sin x_j \right|.$$

(You may assume the Triangle Inequality for real numbers, which says that for any real numbers a and b, $|a+b| \leq |a|+|b|$. This "should" have been proven for you when you took calculus. If you can prove the Triangle Inequality, write a proof of that as well.)

7. Prove that for every natural number greater than 5,

$$\frac{n^n}{3^n} < n! < \frac{n^n}{2^n}.$$

Warning: this problem is very hard.

8. The *Strong Form of Mathematical Induction* says: Let $m \in \mathbb{Z}$ and let P_m, P_{m+1}, P_{m+2}, ... be a sequence of statements which each have a truth value. If
   (a) P_m is true, and
   (b) for every integer $k \geq m$, the truth of all the statements $(P_m, P_{m+1}, ..., P_k)$ implies the truth of P_{k+1},

   then P_n is true for every $n \geq m$.

   Use this to prove that every amount of postage greater than or equal to 12 cents can be made by a combination of 4-cent and 5-cent stamps.

2 Complex numbers

1. Let z be a complex number. Prove $\Re(z) = \frac{z+\bar{z}}{2}$ and $\Im(z) = \frac{z-\bar{z}}{2i}$. (Notation: $\Re(z)$ means the real part of z; $\Im(z)$ means the imaginary part of z.)

2. Let z and w be complex numbers. Prove $\overline{z+w} = \overline{z} + \overline{w}$, $\overline{zw} = \overline{z} \cdot \overline{w}$ and $\bar{\bar{z}} = z$.

3. Let z be a complex number. Prove the following statements:
   (a) $|z|^2 = z \cdot \bar{z}$
   (b) $|\bar{z}| = |z|$
   (c) For any other complex number $w \neq 0$, $\frac{z}{w} = \frac{z\,\bar{w}}{|w|^2}$
   (d) $z^{-1} = \frac{\bar{z}}{|z|^2}$
   (e) $|z| = 1$ if and only if $\bar{z} = z^{-1}$

4. Let z and w be complex numbers. Prove the following statements:
   (a) $\Re(z) \leq |\Re(z)| \leq |z|$
   (b) $\Im(z) \leq |\Im(z)| \leq |z|$
   (c) $|zw| = |z|\,|w|$

5. We saw in class that for a complex number z, its complex conjugate $\overline{z}$ can be obtained by reflecting z through the real axis. For each of these numbers, give a similar description of how they can be obtained from z "graphically":

   (a) $\mathrm{Re}(z)$

   (b) $3z$

   (c) $-z$

   (d) iz

   (e) $-\overline{z}$

6. Prove the following statements, where t is a real number:

   (a) $|e^{it}| = 1$

   (b) $\cos t = \frac{e^{it}+e^{-it}}{2}$

   (c) $\sin t = \frac{e^{it}-e^{-it}}{2i}$

7. Prove DeMoivre's Theorem, which says that for all $n \in \mathbb{N}$ and all $\theta \in \mathbb{R}$,

$$(\cos\theta + i\sin\theta)^n = \cos(n\theta) + i\sin(n\theta).$$

8. Find all solutions to the equation $z^8 = 1$ (z is allowed to be any complex number). Graph the solutions on the same plane. *Hint:* DeMoivre's Theorem may be useful here.

9. Prove the *Triangle Inequality*, which says that for any complex numbers z and w, we have

$$|z+w| \le |z| + |w|.$$

10. Prove the *Generalized Triangle Inequality*, which says that for any complex numbers $z_1, z_2, ..., z_n$, we have

$$|z_1 + ... + z_n| \le |z_1| + ... + |z_n|.$$

11. Prove the formula for a geometric sum, which states that for any complex number $z \ne 1$ and any natural number n,

$$\sum_{j=0}^{n} z^j = \frac{1-z^{n+1}}{1-z}.$$

3 Fields

1. Determine whether or not the following sets are fields under the usual operations of addition and multiplication:
   (a) The *complex rationals*, i.e. the set $\{x + iy : x, y \in \mathbb{Q}\}$.
   (b) The *Gaussian integers*, denoted $\mathbb{Z}[i]$, which is the set $\{x + iy : x, y \in \mathbb{Z}\}$.
   (c) $S = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$.
   (d) $T = \{a + b\sqrt{2} + c\sqrt{3} : a, b, c \in \mathbb{Q}\}$.
   (e) $U = \{a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} : a, b, c, d \in \mathbb{Q}\}$.
2. In our definition of field, we stated that there must be an identity element (called 0) for the addition operation. We did not assume in the definition, however, that this identity element was *unique*, that is, that there is only one such identity element for the addition operation. Use the other defining properties of a field to deduce that this element is unique, that is, that if $x \in F$ is such that $x + y = y$ for all $y \in F$, then $x = 0$.
3. Prove the multiplicative identity element of a field is unique.
4. Prove that the additive and multiplicative inverses of elements of a field are unique.
5. Let 0 be the multiplicative identity element of a field F. Prove $0x = 0$ for all $x \in F$.
6. Let F be a field. By definition, what does the symbol -1 refer to? Prove that for all $x \in F$, $-x = -1 \cdot x$.
7. (Challenge) Does there exist a field with 5 elements? Does there exist a field with 6 elements?

4 Vector space axioms

1. Prove that the additive identity element (called $\mathbf{0}$) of a vector space is unique, i.e. that if $\mathbf{x} + \mathbf{y} = \mathbf{y}$ for all $\mathbf{y} \in V$, then $\mathbf{x} = \mathbf{0}$.
2. Prove that if there is even a single $\mathbf{y} \in V$ such that $\mathbf{x} + \mathbf{y} = \mathbf{y}$, then $\mathbf{x} = \mathbf{0}$.
3. Show that the additive inverse of any vector is unique.
4. Let V be a vector space over a field F; prove that for all $\mathbf{v} \in V$ and all $c \in F$:
   (a) $c\mathbf{0} = \mathbf{0}$.
   (b) $0\mathbf{v} = \mathbf{0}$.
   (c) $-\mathbf{v} = (-1)\mathbf{v}$.

(d) $(-c)\mathbf{v} = c(-\mathbf{v}) = -(c\mathbf{v})$.

5. Comment on the following "proof" of part (a) of the previous problem:

$$c\mathbf{0} = c(0, 0, ..., 0) = (c0, ..., c0) = (0, ..., 0) = \mathbf{0}.$$

6. Prove that the following are vector spaces under the "usual" operations (you don't need to write everything out in detail, but you should describe explicitly what the zero element and the additive inverse of a vector look like, and you should verify one or two axioms):

   (a) The set of all functions from $\mathbb{R}$ to $\mathbb{R}$ (called the set of *real functions*). *Challenge:* Generalize this: do you really need the domain and range to be $\mathbb{R}$?

   (b) The set of infinite sequences of real numbers, where addition and scalar multiplication are defined coordinatewise. *Challenge:* Generalize this.

   (c) The set $\mathbb{P}_n$ of polynomials of degree $\leq n$ with real coefficients. P.S. What is the degree of the zero element of this vector space?

   (d) The set of complex numbers $\mathbb{C}$, taken as a vector space over $\mathbb{R}$.

   (e) The set $\{\mathbf{0}\}$, consisting of only the zero vector (over any field).

7. Consider the set of all vectors in $\mathbb{C}^2$ of the form $(z, \bar{z})$, with the usual addition and scalar multiplication. Is this set a vector space over $\mathbb{C}$? Why or why not? Is this set a vector space over $\mathbb{R}$? Why or why not?

8. Prove that the set of real numbers is a vector space (over itself) with the following strange definitions of addition and scalar multiplication (we'll use $\oplus$ and $\otimes$ to represent the new addition and scalar multiplication), and $+$ to represent the usual addition):

$$x \oplus y := x + y + 1; \quad k \otimes x := kx + k - 1$$

9. Show that the set of positive real numbers forms a vector space over $\mathbb{R}$, *if addition and multiplication are defined in a funny way.* Figure out a way to define addition and multiplication to make $(0, \infty)$ into a vector space over $\mathbb{R}$. *Hint:* You can define addition to be what multiplication normally is.

10. Here's another strange (possible) vector space. The vectors are the set of all *nonzero* real numbers, and we define addition (denoted $\oplus$) and scalar multiplication (denoted $\otimes$) by

$$x \oplus y := xy; \quad k \otimes x := \operatorname{sign}(x)|x|^k$$

    where $\operatorname{sign}(x)$ is equal to 1 if $x > 0$ and equal to -1 if $x < 0$. Prove or disprove: this is a vector space.

11. Prove the Cancellation Law (for vector addition), which states that if $\mathbf{x}, \mathbf{y}$ and $\mathbf{z}$ belong to vector space V, then

$$\text{if } \mathbf{x} + \mathbf{y} = \mathbf{x} + \mathbf{z}, \text{ then } \mathbf{y} = \mathbf{z}.$$

12. Consider the power set $2^{\mathbb{Z}}$ consisting of all subsets of integers. Suppose we were to define addition on this set to be set union, i.e $\{1, 2, 3\} + \{2, 3, 5, 8\} = \{1, 2, 3, 5, 8\}$. Show that no matter how scalar multiplication is subsequently defined (and no matter what the underlying field is), this set cannot be made into a vector space.

13. True or false (prove your answer): if V is a vector space over F and $\mathbf{v} \in V$ and $c \in F$ are such that $c\mathbf{v} = \mathbf{0}$, then either $c = 0$ or $\mathbf{v} = \mathbf{0}$.

14. Prove that any vector space over $\mathbb{R}$, other than the vector space $\{\mathbf{0}\}$, has infinitely many vectors in it. *Hint:* First prove that if $\mathbf{x} \neq 0$ and $c \neq d$, then $c\mathbf{x} \neq d\mathbf{x}$.

15. Let U and V be vector spaces over the same field F. Prove $U \times V$ is a vector space over F where the addition and scalar multiplication are defined coordinatewise.

5 Subspaces

1. True or false? Provide enough justification to satisfy yourself:
   (a) The set of polynomials of even degree is a subspace of $\mathbb{P}_n$.
   (b) The $y-$axis is a subspace of $\mathbb{R}^2$.
   (c) The set of functions from $\mathbb{R}$ to $\mathbb{R}$ which are everywhere differentiable is a subspace of the vector space of continuous functions from $\mathbb{R}$ to $\mathbb{R}$.
   (d) The set of vectors with rational coordinates is a subspace of $\mathbb{R}^n$.
   (e) Every vector space is a subspace of itself.
   (f) The set of vectors $W = \{(x, y, z) : x + y + z = 0\}$ is a subspace of $\mathbb{R}^3$.
   (g) The set of vectors $U = \{(x, y, z) : x + y + z = 1\}$ is a subspace of $\mathbb{R}^3$.
   (h) The set of polynomials satisfying $p(2) = 0$ is a subspace of $\mathbb{P}_n$.
   (i) The set of (twice-differentiable) real functions satisfying $3f''(x)+x^2f'(x)-f(x) = 0$ is a subspace of the set of all twice-differentiable real functions.

2. Let W_1 and W_2 be subspaces of a vector space V. Prove that $W_1 \cap W_2$ is also a subspace.

3. Let $W_1, ..., W_n$ be subspaces of a vector space V. Prove that $W = \bigcap_{j=1}^{n} W_j$ is also a subspace.

4. Prove (by an example) that the union of two subspaces of a vector space need not be a subspace.

5. Let W_1 and W_2 be subsets of a vector space V. Define the *sum* of these subsets to be

$$W_1 + W_2 = \{\mathbf{w}_1 + \mathbf{w}_2 : \mathbf{w}_1 \in W_1, \mathbf{w}_2 \in W_2\}.$$

(a) Let $V = \mathbb{R}^3$, let X be the $x-$axis, and let Y be the $y-$axis. What is $X + Y$?

(b) Let $V = \mathbb{R}^4$ (where the coordinate directions are called x, y, z and w), let U_1 be the $xy-$plane, and let U_2 be the $yz-$plane. Describe $U_1 + U_2$.

(c) Why in the definition above do we require that W_1 and W_2 are subsets of the same vector space?

(d) Prove or disprove: $W_1 + W_2 = W_2 + W_1$.

(e) Prove that if W_1 and W_2 are subspaces (not just subsets), then $W_1 + W_2$ is a subspace.

(f) Let W be a subspace of V. Describe $W + W$, $W + \{\mathbf{0}\}$, and $W + V$.

(g) Prove or disprove: if U, W_1 and W_2 are all subspaces of V such that $U + W_1 = U + W_2$, then $W_1 = W_2$.

6. Let V be a vector space over a field F and for any $n \geq 1$, let $\mathbf{v}_1, ..., \mathbf{v}_n$ be any elements of V. Let W be the *span* of these vectors, that is, let

$$W = Span(\mathbf{v}_1, ..., \mathbf{v}_n) = \left\{ \sum_{j=1}^{n} c_j \mathbf{v}_j : c_j \in F \right\}.$$

Prove that W is a subspace of V. *Note:* W is also denoted $< \mathbf{v}_1, ..., \mathbf{v}_n >$.

7. Describe in words the following subsets of $\mathbb{R}^3$ (the first one is "the $x-$axis"):

(a) $W =< (1, 0, 0) >$.

(b) $W =< (1, 2, 3) >$.

(c) $W =< (1, 0, 0), (0, 1, 0) >$.

(d) $W =< (1, 0, 0), (0, 1, 0), (1, 4, 0) >$.

(e) $W =< (1, 0, 0), (0, 1, 0), (0, 0, 1) >$.

8. Prove or disprove: let $\mathbf{v}$ and $\mathbf{w}$ be vectors in some vector space V. Then

$$< \mathbf{v} > + < \mathbf{w} >=< \mathbf{v}, \mathbf{w} > .$$

9. The functions in the following set are solutions to $f''(x) = -f(x)$:

$$S = \{A \sin x + B \cos x : A, B \in \mathbb{R}\}.$$

But in physics, one usually argues that the following are solutions:

$$T = \{A \sin(x + c) : A, c \in \mathbb{R}\}.$$

The argument goes like this: the differential equation $f'' = -f$ is a special case of Hooke's Law, with a fixed spring constant $k = 1$. All solutions to Hooke's Law are sinusoidal; since k is fixed, so is the frequency of the solution, but the amplitude A and phase shift c are arbitrary.

(a) One of S and T is clearly a vector space (by problem 5.6). Which one?

(b) Determine the relationship between S and T (for example, is one contained in the other?)

6 Affine subspaces, lines and planes

1. Let W be a subset of a vector space V. Define the *translation* of W by the vector $\mathbf{p} \in V$ to be

$$\mathbf{p} + W = \{\mathbf{p} + \mathbf{w} : \mathbf{w} \in W\};$$

observe that $\mathbf{p} + W = \{\mathbf{p}\} + W$ as defined in problem 5.5, so by problem 5.5 (e), $\mathbf{p} + W$ and $W + \mathbf{p}$ are the same thing. Also, for any scalar $k \in F$, define

$$kW = \{k\mathbf{w} : \mathbf{w} \in W\}.$$

Prove or disprove the following (where capital letters represent arbitrary subsets of a vector space and lowercase letters represent arbitrary scalars):

(a) $\mathbf{0} + A = A$

(b) $A - A = \{\mathbf{0}\}$

(c) $(a + b)W = aW + bW$

(d) $(ab)S = a(bS)$

(e) $a(\mathbf{p} + S) = a\mathbf{p} + aS$

(f) $1A = A$

(g) $0A = \{\mathbf{0}\}$

Remark: This problem is one of the first where you are asked to prove the equality of two sets. To prove two sets (say Q and R) are equal, there are two typical ways to proceed:

- Prove two things: first, that $Q \subseteq R$, and second, that $R \subseteq Q$ (these two things together imply $Q = R$). To prove $Q \subseteq R$, start by assuming x is some arbitrary element of Q and try to conclude by some argument that $x \in R$. To prove $R \subseteq Q$, start by assuming x is some arbitrary element of R and try to conclude by some argument that $x \in Q$.
- Start with the statement $x \in Q$ and end with the statement $x \in R$, making sure all your steps are reversible. This works in some parts of this problem, but won't always work.

2. Let $V = \mathbb{R}^2$; let X be the $x-$axis, Y be the $y-$axis, and C be the unit circle (centered at the origin). Describe the following sets:

(a) $5X$

(b) $2C$

(c) $-Y$

(d) $Y + C$

(e) $(0, 1) + X$

(f) $5((0, 1) + X)$

(g) $C + 2C$

3. Suppose W is a subspace of vector space V and k is a nonzero scalar. Prove $kW = W$.

4. Let V be a vector space. Define an *affine subspace* of V to be a subset $A \subseteq V$ which is a translate of a subspace of V.

   (a) Let V be a vector space over $\mathbb{R}$, and let A be an affine subspace of V. Prove that given any two vectors $\mathbf{p}$ and $\mathbf{q}$ in A, $(t\mathbf{p} + (1-t)\mathbf{q})$ is also in A for all $t \in \mathbb{R}$.

   (b) Suppose W_1 and W_2 are subspaces of vector space V and suppose also that $\mathbf{p}_1$ and $\mathbf{p}_2$ are vectors such that
   $$\mathbf{p}_1 + W_1 = \mathbf{p}_2 + W_2.$$
   Prove $W_1 = W_2$.

   (c) In the context of part (b), is it necessarily true that $\mathbf{p}_1 = \mathbf{p}_2$?

   (d) Show by example that the fact proven in part (b) fails if W_1 and W_2 are assumed only to be subsets of V rather than subspaces.

   (e) Show that given any vector $\mathbf{v} \in A$, $A - \mathbf{v}$ is a subspace of V (and by part (b), must be the same subspace no matter the choice of $\mathbf{v}$).

   (f) Prove the converse of (a), i.e. if A is a nonempty subset of V such that given any two vectors $\mathbf{p}$ and $\mathbf{q}$ in A, $(t\mathbf{p} + (1-t)\mathbf{q})$ is also in A for all $t \in \mathbb{R}$, then A is an affine subspace of V.

   (g) Explain why the facts illustrated in parts (a) and (f) give a nice geometric characterization of affine subspaces.

5. Give an example of two parallel vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^4$ such that there is no $c \in \mathbb{R}$ such that $\mathbf{x} = c\mathbf{y}$.

6. Let $V = \mathbb{R}^3$. Consider the following vectors (or points) in V:
   $$\mathbf{v} = (2, 1, 2), \mathbf{w} = (0, 3, 4), \mathbf{x} = (-2, 3, 0).$$

   (a) Find parametric equations of the plane passing through these three points.

   (b) Find parametric equations of the line passing through $\mathbf{v}$ and $\mathbf{w}$.

7. Here are the parametric equations for two lines in $\mathbb{R}^3$:
   $$\begin{cases} x_1 = 2 + 3t \\ x_2 = 1 - t \\ x_3 = 4 + 7t \end{cases} ; \quad \begin{cases} x_1 = 3 - 7t \\ x_2 = -2 + 5t \\ x_3 = -6 - 4t \end{cases}$$

Prove that these two lines intersect in a point. Find the coordinates of this point. *Hint:* If two lines intersect, they must meet at the same (x_1, x_2, x_3), but the t doesn't have to be the same for both lines (think about why this is).

8. Here are the parametric equations for two lines in $\mathbb{R}^3$:

$$\begin{cases} x_1 = 3t \\ x_2 = 2 - t \\ x_3 = -1 + t \end{cases} ; \quad \begin{cases} x_1 = 1 + 4t \\ x_2 = -2 + t \\ x_3 = -3 - 3t \end{cases}$$

Prove that these two lines do not intersect.

9. Here are two sets of parametric equations:

$$\begin{cases} x_1 = 2t \\ x_2 = 2 - t \\ x_3 = -1 + 4t \end{cases} ; \quad \begin{cases} x_1 = 2 - 6t \\ x_2 = 1 + 3t \\ x_3 = 3 - 12t \end{cases}$$

Prove that these are parametric equations for the same line in $\mathbb{R}^3$.

10. Two airplanes fly along straight lines. At time t, plane 1 is at $(75, 50, 25) + t(5, 10, 1)$ and plane 2 is at $(60, 80, 34) + t(10, 5-1)$. Do the flight paths of these planes intersect? Do the planes crash into one another?

7 Linear independence, span, basis and dimension

7.1 Linear independence and linear dependence

1. Given each of the following vector spaces V and lists $\mathcal{S}$ of vectors , determine whether or not $\mathcal{S}$ is a linearly independent set.

   (a) $V = \mathbb{R}^2$; $\mathcal{S} = \{(-3, 7), (4, -10)\}$.
   (b) $V = \mathbb{R}^4$; $\mathcal{S} = \{(1, 1, 1, 1), (2, 2, 2, 2), (1, 3, 5, 8), (0, -2, 5, 7)\}$.
   (c) $V = \mathbb{R}^3$; $\mathcal{S} = \{(1, 2, 7), (-1, 7, 5), (10, -3, 6), (8, 4, -1)\}$.
   Remark: An easy solution to this problem requires concepts not yet developed. You could solve this now by solving a complicated system of equations, but that's hard.
   (d) $V = \mathbb{R}^7$; $\mathcal{S} = \{(2, 3, 8, -5, 0, 0, 0)\}$.
   (e) $V = \mathbb{R}^4$; $\mathcal{S} = \{(1, 2, 3, 4), (5, 6, 7, 8), (0, 0, 0, 0)\}$
   (f) $V = \mathbb{R}^3$; $\mathcal{S} = \{(0, 1, 4), (0, 0, 1), (0, 7, -3)\}$.
   (g) $V = \mathbb{R}^3$; $\mathcal{S} = \{(0, 1, 2), (0, -3, -6), (2, 5, -8)\}$.
   (h) $V = \mathbb{C}^2$ (taken as a vector space over $\mathbb{C}$); $\mathcal{S} = \{(1, i), (i, -1)\}$

(i) $V = \mathbb{C}^2$ (taken as a vector space over $\mathbb{R}$); $\mathcal{S} = \{(1, i), (i, -1)\}$

2. Let $\mathbf{v}_1, ..., \mathbf{v}_n$ be vectors in some vector space. Prove that if $\sum_{j=1}^n c_j \mathbf{v}_j = \sum_{j=1}^n d_j \mathbf{v}_j$ for scalars $c_1, ..., c_n$ and $d_1, ..., d_n$ where the $c_j \neq d_j$ for some j, then $\{\mathbf{v}_1, ..., \mathbf{v}_n\}$ is linearly dependent.

   Note: This problem tells you that if you can write a vector $\mathbf{w}$ as a linear combination of the $\mathbf{v}_j$ in two different ways, then $\{\mathbf{v}_1, ..., \mathbf{v}_n\}$ is linearly dependent.

   Equivalently (remember this), it says that given a lin. indep. set S and given any vector $\mathbf{x}$, there is **at most one way** to write $\mathbf{x}$ as a linear combination of vectors in S.

3. Let $V = C^\infty(\mathbb{R}, \mathbb{R})$ be the vector space of functions from $\mathbb{R}$ to $\mathbb{R}$ which are differentiable infinitely many times (this is a vector space over $\mathbb{R}$ since it is a nonempty subset of $\mathbb{R}^\mathbb{R}$ which is closed under addition and scalar multiplication).

   (a) Is the set of functions $\{\sin^2 x, \cos^2 x, 1\}$ linearly independent? Why or why not? (Check your answer before you proceed to part (b).)

   (b) Is the set of functions $\{e^x, e^{-x}, 1\}$ linearly independent? *Note:* Just because we aren't aware of an identity of the form $c_1 e^x + c_2 e^{-x} = 1$ doesn't mean one doesn't exist. To prove that no such identity exists, we will carry out the following steps:

      i. Suppose we write 0 (the constant function 0) as a linear combination of $\{e^x, e^{-x}, 1\}$:
      $$c_1 e^x + c_2 e^{-x} + c_3 1 = 0 \qquad (1)$$
      Show that it must be the case that $c_1 + c_2 + c_3 = 0$. *Hint:* if the above equation holds as an equality between functions, then it is supposed to hold for all particular values of x.

      ii. Differentiate both sides of equation (1) and subsequently show that $c_1 - c_2 = 0$.

      iii. Differentiate both sides of the equation again, and show that $c_1 + c_2 = 0$.

      iv. Explain why $c_1 = c_2 = c_3 = 0$ (and thus why the three functions $\{e^x, e^{-x}, 1\}$ are linearly independent).

   (c) Suppose $f, g \in V$ are functions such that $f(17) = 2, f'(17) = 1, g(17) = 3$ and $g'(17) = 2$. Are f and g linearly independent? Why or why not?

4. Prove that the set of functions $\{\sin x, \sin 2x, \sin 3x\}$ is linearly independent (in $C^\infty(\mathbb{R}, \mathbb{R}$, taken as a vector space over $\mathbb{R}$).

5. True or False? If the sets $\{\mathbf{u}, \mathbf{v}\}$, $\{\mathbf{u}, \mathbf{w}\}$ and $\{\mathbf{v}, \mathbf{w}\}$ are all linearly independent, then the set $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is linearly independent. Justify your answer.

7.2 More on spans

6. Consider the set W of functions from $\mathbb{R}$ to $\mathbb{R}$ satisfying the differential equation $f''(x) - 5f'(x) + 6f(x) = 0$. W is a subspace of $C^\infty(\mathbb{R}, \mathbb{R})$ (the proof is similar to the one in Problem 5.1 (i)).

(a) Verify that $f(x) = e^{3x}$ and $f(x) = e^{2x}$ both belong to W.

(b) Based on the fact that W is a subspace, give a description of a large collection of functions which must belong to W.

(c) Based on the description in part (b), find a solution $f(x)$ to the differential equation $f''(x) - 5f'(x) + 6f(x) = 0$ satisfying $f(0) = 1$ and $f(1) = 4$. *Important:* Notice that the ideas of linear algebra reduce the problem of solving this differential equation to solving linear equations.

7. Let F be a field and let $V = F$ (i.e. we are taking V to be the field, taken as a vector space over itself). Let $\{\mathbf{x}_1, \mathbf{x}_2, ...\mathbf{x}_n\}$ be a collection of vectors in V; describe what the span of these vectors is. (*Hint:* there are only two possible cases.) Describe under what circumstances this collection of vectors forms a linearly independent set.

8. Prove that two vectors (a, b) and (c, d) span $\mathbb{R}^2$ if and only if $ad - bc \neq 0$.

9. Prove or disprove: If S and S' are two finite sets of vectors in some v.s. V, then $Span(S \cap S') = Span(S) \cap Span(S')$.

10. Prove or disprove: If S and S' are two finite sets of vectors in some v.s. V, then $Span(S \cup S') = Span(S) \cup Span(S')$.

11. Prove or disprove: If S and S' are two finite sets of vectors in some v.s. V, then $Span(S \cup S') = Span(S) + Span(S')$.

7.3 Basis and dimension

12. Given each of the vector spaces (or subspaces) W, find a basis of W, and the dimension of W.

(a) $W = \mathbb{C}$, taken as a vector space over $\mathbb{R}$.

(b) $W = \mathbb{C}$, taken as a vector space over itself.

(c) $W = M_2(\mathbb{R})$ (2×2 matrices with real entries).

(d) $W = \mathbb{P}_3$ (this is the set of polynomials with real coefficients of degree ≤ 3).

(e) $W = \{f \in \mathbb{P}_2 : f(3) = 0\}$

(f) $W = \{\mathbf{w} = (w_1, w_2, w_3) \in \mathbb{R}^3 : 2w_1 + w_2 - 3w_3 = 0\}$.

(g) $W =$ the line in $\mathbb{R}^4$ with parametric equations $x_1 = 4t, x_2 = -3t, x_3 = 2t, x_4 = 0$.

(h) $V = C^1(\mathbb{R}, \mathbb{R})$; W is the subspace of V consisting of functions f satisfying the differential equation $f'(x) = f(x)$.

(i) $V = \mathbb{R}^3$; $W = \left\{ \begin{pmatrix} 2s - 5t \\ s \\ 4t \end{pmatrix} : s, t \in \mathbb{R} \right\}$

(j) V is the strange vector space from Problem 4.9 (i.e. $V = (0, \infty)$ where addition is $x \oplus y = xy$ and $c \otimes x = x^c$).

(k) V is the even stranger vector space of Problem 4.8.

13. Prove the following statements, where V is a vector space with $\dim(V) = n < \infty$:
    (a) If $V \neq \{\mathbf{0}\}$, then V has a basis.
    (b) Any linearly independent set of n vectors is a basis of V. *Hint:* What can be done to any lin. indep. set of vectors in a finite-dimensional vector space?
    (c) Any set of n vectors which span V must be a basis of V.
    (d) If W is a subspace of V, then $\dim(W) \leq \dim(V)$.
    (e) If W is a subspace of V and $\dim(W) = \dim(V)$, then $W = V$. **Warning:** This is not true if $\dim(V) = \infty$.

14. Show that the only affine subspaces of $\mathbb{R}^2$ are single points, lines, and all of $\mathbb{R}^2$. *Hint:* Start by characterizing all the subspaces of $\mathbb{R}^2$. What are the possible dimensions of such a subspace? What is true about a subspace if its dimension is zero? What if its dimension is one? Etc.

15. Characterize all the affine subspaces of $\mathbb{R}^3$; the answer should be "they are points or lines or ..." (start by characterizing all the subspaces of $\mathbb{R}^3$).

16. Show that any subset of $\mathbb{R}^3$ of the form

$$\{(x, y, z) \in \mathbb{R}^3 : ax + by + cz = d\}$$

(where a, b and c are not all zero) is a plane. *Hint:* Start by showing that any subset of the form $\{(x, y, z) \in \mathbb{R}^3 : ax + by + cz = 0\}$ is a plane.

Remark: We will show in a future homework problem that every plane in $\mathbb{R}^3$ can be described by such an equation.

17. Earlier this semester one of the homework questions was to determine whether the following set

$$T = \{a + b\sqrt{2} + c\sqrt{3} : a, b, c \in \mathbb{Q}\}$$

was a field (under the usual operations of $+$ and $\cdot$). You folks correctly told me that T is not a field because it is not closed under multiplication... $\sqrt{2} \cdot \sqrt{3} = \sqrt{6}$ which is not in T, i.e. there is no solution to the equation

$$a + b\sqrt{2} + c\sqrt{3} = \sqrt{6} \tag{2}$$

where a, b, c are rational. We can restate the context of this problem using the language of vector spaces: consider the set

$$V = \{a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} : a, b, c, d \in \mathbb{Q}\};$$

this is a vector space over $\mathbb{Q}$ since $V = Span(1, \sqrt{2}, \sqrt{3}, \sqrt{6})$.

   (a) Restate the assertion "$\sqrt{6}$ is not in T" using the language of linear algebra (use a term or terms like "linear independence", "span", "basis", etc.)

(b) In fact, the following statement is also true:

$$a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} = 0 \ (a, b, c, d \in \mathbb{Q}) \Rightarrow a = b = c = d = 0.$$

Restate this claim in linear algebra language.

(c) Show that the assertion "$\sqrt{6}$ is not in T" follows from the statement in part (b).

(d) On the other hand, equation (2) above has lots of solutions a, b, c if a, b and c are allowed to be real numbers. Give three different triples (a, b, c) of real numbers which satisfy (2).

(e) Consider the set

$$Span(1, \sqrt{2}, \sqrt{3}) = \{a + b\sqrt{2} + c\sqrt{3} : a, b, c \in \mathbb{R}\}.$$

(Here we are taking the span over the reals rather than over the rationals.) What is another name for this set?

8 Matrix theory

8.1 Matrix arithmetic

1. Let F be a field. Prove that for fixed m and n, $M_{mn}(F)$ is a vector space over F (i.e. verify all the axioms).

2. Let $A = \begin{pmatrix} i & 1+i \\ 0 & 2 \end{pmatrix}$; $B_{2\times 1} = \begin{pmatrix} 2-i & 3i \end{pmatrix}$ and $C = \begin{pmatrix} i & 1+i & 0 \\ 0 & 2 & 1+i \end{pmatrix}$. Compute the following or explain that they do not exist:

$$C^2 \qquad AC \qquad CA \qquad (AB^T)^T \qquad BA \qquad A^2 \qquad \overline{C}B^H$$

3. Prove the associative law of matrix multiplication, which says that for any matrices A, B and C, if $(AB)C$ exists, then so does $A(BC)$ and in this case, $A(BC) = (AB)C$.

4. Prove that "scalars pass through matrix multiplication", which says that for any $A \in M_{mn}(F)$, $B \in M_{np}(F)$ and $k \in F$, we have $k(AB) = (kA)B = A(kB)$.

5. The commutative property $AB = BA$ cannot hold for all matrices, because it can be the case where BA exists, but AB does not exist. Even when both AB and BA exist, however, they might be different sizes (hence unequal). Even worse, if AB and BA both exist and are the same size, it still is not necessarily the case that $AB = BA$. Prove this by finding a specific example of an A and a B where AB and BA both exist, are the same size, but are different matrices. What is the smallest possible size of such an A and B?

6. Let $A \in M_{mn}(F)$ and $B, C \in M_{np}(F)$. Prove $A(B + C) = AB + AC$.

7. Let $A, B \in M_{mn}(F)$ and $C \in M_{np}(F)$. Prove $(A + B)C = AC + AC$.

8. If A and B are matrices where AB and BA are both defined and we have $AB = BA$, we say A and B *commute*.

   (a) If A and B commute, what must be true about the sizes of A and B?

   (b) Suppose A commutes with B and C. Is it necessarily the case that A commutes with $B + C$? Is it necessarily the case that A commutes with B^2? Does A commute with kB for any scalar k? Does A commute with BC?

   (c) Supose A and B commute. Prove that A^2 and B^2 commute.

   (d) Prove or disprove: if A and B commute, then for all nonnegative integers p and q, the matrices A^p and B^q commute.

9. True or false: let $A, B \in M_n(F)$. Then $A^2 - B^2 = (A + B)(A - B)$.

10. In this problem, for each k let I_k be the $k \times k$ identity matrix. Prove that if $A \in M_{mn}(F)$, then
$$I_m A = A \quad \text{and} \quad AI_n = A.$$

11. Prove, for matrices $A, B \in M_{mn}(F)$, $C \in M_{np}(F)$, $D \in M_{nm}(F)$ and scalar $r \in F$:

    (a) $(A^T)^T = A$ and $(A^H)^H = A$;

    (b) $(A^T)^H = (A^H)^T = \overline{A}$;

    (c) If A is square, $tr(A^T) = tr(A)$ and $tr(A^H) = tr(\overline{A}) = \overline{tr(A)}$;

    (d) $(rA)^T = rA^T$ and $(rA)^H = \overline{r}A^H$;

    (e) $(A + B)^T = A^T + B^T$ and $(A + B)^H = A^H + B^H$;

    (f) $(AC)^T = C^T A^T$ and $(AC)^H = C^H A^H$;

    (g) $tr(A + B) = trA + trB$;

    (h) $tr(AD) = tr(DA)$.

12. Regardless of the size of a matrix A, the products AA^T and $A^T A$ both exist. What are the sizes of these matrices, in terms of the size of A?

8.2 Invertibility of matrices

13. We say that a matrix $A \in M_n(F)$ is *invertible* if there exists another matrix $B \in M_n(F)$ such that $AB = I$ (where I is the $n \times n$ identity matrix). Prove that for the same matrix B, $BA = I$ (note, we only assume $AB = I$).

    Remark: If A is invertible, we call the matrix B of this problem the *inverse* of A and denote it A^{-1}. In particular, if A is invertible, this problem shows that there is a matrix A^{-1} such that $A^{-1}A = AA^{-1} = I$.

14. Prove that the inverse of an invertible matrix is unique.

15. Suppose $A, B \in M_n(F)$ are both invertible. Prove AB is invertible and find an explicit formula for its inverse.

16. Suppose $A \in M_n(F)$ is invertible. Prove A^T is invertible and $(A^T)^{-1} = (A^{-1})^T$; prove $\overline{A}$ is invertible and $\left(\overline{A}\right)^{-1} = \overline{A^{-1}}$; prove A^H is invertible and $(A^H)^{-1} = (A^{-1})^H$. Prove that for any nonnegative integer n, A^n is invertible; what is the inverse of A^n? Prove also that A^{-1} is invertible; what is the inverse of A^{-1}?

17. Prove the following statements (assume the matrices are of the appropriate size so that everything is defined):

    (a) $(A_1 + A_2 + ... + A_n)^T = A_1^T + ... + A_n^T$;
    (b) $(A_1 + A_2 + ... + A_n)^H = A_1^H + ... + A_n^H$;
    (c) $(A_1 A_2 \cdots A_n)^T = A_n^T \cdots A_1^T$;
    (d) $(A_1 A_2 \cdots A_n)^H = A_n^H \cdots A_1^H$;
    (e) If the A_j are all invertible, then $(A_1 A_2 \cdots A_n)^{-1} = A_n^{-1} \cdots A_1^{-1}$

18. Show that a 2×2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is invertible if and only if $ad - bc \neq 0$, in which case

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

    Remember this formula for the inverse of a 2×2 matrix.

19. Show by an explicit counterexample (without using the zero matrix as any of your counterexamples) that each of the following statements is false:

    (a) If A, B and C are matrices such that the products AB and AC both exist and $AB = AC$, then $B = C$.
    (b) If A, B are matrices such that $AB = 0$ (the zero matrix), then $A = 0$ or $B = 0$.
    (c) If $A \in M_n(F)$ is not the zero matrix, then A is invertible.
    (d) If $A, X \in M_n(F)$ are such that $AX = A$, then $X = I$.

20. Prove the *Cancellation Law* for matrices, which says that if $A \in M_n(F)$ is an invertible matrix, then:

    (a) If $AB = AC$, then $B = C$; and
    (b) if $BA = CA$, then $B = C$.

    Note: This does not mean that if $AB = CA$, then $B = C$. Why not?

21. Suppose A, B and X are $n \times n$ matrices such that A, X and $A - AX$ are invertible. If $(A - AX)^{-1} = X^{-1}B$, find X in terms of A and B.

22. Suppose A and B are invertible matrices of the same size. Show $A^{-1} + B^{-1} = A^{-1}(A+B)B^{-1}$. What fact from high-school (or earlier) algebra does this generalize?

8.3 Special matrices

23. A matrix $A \in M_n(F)$ is called *diagonal* if $A_{jk} = 0$ whenever $j \neq k$ (i.e. all the entries of A off its diagonal are zero). Prove that the set of diagonal $n \times n$ matrices is a subspace of $M_n(F)$; what is its dimension?

24. Prove that any two diagonal matrices (of the same size) commute.

25. A matrix A is called *symmetric* if $A = A^T$. Explain why every symmetric matrix must be square. Now, denote by $Sym_n(F)$ the set of symmetric $n \times n$ matrices with entries in the field F. Prove $Sym_n(F)$ is a subspace of $M_n(F)$. What is the dimension of $Sym_n(F)$ (as a vector space over F)?

26. Let $A \in M_n(F)$. Prove AA^T and A^TA are symmetric.

27. A matrix $A \in M_n(\mathbb{C})$ is called *Hermitian* if $A = A^H$. Prove that the set of Hermitian $n \times n$ matrices with entries in $\mathbb{C}$, called $H_n(\mathbb{C})$ or just H_n, is not a subspace of $M_n(\mathbb{C})$ if the underlying field is $\mathbb{C}$, but is a subspace of $M_n(\mathbb{C})$ if the underlying field is $\mathbb{R}$. *Hint:* What must be true about all the diagonal entries of a Hermitian matrix?

28. Prove that if $A \in Sym_n(F)$, then $A^2 \in Sym_n(F)$. Prove or disprove: if $A, B \in Sym_n(F)$, then $AB \in Sym_n(F)$. Prove or disprove: if $A, B \in Sym_n(F)$, then $A+B \in Sym_n(F)$.

29. Let $A \in M_n(F)$. Prove AA^H and A^HA are Hermitian.

30. A matrix A is called *skew symmetric* if $A = -A^T$. The set of skew-symmetric $n \times n$ matrices is called $Skew_n(F)$; this is a subspace of $M_n(F)$ (you don't need to prove this). What is its dimension (as a vector space over F?

31. Suppose $A \in Skew_n(F)$. What kind of matrix is A^2?

32. Prove every square matrix is the sum of a symmetric matrix and a skew-symmetric matrix. *Hint:* You can give a slick proof of this using dimensions of the spaces. What is $Sym_n(F) \bigcap Skew_n(F)$?

8.4 More insight into matrix multiplication

33. Let $B \in M_{3n}(F)$. This means B has three rows and each row of B, by itself, constitutes a vector in F^n. Let these rows (from top to bottom) be $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ (these $\mathbf{r}_j$ are vectors in F^n, not numbers). We write this by saying

$$B = \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{r}_3 \end{pmatrix}.$$

(a) Is there a matrix A such that AB has three rows, and the rows of AB are exactly

$$\begin{pmatrix} -2\mathbf{r}_2 \\ \mathbf{r}_2 + 3\mathbf{r}_3 \\ 4\mathbf{r}_1 - 7\mathbf{r}_2 + 5\mathbf{r}_3 \end{pmatrix}?$$

If so, find A. If not, explain why not.

(b) Is there a matrix C such that BC has three rows, and the rows of BC are exactly

$$\begin{pmatrix} -2\mathbf{r}_2 \\ \mathbf{r}_2 + 3\mathbf{r}_3 \\ 4\mathbf{r}_1 - 7\mathbf{r}_2 + 5\mathbf{r}_3 \end{pmatrix}?$$

If so, find C. If not, explain why not.

34. Let $B \in M_{m4}(F)$. This means each column of B, by itself, constitutes a vector in F^m. Let these columns (from left to right) be $\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4$ (these are vectors, not numbers). We write this by saying

$$B = \begin{pmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{c}_3 & \mathbf{c}_4 \end{pmatrix}.$$

(a) Is there a matrix A such that AB has four columns, and the columns of AB are exactly

$$\begin{pmatrix} -2\mathbf{c}_2 & \mathbf{c}_2 + 3\mathbf{c}_3 & 4\mathbf{c}_1 - 7\mathbf{c}_2 + 5\mathbf{c}_3 & -\mathbf{c}_3 + \mathbf{c}_4 \end{pmatrix}?$$

If so, find A. If not, explain why not.

(b) Is there a matrix C such that BC has four columns, and the rows of BC are exactly

$$\begin{pmatrix} -2\mathbf{c}_2 & \mathbf{c}_2 + 3\mathbf{c}_3 & 4\mathbf{c}_1 - 7\mathbf{c}_2 + 5\mathbf{c}_3 & -\mathbf{c}_3 + \mathbf{c}_4 \end{pmatrix}?$$

If so, find C. If not, explain why not.

35. Let

$$E_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \in M_3(F).$$

Let $A \in M_3(F)$. What is the relationship between A and $E_{12}A$? What is the relationship between A and AE_{12}? Generalize this.

36. Suppose mystery matrix B satisfies

$$B\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \quad \text{and} \quad B\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ -2 \end{pmatrix}.$$

(a) (Quickly) find X such that

$$BX = \begin{pmatrix} 5 & 4 \\ -2 & 6 \end{pmatrix}.$$

(b) (Quickly) find B.

37. Suppose $C \in M_2(\mathbb{R})$ is such that

$$C\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} = I.$$

Quickly find $\mathbf{x}$ and $\mathbf{y}$ so that

$$C\mathbf{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad C\mathbf{y} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

8.5 Partitioned matrices

38. If we write something like this:

$$M = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right)$$

then this means $A \in M_{mp}(F)$, $B \in M_{mq}(F)$, $C \in M_{np}(F)$ and $D \in M_{nq}(F)$ are all matrices that when concatenated together, make the matrix $M \in M_{m+n,p+q}(F)$. M, when written this way, is called a *partitioned matrix* and A, B, C and D are called *blocks* in M. For example, if $A = I_{2\times 2}$, $B = (3,4)$, C is the 2×2 zero matrix, and $D = (1,-1)$, then

$$\left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right) = \left(\begin{array}{c|c} I & B \\ \hline 0 & D \end{array}\right) = \left(\begin{array}{cc|c} 1 & 0 & 3 \\ 0 & 1 & 4 \\ \hline 0 & 0 & 1 \\ 0 & 0 & -1 \end{array}\right) = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \end{pmatrix}.$$

Explain why (you don't need a detailed proof) the following equation holds:

$$\left(\begin{array}{c} A \\ \hline C \end{array}\right)\left(\begin{array}{c|c} B & D \end{array}\right) = \left(\begin{array}{c|c} AB & AD \\ \hline CB & CD \end{array}\right).$$

39. Compute the partitioned form of the matrix product (assume all the blocks are 2×2):

$$\left(\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right)\left(\begin{array}{c|c} I & 0 \\ \hline D & I \end{array}\right)$$

40. Find formulas for X, Y and Z in terms of A and B (assume A and B are invertible) if

$$\left(\begin{array}{c|c|c} X & 0 & 0 \\ \hline Y & 0 & I \end{array}\right)\left(\begin{array}{c|c} A & Z \\ \hline 0 & 0 \\ \hline B & I \end{array}\right) = \left(\begin{array}{c|c} I & 0 \\ \hline 0 & I \end{array}\right)$$

41. Show that if A_1 and A_2 are invertible matrices, then

$$\left(\begin{array}{c|c} A_1 & B \\ \hline 0 & A_2 \end{array}\right)$$

is invertible. Calculate the partitioned form of its inverse.

42. Show that if $A_1, ..., A_n$ are all invertible, then

$$\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_n \end{pmatrix}$$

is invertible. Calculate the partitioned form of its inverse.

8.6 Fundamental subspaces of a matrix

43. Given $A \in M_{mn}(F)$, define the *row space* $R(A)$ of A to be the span of the rows of A, and define the *column space* $C(A)$ of A to be the span of the columns of A. For example, if $A = \begin{pmatrix} 0 & 1 \\ 2 & 4 \end{pmatrix}$, the row space is $R(A) = Span((0,1),(2,4))$ and the column space is $C(A) = Span((0,2),(1,4))$.

    As the row space and column space are both spans, they are subspaces. For an arbitrary matrix $A \in M_{mn}(F)$, for what value of p is $R(A)$ a subspace of F^p? For a matrix $A \in M_{mn}(F)$, for what value of p is $C(A)$ a subspace of F^p?

44. Prove that $\mathbf{x} \in C(A)$ if and only if $\mathbf{x} = A\mathbf{z}$ for some vector $\mathbf{z}$.

45. Given $A \in M_{mn}(F)$, the *null space* $N(A)$ of A is the set of vectors $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{0}$. Prove that $N(A)$ is a subspace of F^p for some p. (What is the value of p)?

46. Given $A \in M_{mn}(F)$ where $F = \mathbb{R}$ or $\mathbb{C}$, the *left null space* $N(A^H)$ of A is the set of vectors $\mathbf{y}$ such that $A^H\mathbf{y} = \mathbf{0}$ (i.e. it is the null space of the Hermitian of A). Prove that $N(A^H)$ is a subspace of F^p for some p (use the preceding problem). (What is the value of p)?

9 Inner products and geometry

9.1 Inner products

1. Let $\mathbf{u} = (1,2,3)$, $\mathbf{v} = (1,0,-1)$ and $\mathbf{w} = (0,2,1)$. Think of these vectors as being elements of $\mathbb{R}^3$, which we endow with the usual inner product (i.e. dot product). Compute $(\mathbf{u}+\mathbf{v}) \cdot \mathbf{w}$ and $3(\mathbf{v} \cdot \mathbf{w})$. *Remark:* the first thing you are asked to compute here could also be written $< \mathbf{u}+\mathbf{v}, \mathbf{w} >$.

2. Let V be the vector space $\mathbb{C}^4$, endowed with the usual Hermitian inner product $<,>$. Compute $<(1, 2+i, -3i, 0), (1+i, 2-3i, 1+i, 2-i)>$.

3. Let V be $\mathbb{R}^3$ with the usual inner product $<,>$. Suppose that for *some* $\mathbf{v} \neq \mathbf{0} \in V$, $<\mathbf{v}, \mathbf{a}>=<\mathbf{v}, \mathbf{b}>$. Is it necessarily the case that $\mathbf{a} = \mathbf{b}$? Explain.

4. Let V be the vector space $\mathbb{R}^3$ with the usual inner product $<,>$. Suppose that for *every* $\mathbf{v} \in V$, $<\mathbf{v}, \mathbf{a}>=<\mathbf{v}, \mathbf{b}>$. Is it necessarily the case that $\mathbf{a} = \mathbf{b}$? Justify your answer.

5. For each of the following vector spaces, determine whether the given formula for $<,>$ actually defines an inner product. Justify your answer.

   (a) $V = \mathbb{C}^2$. Given $\mathbf{z} = (z_1, z_2)$ and $\mathbf{w} = (w_1, w_2)$, set
   $$<\mathbf{z}, \mathbf{w}>= z_1 w_1 + z_2 w_2.$$

   (b) V is the set of continuous functions from $\mathbb{R}$ to $\mathbb{R}$ which are $2\pi-$ periodic, that is, that $V = \{f \in C(\mathbb{R}, \mathbb{R}) : f(x) = f(x+2\pi) \text{ for all } x \in \mathbb{R}\}$. For $f, g \in V$, set
   $$<f, g>= \int_0^{2\pi} f(x)g(x)\, dx.$$

   (c) $V = \mathbb{R}^2$. Given $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$, set
   $$<\mathbf{x}, \mathbf{y}>= x_1 y_1.$$

   (d) $V = \mathbb{P}_3$. For $f, g \in V$, set
   $$<f, g>= f(0)g(0).$$

   (e) $V = \mathbb{R}^2$. Given $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$, set
   $$<\mathbf{x}, \mathbf{y}>= 4x_1 y_1 + x_1 y_2 + x_2 y_1 + 4x_2 y_2.$$

   (f) $V = \mathbb{P}_2$. For $f, g \in V$, set
   $$<f, g>= f(0)g(0) + f(1)g(1) + f(2)g(2).$$

6. Let $V = \mathbb{C}^n$. Give an example of a function $\llcorner, \lrcorner : V \times V \to \mathbb{C}$ which is antisymmetric, linear in the first coordinate, and which satisfies $\llcorner \mathbf{v}, \mathbf{v} \lrcorner \geq 0$ for all $\mathbf{v} \in V$ but is *not* an inner product.

7. Prove that if $<,>$ is an arbitrary inner product on a vector space V, then for all $d_1, ..., d_n \in F$ and all $\mathbf{v}, \mathbf{w}_1, ..., \mathbf{w}_n$ we have
$$\left\langle \mathbf{v}, \sum_{j=1}^n d_j \mathbf{w}_j \right\rangle = \sum_{j=1}^n \overline{d_j} < \mathbf{v}, \mathbf{w}_j > .$$

8. Let V be a vector sapce with inner product $<,>$. Suppose $<\mathbf{u}, \mathbf{v}>= 1+i$ and $<\mathbf{v}, \mathbf{w}>= 3i$ for vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$. Compute $<(2-i)\mathbf{v}, 2i\mathbf{u} - (1+i)\mathbf{w}>$.

9. Let V be a vector space and let $W = Span(\mathbf{w}_1, ..., \mathbf{w}_n)$. Prove that $<\mathbf{v}, \mathbf{w}>= 0$ for every $w \in W$ if and only if $<\mathbf{v}, \mathbf{w}_j>= 0$ for $j = 1, ..., n$.

9.2 Norms

10. Let $V = \mathbb{R}^4$, endowed with the usual inner product. Let $\mathbf{x} = (1, 2, -1, 4)$ and $\mathbf{y} = (-2, 2, 3, -8)$. Compute the norm of $\mathbf{x}$ and the distance between $\mathbf{x}$ and $\mathbf{y}$.

11. Find a real number k such that $(1, 2, k)$ and $(2, 4, 3)$ are distance 3 apart (in $\mathbb{R}^3$, endowed with the usual inner product). Before you solve this, can you determine by geometric reasoning (or a picture) how many solutions for k will exist?

12. Let V be the vector space $\mathbb{C}^3$, endowed with the usual Hermitian inner product $<,>$. Compute the distance between $(1, 2+i, -3i, 0)$ and $(1+i, 2-3i, 1+i, 2-i)$.

13. Let V be a vector space endowed with some inner product $<,>$ and associated norm $|| \cdot ||$ (we shorthand this sentence by saying V is an *inner product space*). We say that a vector $\mathbf{v} \in V$ is a *unit vector* if $||\mathbf{v}|| = 1$. Show that given any nonzero vector $\mathbf{v} \in V$, there is a unit vector in the same direction as $\mathbf{v}$. Such a unit vector is called a *normalized version* of $\mathbf{v}$, and the process of finding such a vector is called *normalizing* $\mathbf{v}$.

14. Let $V = \mathbb{C}^2$, endowed with the usual Hermitian inner product $<,>$. Normalize $\mathbf{v} = (1-2i, 2+i)$, and find a vector of length 8 in the direction of $\mathbf{v}$.

15. Let V be the inner product space of problem 8.5 (b). Find the norm of $f(x) = \sin x$.

16. Let $V = \mathbb{R}^2$ be endowed with inner product $< \mathbf{x}, \mathbf{y} >= 2x_1y_1 + 4x_2y_2$ (you do not need to prove that this is in fact an inner product). Sketch a picture of the set of points (x, y) such that $\mathbf{x} = (x, y)$ is a unit vector in V relative to this inner product. Find the norm of $(1, 2)$ relative to this inner product.

17. Prove the *Polarization Identities*, which say that for any inner product space V, and any vectors $\mathbf{v}, \mathbf{w} \in V$,

    (a) $\Re(< \mathbf{v}, \mathbf{w} >) = \frac{1}{4}\left(||\mathbf{v}+\mathbf{w}||^2 - ||\mathbf{v}-\mathbf{w}||^2\right)$.
    (b) $\Im(< \mathbf{v}, \mathbf{w} >) = \frac{1}{4}\left(||\mathbf{v}+i\mathbf{w}||^2 - ||\mathbf{v}-i\mathbf{w}||^2\right)$.

    Note: If V is a vector space over $\mathbb{R}$ (as opposed to $\mathbb{C}$), the first identity gives us $< \mathbf{v}, \mathbf{w} >= \frac{1}{4}\left(||\mathbf{v}+\mathbf{w}||^2 - ||\mathbf{v}-\mathbf{w}||^2\right)$.

18. Prove the *Parallelogram Law*, which says that for any inner product space V, and any vectors $\mathbf{v}, \mathbf{w} \in V$,

$$||\mathbf{v}+\mathbf{w}||^2 + ||\mathbf{v}-\mathbf{w}||^2 = 2\left(||\mathbf{v}||^2 + ||\mathbf{w}||^2\right).$$

9.3 Dual relations and classification of inner products on $\mathbb{C}^n$ and $\mathbb{R}^n$

19. A matrix $A \in M_n(\mathbb{C})$ is called *positive definite* if it is Hermitian and:

    - for every $\mathbf{v} \in \mathbb{C}^n$, the product $\mathbf{v}^H A\mathbf{v} \geq 0$ (in particular, this means $\mathbf{v}^H A\mathbf{v}$ is real for every $\mathbf{v}$), and
    - if $\mathbf{v} \in \mathbb{C}^n$ is such that $\mathbf{v}^H A\mathbf{v} = 0$, then $\mathbf{v} = \mathbf{0}$.

(a) Prove that for any positive definite, Hermitian matrix A, the formula $< \mathbf{v}, \mathbf{w} >= \mathbf{w}^H A\mathbf{v}$ defines an inner product on $\mathbb{C}^n$.

(b) Find a positive definite, Hermitian matrix A such that $\mathbf{w}^H A\mathbf{v}$ defines the usual (Hermitian) inner product on $\mathbb{C}^n$.

(c) Let $\{\mathbf{e}_1, ..., \mathbf{e}_n\}$ be the standard basis of $\mathbb{C}^n$, and let $< \mathbf{v}, \mathbf{w} >= \mathbf{w}^H A\mathbf{v}$ for some positive definite, Hermitian matrix $A \in M_n(\mathbb{C})$. Calculate (in terms of the entries of A) $< \mathbf{e}_j, \mathbf{e}_k >$ for all choices of j and k.

(d) Prove that every inner product on $\mathbb{C}^n$ is of the form $< \mathbf{v}, \mathbf{w} >= \mathbf{w}^H A\mathbf{v}$ for a suitably chosen positive definite, Hermitian matrix A. *Hint:* use the result of part (c) to define the entries of A in terms of the values of some inner products.

(e) Prove that the matrix A in part (d) is unique, given the inner product $<, >$. *Hint:* use part (c).

20. A matrix $A \in M_n(\mathbb{R})$ is called *positive definite* if it is symmetric and:
    - for every $\mathbf{v} \in \mathbb{R}^n$, the product $\mathbf{v}^T A\mathbf{v} \geq 0$, and
    - if $\mathbf{v} \in \mathbb{R}^n$ is such that $\mathbf{v}^T A\mathbf{v} = 0$, then $\mathbf{v} = \mathbf{0}$.

    (a) Prove that for any positive definite, symmetric matrix A, the formula $< \mathbf{v}, \mathbf{w} >= \mathbf{w}^T A\mathbf{v}$ defines an inner product on $\mathbb{R}^n$.

    (b) Prove that every inner product on $\mathbb{R}^n$ is of the form $< \mathbf{v}, \mathbf{w} >= \mathbf{w}^T A\mathbf{v}$ for a suitably chosen positive definite, symmetric matrix A.

    (c) Prove that the matrix A in part (b) is unique, given the inner product $<, >$.

    (d) Let $A = \begin{pmatrix} 3 & 1 \\ 1 & 5 \end{pmatrix}$ and define an inner product on $\mathbb{R}^2$ by $< \mathbf{v}, \mathbf{w} >= \mathbf{w}^T A\mathbf{v}$. Compute $< (2, 1), (-1, 3) >$.

    (e) Find the symmetric, positive definite matrix A such that the inner product of problem 8.5 (e) is of the form $< \mathbf{x}, \mathbf{y} >= \mathbf{y}^T A\mathbf{x}$.

21. Let $V = F^n$ where $F = \mathbb{R}$ or $\mathbb{C}$, with the usual (dot or Hermitian) inner product. Let $M \in M_n(F)$. Show that for any $\mathbf{x}, \mathbf{y} \in V$, $< M\mathbf{x}, \mathbf{y} >=< \mathbf{x}, M^H\mathbf{y} >$. (This means that if $F = \mathbb{R}$, $< M\mathbf{x}, \mathbf{y} >=< \mathbf{x}, M^T\mathbf{y} >$.)

    Remember these facts. They are used to prove the Fundamental Theorem of Linear Algebra.

22. Let $V = F^n$ where $F = \mathbb{R}$ or $\mathbb{C}$, with *some* inner product on it. Let $M \in M_n(F)$. Under what conditions is it true that for all $\mathbf{x}, \mathbf{y} \in V$, $< M\mathbf{x}, \mathbf{y} >=< \mathbf{x}, M^H\mathbf{y} >$?

9.4 Orthogonality

23. Find a complex number z such that the two vectors

$$\begin{pmatrix} 2+3i \\ 5-i \end{pmatrix} \quad \text{and} \begin{pmatrix} 1+i \\ z \end{pmatrix}$$

are orthogonal (with respect to the usual Hermitian inner product).

24. Let $V = \mathbb{R}^2$ be endowed with inner product $< \mathbf{x}, \mathbf{y} >= 4x_1y_1 + x_1y_2 + x_2y_1 + 4x_2y_2$ (this is the inner product of Problem 8.5 (e)). Sketch a picture of the set of points (x_1, x_2) such that $\mathbf{x} = (x_1, x_2)$ is orthogonal to $(1, 0)$.

25. Let $V = \mathbb{R}^3$ with the usual inner product (i.e. dot product). Given any two vectors $\mathbf{a} = (a_1, a_2, a_3)$ and $\mathbf{b} = (b_1, b_2, b_3)$ in V, define the *cross product* of these two vectors by

$$\mathbf{a} \times \mathbf{b} = (a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1).$$

In particular, the cross product of two vectors in $\mathbb{R}^3$ is itself a vector in $\mathbb{R}^3$.

(a) Compute $(1, 2, -1) \times (2, -1, 3)$.

(b) Prove that for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$, $(\mathbf{a} \times \mathbf{b}) \perp \mathbf{a}$ and $(\mathbf{a} \times \mathbf{b}) \perp \mathbf{b}$.

(c) What is the relationship between $\mathbf{a} \times \mathbf{b}$ and $\mathbf{b} \times \mathbf{a}$?

(d) Prove that $\mathbf{a} \times \mathbf{b} = \mathbf{0}$ if and only if $\mathbf{a} \,||\, \mathbf{b}$.

26. Prove the *Pythagorean Theorem* (for vectors), which says that if $\mathbf{v}$ and $\mathbf{w}$ belong to some inner product space V and $\mathbf{v} \perp \mathbf{w}$, then $||\mathbf{v} + \mathbf{w}||^2 = ||\mathbf{v}||^2 + ||\mathbf{w}||^2$.

27. Here is the converse of the Pythagorean Theorem: If $\mathbf{v}$ and $\mathbf{w}$ are vectors in some inner product space V such that $||\mathbf{v} + \mathbf{w}||^2 = ||\mathbf{v}||^2 + ||\mathbf{w}||^2$, then $\mathbf{v} \perp \mathbf{w}$.

(a) Show that the converse of the Pythagorean Theorem is false. *Hint:* to find a counterexample, you can set $V = \mathbb{C}$ (as a vector space over itself), with the usual Hermitian inner product.

(b) Show that if V is an inner product space over the field $\mathbb{R}$, then the converse of the Pythagorean Theorem is true.

28. Give a proof (using vectors and inner products) that the diagonals of a (Euclidean) rhombus are perpendicular. *Hint:* Draw a rhombus and think of the sides as vectors; give them names like $\mathbf{v}$ and $\mathbf{w}$. Since the shape you drew is a rhombus, what is true about $\mathbf{v}$ and $\mathbf{w}$? Draw the diagonals of the rhombus, figure out what they are in terms of the sides, and show they are orthogonal.

29. Prove (using vectors and inner products) that if the diagonals of a (Euclidean) parallelogram have the same length, then the parallelogram is a rectangle.

30. Prove (using vectors and inner products) that if A and B are ends of the diameter of a circle and if C is any other point on the same circle, then AC is perpendicular to BC.

9.5 Orthogonal decomposition, projections and orthogonal complements

31. Let V be an inner product space. What are $V^\perp$ and $\{\mathbf{0}\}^\perp$?

32. Let V be an inner product space and let $W = \text{Span}(\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_n)$. Show that $\mathbf{v} \in W^\perp$ if and only if $\mathbf{v} \perp \mathbf{w}_j$ for all $j = 1, ..., n$. *Hint:* This is pretty easy if you look back at Problem 8.9.

33. Let V be an inner product space. If $\mathbf{v}$ is a vector which is orthogonal to itself, what must be true about $\mathbf{v}$?

34. Let V be an inner product space. Let W be any subspace of V. What is $W \cap W^\perp$?

35. Let $S = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ be a set of nonzero vectors in an inner product space such that for all $i \neq j$, $\mathbf{v}_i \perp \mathbf{v}_j$. Show that S is a linearly independent set of vectors.

36. Let W be a subspace of an inner product space V. Prove that $W \subseteq (W^\perp)^\perp$. **Warning:** It is not true in general that $W = (W^\perp)^\perp$. More on this later.

37. Let $A \in M_{mn}(F)$, where $F = \mathbb{R}$ or $\mathbb{C}$. Prove $C(A) \subseteq [N(A^H)]^\perp$ and $R(A) \subseteq [N(A)]^\perp$. *Hint:* Use the dual relations.

38. Let V be $\mathbb{R}^3$ with the usual inner product. Consider the following vectors in V:

$$\mathbf{v} = (2, 1, 2) \quad \text{and} \quad \mathbf{w} = (0, 3, 4).$$

   (a) Find the projection of $\mathbf{v}$ onto $\mathbf{w}$.
   (b) Find the "vector component of $\mathbf{v}$ orthogonal to $\mathbf{w}$" (we called this $\mathbf{v}^\perp$ in class).
   (c) Find the angle θ between $\mathbf{v}$ and $\mathbf{w}$.
   (d) Let $W = \text{Span}(\mathbf{v}, \mathbf{w})$. Prove that $W^\perp$ is a line by proving that $W^\perp = Span(\mathbf{v} \times \mathbf{w})$.

39. Prove that $\text{proj}_{\mathbf{w}}\mathbf{v} = <\mathbf{v}, \mathbf{u}> \mathbf{u}$ where $\mathbf{u}$ is the normalized version of $\mathbf{w}$.

40. Prove the following statements (assume $\mathbf{w} \neq \mathbf{0}$):

   (a) $\mathbf{v} \perp \mathbf{w} \Leftrightarrow \text{proj}_{\mathbf{w}}\mathbf{v} = \mathbf{0}$.
   (b) $\mathbf{v} || \mathbf{w} \Leftrightarrow \text{proj}_{\mathbf{w}}\mathbf{v} = \mathbf{v}$.
   (c) $\text{proj}_{\mathbf{w}}\mathbf{v} \perp \mathbf{v} - \text{proj}_{\mathbf{w}}\mathbf{v}$.

41. Let θ be the angle between the vectors $\mathbf{v}$ and $\mathbf{w}$. Prove that

$$||\mathbf{v}^\perp|| = ||\mathbf{v}|| \sin\theta$$

where $\mathbf{v}^\perp = \mathbf{v} - \text{proj}_{\mathbf{w}}\mathbf{v}$.

42. Prove the *Law of Cosines*, which says that given any vectors $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{R}^n$, if θ is the angle between $\mathbf{v}$ and $\mathbf{w}$ then

$$||\mathbf{v} - \mathbf{w}||^2 = ||\mathbf{v}||^2 + ||\mathbf{w}||^2 - 2||\mathbf{v}|| \, ||\mathbf{w}|| \cos\theta.$$

43. Let $V = \mathbb{R}^2$ be endowed with inner product $< \mathbf{x}, \mathbf{y} >= 2x_1y_1 + 4x_2y_2$ (you do not need to prove that this is in fact an inner product). Relative to this inner product, find the measure of the angle between $(1, 0)$ and $(1, 1)$. Is it greater than, less than, or equal to $\pi/4$?

44. Prove the *Generalized Triangle Inequality*, which says that given any vectors $\mathbf{v}_1, ..., \mathbf{v}_n$, we have
$$\left|\left|\sum_{j=1}^{n} \mathbf{v}_j\right|\right| \leq \sum_{j=1}^{n} ||\mathbf{v}_j||.$$

45. In class, we proved the Triangle Inequality using the Cauchy-Schwarz Inequality (which was derived independently). Prove the Cauchy-Schwarz Inequality, assuming that the Triangle Inequality is true.

46. Let V be any vector space with inner product $<, >$. Prove the *(General) Orthogonal Decomposition Theorem*, which says that given any finite-dimensional subspace W of V, and any vector $\mathbf{v} \in V$, we can write $\mathbf{v} = \mathbf{v}^W + \mathbf{v}^\perp$ where $\mathbf{v}^W \in W$ and $\mathbf{v}^\perp \in W^\perp$.

    An outline of how to proceed: First, show that any finite-dimensional vector space (or subspace) has a basis where the first vector in the basis is orthogonal to each of the other elements. (To do this, take any basis $\{\mathbf{v}_1, \mathbf{v}_2, ...\}$ and replace all the vectors other than $\mathbf{v}_1$ with their vector component orthogonal to $\mathbf{v}_1$; prove that when you do this you still have a basis of V.)

    Second, to prove the theorem, use induction on the dimension of W. The base case $\dim(W) = 1$ was proven in class (you don't need to reprove it). The inductive step goes like this: suppose the result is true for all subspaces of dimension k and suppose $\dim(W) = k + 1$. By the result in the preceding paragraph, W has a basis where the first vector in the basis is orthogonal to all the other elements in the basis. Now let W' be the span of all the vectors in this basis *other than the first vector*; this is a subspace of dimension k. Apply the inductive hypothesis to write $\mathbf{v}$ as the sum of something in W' (hence in W) and something in $(W')^\perp$.

    Finally, figure out how to write the something in $(W')^\perp$ as a sum of something in W and something in $W^\perp$. Then explicitly write $\mathbf{v}$ as the sum of something in W plus something in $W^\perp$ and check that everything works.

    Remark: Your proof here doesn't really tell you how to compute $\mathbf{v}^W$ and $\mathbf{v}^\perp$. We'll talk about how to do that later–you use matrix operations to do it.

47. Prove that the choice of $\mathbf{v}^W$ and $\mathbf{v}^\perp$ in the Orthogonal Decomposition Theorem are uniquely determined by $\mathbf{v}$ and W. *Hint:* Suppose you have two different decompositions
$$\mathbf{v} = \mathbf{v}^W + \mathbf{v}^\perp \quad \text{and} \quad \mathbf{v} = \mathbf{x}^W + \mathbf{x}^\perp$$
    where $\mathbf{v}^W, \mathbf{x}^W \in W$ and $\mathbf{v}^\perp, \mathbf{x}^\perp \in W^\perp$. Subtract these equations and explain why it must be that $\mathbf{v}^W - \mathbf{x}^W \in W \bigcap W^\perp$. But what is $W \bigcap W^\perp$?

48. Let V be any finite-dimensional inner product space. Let W be a subspace of V.

(a) Prove that given any basis $\mathcal{B}$ for W and any basis $\mathcal{B}^\perp$ for $W^\perp$, $\mathcal{B} \cup \mathcal{B}^\perp$ is a basis of V.

(b) Use part (a) to find a formula relating $\dim(V), \dim(W)$ and $\dim(W^\perp)$.

(c) Use part (b) to find the dimension of $(W^\perp)^\perp$, in terms of the dimension of W.

(d) Prove that $W = (W^\perp)^\perp$. *Hint:* A part of problem 7.13 may be helpful here.

(e) *Note:* The statement $(W^\perp)^\perp = W$ does not always hold if $\dim(V) = \infty$. Explain what part of the argument in parts (a)-(d) of this question breaks down if $\dim(V) = \infty$.

49. Let W be a plane in $\mathbb{R}^3$ containing $\mathbf{0}$. Why must $W^\perp$ be a line? *Hint:* One of the parts of Problem 42 is relevant here. Show that for every plane in $\mathbb{R}^3$ containing $\mathbf{0}$, there is a vector $\mathbf{n}$ called a *normal vector* to the plane such that

$$\mathbf{x} \text{ lies in the plane } \Leftrightarrow \mathbf{n} \cdot \mathbf{x} = 0.$$

We know by definition that W is the span of two vectors, say $\mathbf{a}$ and $\mathbf{b}$. Given these vectors, how would you compute $\mathbf{n}$?

50. Let P be a plane in $\mathbb{R}^3$. Show that there is a vector $\mathbf{n}$ and a scalar d such that the plane can be characterized as

$$\mathbf{x} = (x, y, z) \in P \Leftrightarrow \mathbf{n} \cdot \mathbf{x} = d.$$

As before, $\mathbf{n}$ is called a *normal vector* to the plane. In particular, if you are given a definition of P as

$$P = \{\mathbf{p} + s\mathbf{v} + t\mathbf{w} : s, t \in \mathbb{R}\},$$

how do you compute $\mathbf{n}$ and d in terms of the given information $\mathbf{p}, \mathbf{v}, \mathbf{w}$?

Note: If you let $\mathbf{n} = (a, b, c)$, then the equation $\mathbf{n} \cdot \mathbf{x} = d$ becomes $ax + by + cz = d$; this proves that every plane in $\mathbb{R}^3$ has an equation of the form $ax + by + cz = d$.

51. Find the equation (in $ax + by + cz = d$ form) of the plane containing the points $(1, 2, -3)$, $(0, 1, 6)$ and $(-1, 1, 4)$.

52. Let V be an inner product space. A *hyperplane* in V is a set $H \subseteq V$ of the form

$$H = \{\mathbf{x} \in V : < \mathbf{x}, \mathbf{n} > = d\}$$

for some fixed vector $\mathbf{n} \neq 0$ (called a *normal vector* to the hyperplane) and some scalar d.

(a) Prove that every hyperplane is an affine subspace (and is a subspace if $d = 0$).

(b) Suppose V is finite-dimensional; find the dimension of a hyperplane H in terms of the dimension of V.

(c) Describe all hyperplanes in $V = \mathbb{R}$, all hyperplanes in $V = \mathbb{R}^2$ and all hyperplanes in $V = \mathbb{R}^3$ (all taken as vector spaces over $\mathbb{R}$).

(d) Formulate a definition of what it means for two hyperplanes to be *parallel.*

(e) Formulate a definition of the *angle* between two hyperplanes.

9.6 Coordinate systems and orthonormal bases

53. Find the coordinates of the function $2x^2 - 3x + 1$ relative to the basis $\{4 - x, 2 + 3x, x^2 - 1\}$ of $\mathbb{P}_2$.

54. Find the coordinates of the vector $(1, 4, 6, 0)$ relative to the basis

$$\mathcal{B} = \{(1, 0, 1, 0), (1, 0, -1, 0), (0, 1, 0, 1), (0, 1, 0, -1)\}$$

of $\mathbb{R}^4$.

55. Let $\mathbf{x}, \mathbf{y} \in V$ where V is some vector space and suppose $\mathcal{B} = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ is a basis of V. Prove:

    (a) $[\mathbf{x} + \mathbf{y}]_{\mathcal{B}} = [\mathbf{x}]_{\mathbf{B}} + [\mathbf{y}]_{\mathcal{B}}$.

    (b) $[r\mathbf{x}]_{\mathcal{B}} = r[\mathbf{x}]_{\mathcal{B}}$ for any scalar r.

56. (a) Suppose $\mathcal{B} = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ be an orthogonal basis of an inner product space V. Let $\mathbf{x} \in V$ and suppose $[\mathbf{x}]_{\mathcal{B}} = (c_1, ..., c_n)$ (i.e. $\mathbf{x} = \sum_{j=1}^{n} c_j \mathbf{v}_j$). Find a formula for the coordinates c_j in terms of $\mathbf{x}$ and the basis vectors. *Hint:* Inner products may be useful here.

    (b) Repeat part (a) if the basis $\mathcal{B}$ is assumed to be an orthonormal basis (as opposed to just an orthogonal basis). **The formula you get here should be remembered.**

57. Let $W = \mathit{Span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ where

$$\mathbf{v}_1 = (3, 1, -1, 3); \quad \mathbf{v}_2 = (-5, 1, 5, -7); \quad \mathbf{v}_3 = (1, 1, -2, 8)$$

(W is taken as a subspace of $\mathbb{R}^4$). Use the Gram-Schmidt procedure to find an orthonormal basis for W.

58. In problem 46 we proved the Orthogonal Decomposition Theorem, which says that given any finite-dimensional subspace W of an inner product space V, and given any $\mathbf{v} \in V$, we can write $\mathbf{v} = \mathbf{v}^W + \mathbf{v}^\perp$ where $\mathbf{v}^W \in W$ and $\mathbf{v}^\perp \in W^\perp$. We know that if $W = \mathit{Span}(\mathbf{w})$, then we can compute this decomposition because $\mathbf{v}^W = \text{proj}_{\mathbf{w}}\mathbf{v}$. In this problem we show (one method) of actually computing $\mathbf{v}^W$ in the general case:

    Suppose $\mathbf{v} \in V$ and $\{\mathbf{w}_1, ..., \mathbf{w}_m\}$ is an orthonormal basis for W. (Hypothetically, you could find an orthonormal basis by starting with any basis of W and using Gram-Schmidt on that basis.)

    (a) Let $\mathbf{x} = \sum_{k=1}^{m} < \mathbf{v}, \mathbf{w}_k > \mathbf{w}_k$. Explain why $\mathbf{x} \in W$.

    (b) Show that $\mathbf{v} - \mathbf{x} \in W^\perp$.

    (c) Explain how the work in (a) and (b) gives formulas for $\mathbf{v}^W$ and $\mathbf{v}^\perp$.

    (d) Let V be Euclidean 4−dimensional space. Compute the projection of $\mathbf{v} = (1, 2, 5, 6)$ onto the subspace W spanned by $(1, -4, 0, 1)$ and $(7, -7, -4, 1)$. (Use a calculator.)

59. Given a vector $\mathbf{v} \in V$ where V is an inner product space, and given a subspace $W \subseteq V$, define *the distance from* $\mathbf{v}$ *to* W to be

$$\text{dist}(\mathbf{v}, W) = \min\{||\mathbf{v} - \mathbf{w}|| : \mathbf{w} \in W\};$$

that is, that the distance from a vector to the subspace is the minimum distance from the vector to any point in the subspace.

   (a) Show that $\text{dist}(\mathbf{v}, W) = ||\mathbf{v}^\perp||$ where $\mathbf{v}^\perp$ is the vector component of $\mathbf{v}$ orthogonal to W. *Hint:* It is sufficient to show that $||\mathbf{v}^\perp|| \leq ||\mathbf{v}-\mathbf{w}||$ for all $\mathbf{w} \in W$ (convince yourself why). You can show this inequality using the Pythagorean Theorem (or by other means).

   (b) Use your answer to Problem 50 (d) to compute the distance (in Euclidean 4−dimensional space) from $(1, 2, 5, 6)$ to the plane spanned by $(0, 3, -4, 0)$ and $(2, 1, 0, 4/25)$.

   (c) Let W be a hyperplane containing 0 with normal vector $\mathbf{n}$. Prove that the distance from $\mathbf{v}$ to W is $| < \mathbf{n}, \mathbf{v} > |/||\mathbf{n}||$.

9.7 Orthogonal and unitary matrices

60. Let V be $\mathbb{R}^n$, endowed with the usual inner product. An invertible real matrix Q is called *orthogonal* if $Q^T = Q^{-1}$. The set of $n \times n$ orthogonal matrices is denoted O_n. Prove these statements:

   (a) If Q is orthogonal, so is Q^T. (This is not hard.)

   (b) Prove that every orthogonal matrix is invertible, and that the inverse of an orthogonal matrix is also orthogonal.

   (c) If Q is orthogonal, then the columns of Q form an orthonormal set of vectors (hence form an orthonormal basis of $\mathbb{R}^n$).

   (d) If Q is orthogonal, then the rows of Q form an orthonormal set of vectors (hence form an orthonormal basis of $\mathbb{R}^n$).

   (e) If Q is orthogonal, then for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $< Q\mathbf{x}, Q\mathbf{y} > = < \mathbf{x}, \mathbf{y} >$.

   (f) If Q is orthogonal, then for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, the angle between $\mathbf{x}$ and $\mathbf{y}$ is the same as the angle between $Q\mathbf{x}$ and $Q\mathbf{y}$.

   (g) Classify all the 2×2 orthogonal matrices in terms of one parameter θ. *Hint:* Why is the parameter called θ?

   (h) Prove the converse of (b), i.e. that if the columns of an $n \times n$ real matrix Q form an orthonormal basis of $\mathbb{R}^n$, then Q is orthonormal.

61. Let V be $\mathbb{C}^n$, endowed with the usual Hermitian inner product. An invertible complex matrix U is called *unitary* if $U^H = U^{-1}$. (Unitary matrices with real entries are exactly the orthogonal matrices.) The set of $n \times n$ unitary matrices is denoted U_n. Prove:

   (a) If U is unitary, so is U^H, $\overline{U}$, U^T and U^{-1} (so in particular, all unitary matrices are invertible).

(b) If U is unitary, then the columns of U form an orthonormal set of vectors (hence form an orthonormal basis of $\mathbb{C}^n$).

(c) If U is unitary, then the rows of U form an orthonormal set of vectors (hence form an orthonormal basis of $\mathbb{C}^n$).

(d) If U is unitary, then for all $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$, $< U\mathbf{x}, U\mathbf{y} >=< \mathbf{x}, \mathbf{y} >$.

(e) Prove the converse of (b), i.e. that if the columns of an $n \times n$ complex matrix U form an orthonormal basis of $\mathbb{C}^n$, then U is unitary.

(f) Prove the converse of (e), i.e. that if $U \in M_n(\mathbb{C})$ and for all $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$, $< U\mathbf{x}, U\mathbf{y} >=< \mathbf{x}, \mathbf{y} >$, then U is unitary.

10 Linear transformations

1. For each of the following functions T, decide whether or not T is a linear transformation from V_1 to V_2. For each function that is a linear transformation, describe the kernel and image of that transformation (by giving bases for those subspaces if necessary).

   (a) $V_1 = V_2 = \mathbb{P}_3$; $T(f) = f'$.

   (b) $V_1 = \mathbb{P}_3$; $V_2 = \mathbb{R}$; $T(f) = f'(0)$.

   (c) $V_1 = C([-1, 1], \mathbb{R})$; $V_2 = \mathbb{R}$; $T(f) = \int_{-1}^{1} (f(x))^2 \, dx$.

   (d) $V_1 = V_2 = \mathbb{R}^3$; $T(x_1, x_2, x_3) = (x_1 - x_2, 0, 0)$.

   (e) $V_1 = \mathbb{R}^3$; $V_2 = \mathbb{R}^2$; $T(x_1, x_2, x_3) = (x_1 + x_2 + 2x_3, x_2 - 2x_3)$.

   (f) $V_1 = V_2 = \mathbb{R}^2$; T reflects points through the line $x = 3$.

   (g) $V_1 = V_2 = \mathbb{R}^3$; $T(\mathbf{x}) = A\mathbf{x}$ where

   $$A = \begin{pmatrix} 1 & 2 & 2 \\ 1 & 3 & 5 \\ 0 & 1 & 3 \end{pmatrix}.$$

   Hint: The third column of A is equal to 3 times the second column minus 4 times the first column.

   (h) $V_1 = M_{mn}(\mathbb{R})$; $V_2 = M_{nm}(\mathbb{R})$; $T(A) = A^T$.

   (i) $V_1 = M_n(\mathbb{R})$; $V_2 = \mathbb{R}$; $T(A) = \text{trace}(A)$.

   (j) $V_1 = V_2 = M_2(\mathbb{R})$; T defined by

   $$T \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+b & 0 \\ 0 & c+d \end{pmatrix}.$$

   (k) $V_1 = M_{mn}(F)$; $V_2 = F^m$; $T(A)$ is the first column of A.

   (l) $V_1 = V_2 = \mathbb{P}_2$; $T(ax^2 + bx + c) = a + b(x + 1) + b(x + 1)^2$.

2. For each of the linear transformations in the previous problem, calculate $\dim(ker(T)) + \dim(image(T))$. What does this sum have to do with the linear transformation? Make a general conjecture.

3. Let $V = F$ be a vector space over itself. Describe all linear transformations from V to V.

4. Prove that for any linear transformation $T : V_1 \to V_2$,
$$T\left(\sum_{j=1}^{n} c_j \mathbf{v}_j\right) = \sum_{j=1}^{n} c_j T(\mathbf{v}_j)$$
for any vectors $\mathbf{v}_1, ..., \mathbf{v}_n$ and any scalars $c_1, ..., c_n$.

5. How many rows and columns must a matrix A have in order to define a mapping from F^4 into F^5 by the rule $T(\mathbf{x}) = A\mathbf{x}$?

6. Let $T : V_1 \to V_2$ be a linear transformation. Prove that if T maps a set of linearly independent vectors in V_1 to a set of vectors in V_2 which are linearly dependent, then the equation $T(\mathbf{x}) = \mathbf{0}$ has a nontrivial solution (nontrivial means $\mathbf{x} \neq \mathbf{0}$).

   Remark: The contrapositive of this statement is useful to know; it says that if $T(\mathbf{x}) = \mathbf{0}$ has only the solution $\mathbf{x} = 0$ (i.e. if T is injective), then T maps sets of linearly independent vectors to sets of linearly independent vectors.

7. Suppose $T : \mathbb{R}^2 \to \mathbb{R}^3$ is a linear transformation such that $T(1, 0) = (-1, 3, 7)$ and $T(0, 1) = (0, -2, -2)$. Find $T(2, 3)$ and $T(-4, 1)$. What is $T(x_1, x_2)$ for an arbitrary vector (x_1, x_2)?

8. Find a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ such that $T(-1, 2) = (2, 1)$ and $T(1, 3) = (-2, 4)$. *Hint:* Find a matrix A such that $T(\mathbf{x}) = A\mathbf{x}$.

9. Let $Q = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ for some θ. Consider the linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(\mathbf{x}) = Q\mathbf{x}$. Show that T rotates each vector in $\mathbb{R}^2$ counter-clockwise by θ degrees. *Hint:* Consider the polar coordinates of the vector $\mathbf{x}$.

10. Let V_1 and V_2 be vector spaces over the same field F and let $L(V_1, V_2)$ be the set of linear transformations from V_1 to V_2. (Sometimes $L(V_1, V_2)$ is called $Hom(V_1, V_2)$.) Show that $L(V_1, V_2)$ is a vector space. *Hint:* To do this, you need to define an addition and a scalar multiplication, i.e. given linear transformations T and S, what is $T + S$? What is rT? What is the additive identity element? How are additive inverses defined? If you answer these questions correctly, the rest of the vector space axioms don't really need to be checked because they're "obvious".

10.1 Dimension issues with linear transformations; surjectivity, injectivity, isomorphisms, etc.

11. Suppose $T : V_1 \to V_2$ is a linear transformation and suppose W is a subspace of V_1. Prove that $T(W)$, which is defined to be
$$T(W) = \{T(\mathbf{w}) : \mathbf{w} \in W\}$$

is a subspace of V_2 and that $\dim(T(W)) \leq \dim(W)$.

Hint: The "right" way to prove this is to take a basis $\mathcal{B} = \{\mathbf{w}_1, \mathbf{w}_2, ...\}$ of W and then let $T(\mathcal{B}) = \{T(\mathbf{w}_1), T(\mathbf{w}_2), ...\}$. Then, by the result of Problem 4, $T(W)$ is the span of the vectors in $T(\mathcal{B})$. Why is this the "right" proof? (Because you prove something stronger... formulate the stronger statement you have proven.)

Remark: As a special case of this, we see that if $T : V_1 \to V_2$ is a linear transformation, then $\dim T(V_1) \leq \dim(V_1)$.

12. Prove that if a linear transformation $T : V_1 \to V_2$ is injective, then $\dim(V_1) \leq \dim(V_2)$. *Hint:* take a basis $\mathcal{B} = \{\mathbf{w}_1, \mathbf{w}_2, ...\}$ of V_1 and then let $T(\mathcal{B}) = \{T(\mathbf{w}_1), T(\mathbf{w}_2), ...\}$. Show that $T(\mathcal{B})$ is a linearly independent set of vectors in V_2.

13. We say two vector spaces V_1 and V_2 over the same field are *isomorphic* (and write $V_1 \cong V_2$) if there is a bijective linear transformation $T : V_1 \to V_2$; such a transformation T is called an *isomorphism* (of vector spaces). Prove: if $V_1 \cong V_2$, then $\dim(V_1) = \dim(V_2)$.

14. Prove that if V_1 and V_2 are vector spaces over F with the same *finite* dimension, then $V_1 \cong V_2$.

15. Let V_1 and V_2 be finite-dimensional vector spaces of the same finite dimension and let $T : V_1 \to V_2$ be a linear transformation. Prove that T is surjective if and only if T is injective.

16. Prove the conjecture you made in Problem 2.

17. Suppose $T : V_1 \to V_2$ is an isomorphism. Prove that for any basis $\mathcal{B}$ of V_1, $T(\mathcal{B})$ is a basis of V_2.

18. Suppose $T : V_1 \to V_2$ is a linear transformation. Prove that if for any basis $\mathcal{B}$ of V_1, $T(\mathcal{B})$ is a basis of V_2, then T is an isomorphism.

19. Prove that isomorphism of vector spaces is an equivalence relation, i.e. prove for vector spaces V, W, X over the same field F:

    (a) $V \cong V$;

    (b) if $V \cong W$, then $W \cong V$;

    (c) if $V \cong W$ and $W \cong X$ then $V \cong X$.

    Hint: Remember that to say two vector spaces are isomorphic means there is a bijective linear transformation between the two spaces. Part (b) is essentially asking you to prove that the inverse of a bijective linear transformation is also a bijective linear transformation, and part (c) is essentially asking you to prove that the composition of two bijective linear transformations is a bijective linear transformation.

20. Let U, V and W be vector spaces over the same field and let $S : V \to W$ and $T : U \to V$ be linear transformations. Classify the following statements as true or false; prove the statements that are true and show the false statements are false by giving an explicit counterexample:

(a) If ST is injective, then T is injective. (Here, as always in linear algebra, ST means $S \circ T$, not S "times" T.)

(b) If ST is injective, then S is injective.

(c) If ST is surjective, then T is surjective.

(d) If ST is surjective, then S is surjective.

10.2 Standard matrices of linear transformations $F^n \to F^m$

21. Suppose $T : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear transformation that reflects all points in $\mathbb{R}^3$ through the $xy-$plane. Find the standard matrix of T.

22. Suppose $T : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transformation that first reflects points through the $x-$axis, then reflects points through the $y-$axis. Find the standard matrix of T, and show that T is a rotation (find the angle of rotation).

23. Find the standard matrix of each of the following linear transformations.

    (a) $T : \mathbb{R}^3 \to \mathbb{R}^3$ where T is projection onto $(2, 1, -2)$;

    (b) $T : \mathbb{R}^4 \to \mathbb{R}^2$ where $T(x_1, x_2, x_3, x_4) = (x_1 - 5x_3 + x_4, 0)$;

    (c) $T : \mathbb{R}^2 \to \mathbb{R}^2$ where T stretches vectors by a factor of 2, then rotates the plane $\pi/3$ radians counterclockwise;

    (d) $T : \mathbb{R}^3 \to \mathbb{R}^3$ described by $T(1, 0, 0) = (1, 2, 3)$; $T(0, 1, 0) = (1, 1, 1)$; $T(0, 0, 1) = (0, 0, 1)$;

24. Let A be the standard matrix of a linear transformation $S : \mathbb{R}^n \to \mathbb{R}^p$ and let B be the standard matrix of another linear transformation $T : \mathbb{R}^m \to \mathbb{R}^n$. Show that ST is linear (this may have been done in a previous problem) and that the standard matrix of ST is AB.

    Remark: This problem explains why matrix multiplication is defined the way that it is. Multiplying two matrices corresponds to composing the associated linear transformations.

11 Systems of linear equations

1. Write the following system of linear equations as a matrix equation $A\mathbf{x} = \mathbf{b}$ and as a vector equation:

$$\begin{cases} x - 2y + 3z - \frac{w}{2} = 2 \\ -\sqrt{3}x + y + 5z + \pi^2 w = -4 \\ 2x - 3z = 6 \end{cases}$$

2. (This problem is particularly important.) Let $A \in M_{mn}(F)$ and let $\mathbf{b} \in F^m$. Let $S \subseteq F^n$ be the set of solutions to the matrix equation $A\mathbf{x} = \mathbf{b}$. Prove:

    (a) $S \neq \emptyset$ if and only if $\mathbf{b} \in C(A)$.

(b) If $S \neq \emptyset$, then S is an affine subspace of F^n. *Hint:* assuming $S \neq \emptyset$, then there is a vector $\mathbf{p} \in S$. Consider the set $W = S - \mathbf{p}$. Show W is a subspace of F^n; what name have we given to W?

(c) If $N(A) \neq \{\mathbf{0}\}$, then S contains at least two elements if it is nonempty.

(d) If F is infinite and $N(A) \neq \{\mathbf{0}\}$, then S is infinite if it is nonempty.

(e) If $N(A) = \{\mathbf{0}\}$, then S contains at most one element.

(f) The equation $A\mathbf{x} = \mathbf{b}$ has a unique solution $\mathbf{x}$ if and only if ($\mathbf{b} \in C(A)$ and $N(A) = \{\mathbf{0}\}$).

3. Prove that no system of linear equations with more variables than equations can have a unique solution. *Hint:* Write the matrix version of the system and consider the dimensions of the fundamental subspaces associated to the coefficient matrix A.

4. Prove that if $A \in M_n(F)$ is invertible then the equation $A\mathbf{x} = \mathbf{b}$ has exactly one solution $\mathbf{x}$ for every choice of $\mathbf{b} \in F^n$. What is that solution? What does this fact imply about the null space of an invertible matrix? What does this fact imply about the row space of an invertible matrix? What does this fact imply about the column space of a real invertible matrix (hint: if A is invertible, so is A^T)?

5. Suppose A is some matrix with $N(A) = Span((2, -1, 3), (1, 0, 5))$. Suppose $A(1, 1, 4) = (2, 0, 3)$. Describe all solutions to the equation $A\mathbf{x} = (2, 0, 3)$.

6. Given the following subspace $W \subseteq \mathbb{C}^5$, find a matrix A such that $W = C(A)$:

$$W = \{(3 - 2i)r + s - 3t, s + (2 - 7i)t, 5t, -2is - t, (1 - i)r - 3s) : r, s, t \in \mathbb{C}\}$$

7. Assume that the following chart shows the number of grams of nutrients per ounce of food indicated:

	BEEF	POTATO	CABBAGE
PROTEIN	20	5	1
FAT	4	7	12
CARBOHYDRATES	15	20	5

If you eat a meal consisting of 9 ounces of meat, 20 ounces of potatoes, and 5 ounces of cabbage, how many grams of each nutrient do you get? More importantly, why is this problem in our linear algebra course (express the result using linear algebra language)?

8. Continuing with the data from the previous problem, suppose the army desires to use these same delectable foods to feed new recruits a dinner providing 305 grams of protein, 365 grams of fat, and 575 grams of carbohydrates. Write a system of equations which, when solved, will figure out how much of each food should be prepared for each recruit. Be sure to make clear what this problem is doing in a linear algebra course.

11.1 Gaussian elimination

9. Use Gaussian elimination to row reduce the following matrices to a row-echelon form, then to a reduced row-echelon form:

$$A = \begin{pmatrix} 1 & 2 & -3 & 0 \\ 2 & 4 & -2 & 2 \\ 3 & 6 & -4 & 3 \end{pmatrix}$$

$$B = \begin{pmatrix} 2 & 2 & -1 & 6 & 4 \\ 4 & 4 & 1 & 10 & 13 \\ 6 & 6 & 0 & 20 & 19 \end{pmatrix}$$

10. Describe all possible row-echelon forms of a 2×2 matrix. Describe all possible reduced row-echelon forms of a 2×2 matrix.

11. Solve the following systems using Gaussian elimination:

$$\text{(a)} \left\{ \begin{array}{rcrcrcr} x & - & 2y & + & z & = & 7 \\ 2x & - & y & + & 4z & = & 17 \\ 3x & - & 2y & + & 2z & = & 14 \end{array} \right.$$

$$\text{(b)} \left\{ \begin{array}{rcrcrcr} x & + & 2y & - & z & = & 3 \\ x & + & 3y & + & z & = & 5 \\ 3x & + & 8y & + & 4z & = & 17 \end{array} \right.$$

$$\text{(c)} \left\{ \begin{array}{rcrcrcrcr} x_1 & - & 2x_2 & + & x_3 & + & x_4 & = & 7 \\ & & 2x_2 & - & x_3 & - & 7x_4 & = & 6 \\ & & & - & 3x_3 & + & 2x_4 & = & -6 \end{array} \right.$$

$$\text{(d)} \left\{ \begin{array}{rcrcrcr} x & + & 2y & + & 3z & = & 1 \\ 2x & + & 3y & + & 4z & = & 0 \\ 3x & + & 4y & + & 5z & = & 1 \end{array} \right.$$

$$\text{(e)} \left\{ \begin{array}{rcrcrcrcrcr} x_1 & - & 2x_2 & - & 3x_3 & + & 5x_4 & - & 2x_5 & = & 4 \\ & & & & 2x_3 & - & 6x_4 & + & 3x_5 & = & 2 \\ & & & & & & & & 5x_5 & = & 10 \end{array} \right.$$

12. Determine if the following three planes in $\mathbb{R}^3$ have at least one common point of intersection: $x + 2y + z = 4$, $y - z = 1$ and $x + 3y = 0$.

13. Find the unique polynomial of degree at most 3 which goes through the points $(1,1)$,$(2,3)$,$(3,6)$ and $(4,10)$ (please use Gaussian elimination).

14. Express $(1,-2,5)$ as a linear combination of $(1,-3,2)$, $(2,-4,-1)$ and $(1,-5,7)$, if it can be done; if not, explain why this problem is impossible.

15. Let $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathbb{R}^3$ be $\mathbf{v}_1 = (1,-5,-3)$, $\mathbf{v}_2 = (-2,10,6)$ and $\mathbf{v}_3 = (2,-9,h)$.

    (a) For what values of h, if any, is $\mathbf{v}_3 \in Span(\mathbf{v}_1, \mathbf{v}_2)$?

(b) For what values of h, if any, does $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ form a linearly independent set?

16. Let $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathbb{R}^3$ be $\mathbf{v}_1 = (1, -3, 2)$, $\mathbf{v}_2 = (-3, 9, -6)$ and $\mathbf{v}_3 = (5, -7, h)$.

   (a) For what values of h, if any, is $\mathbf{v}_3 \in Span(\mathbf{v}_1, \mathbf{v}_2)$?

   (b) For what values of h, if any, does $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ form a linearly independent set?

11.2 Rank

17. Give the rank of coefficient matrices for each of the systems of equations in problem 11.

18. For each of the following pairs of matrices, assume A is row equivalent to B. In each case, find a basis for $C(A)$, $R(A)$, and $N(A)$.

$$\text{(a) } A = \begin{pmatrix} -1 & -4 & 9 & -7 \\ -1 & 2 & -4 & 1 \\ 5 & -6 & 10 & 7 \end{pmatrix} ; B = \begin{pmatrix} 1 & 0 & -1 & 5 \\ 0 & -2 & 5 & -6 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\text{(b) } A = \begin{pmatrix} -1 & 1 & -3 & 7 & 9 & -9 \\ 1 & 2 & -4 & 10 & 13 & -12 \\ 1 & -1 & -1 & 1 & 1 & -3 \\ 1 & -3 & 1 & -5 & -7 & 3 \\ 1 & -2 & 0 & 0 & -5 & -4 \end{pmatrix} ; B = \begin{pmatrix} 1 & 1 & -3 & 7 & 9 & -9 \\ 0 & 1 & -1 & 3 & 4 & -3 \\ 0 & 0 & 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

19. (a) Suppose $A \in M_{38}(\mathbb{R})$ has rank 3. Find $\dim N(A)$, $\dim R(A)$, $\dim C(A)$ and the rank of A^T.

   (b) Suppose $A \in M_{38}(\mathbb{R})$ has rank 1. Find $\dim N(A)$, $\dim R(A)$, $\dim C(A)$ and the rank of A^T.

   (c) Suppose $A \in M_{63}(\mathbb{C})$ has rank 3. Find $\dim N(A)$, $\dim R(A)$, $\dim C(A)$ and the rank of A^H.

   (d) Suppose $A \in M_{63}(\mathbb{R})$ has rank 2. Find $\dim N(A)$, $\dim R(A)$, $\dim C(A)$ and the rank of A^T.

   (e) Suppose $A \in M_{56}(F)$ has a 4$-$dimensional null space. How many linearly independent columns does A have? How many linearly independent rows does A have? How many pivot columns will A have if one performs Gaussian elimination on A?

   (f) Suppose $A \in M_{35}(F)$. What are the possible dimensions of $C(A)$? What are the possible dimensions of $R(A)$? What are the possible dimensions of $N(A)$?

   (g) Suppose $A \in M_{42}(F)$. What are the possible dimensions of $C(A)$? What are the possible dimensions of $R(A)$? What are the possible dimensions of $N(A)$?

   (h) Suppose $A \in M_{57}(F)$ is such that the subspace of all solutions to $A\mathbf{x} = \mathbf{0}$ has a basis of three vectors. What is the rank of A?

20. Is there a 3×4 matrix such that $\dim C(A) = 2$ and $\dim N(A) = 2$? If so, find such a matrix. If not, explain why not.

21. Give an example of a 4×3 matrix with rank 1.

22. In each of the following, determine (a) if the equation $A\mathbf{x} = \mathbf{0}$ has a nontrivial solution ("nontrivial" means a solution other than $\mathbf{x} = \mathbf{0}$) and (b) if the equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for every choice of $\mathbf{b}$.

    (a) A is a 3×3 matrix with 3 pivot columns.
    (b) A is a 2×4 matrix with rank 2.
    (c) A is an 8×6 matrix with $1-$dimensional null space.
    (d) A is a 5×3 matrix with 3 pivot columns.
    (e) A is a 4×4 matrix with rank 3.

23. Consider the system of equations $A\mathbf{x} = \mathbf{b}$ where

$$A = \begin{pmatrix} 1 & -2 & 4 & 0 \\ -1 & -1 & -1 & 0 \\ 0 & 4 & -4 & 3 \\ 3 & 2 & 4 & -2 \\ 1 & 0 & 2 & 2 \end{pmatrix} \text{ and } \mathbf{b} = \begin{pmatrix} 4 \\ 2 \\ -11 \\ -2 \\ -2 \end{pmatrix}.$$

    Let $\mathbf{a}_j$ represent the j^{th} column of A.

    (a) Find the solution set of this system.
    (b) Find bases for the column space of A, the row space of A, and the null space of A.
    (c) Find the dimensions of $R(A)$, $C(A)$, $N(A)$ and $N(A^T)$.
    (d) Find the rank of A.
    (e) Let $T : \mathbb{R}^a \to \mathbb{R}^b$ be the linear transformation defined by $T(\mathbf{x}) = A\mathbf{x}$. What are a and b? Find bases for the kernel and image of T, and the dimensions of the kernel and image of T.
    (f) Is $\mathbf{b}$ in the span of the columns of A? If so, write $\mathbf{b}$ as a linear combination of the columns of A. If not, explain why not.
    (g) Find a vector which is not in the column space of A.
    (h) Is there any vector $\mathbf{v} \in \mathbb{R}^5$ for which $A\mathbf{x} = \mathbf{v}$ has exactly one solution? If so, find such a vector $\mathbf{v}$. If not, explain why not.
    (i) Do the vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4$ form a linearly independent set? Why or why not?
    (j) Do the vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_4$ form a linearly independent set? Why or why not?
    (k) Do the vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4, \mathbf{b}$ form a linearly independent set? Why or why not?

24. Do the vectors $(1, 1, 0, 0)$, $(0, 1, 1, 0)$, $(0, 0, 1, 1)$ and $(1, 0, 0, 1)$ span $\mathbb{R}^4$? Why or why not? Do they span $\mathbb{C}^4$? Why or why not?

25. Let W be the subspace of $\mathbb{R}^4$ spanned by $(1, 2, -5, 2)$ and $(0, 1, 3, -1)$. Find a basis of $W^\perp$.

11.3 Fundamental theorem of linear algebra

26. Let $T : F^n \to F^m$ be a linear transformation with standard matrix A. Prove:
   (a) $\ker(T) = N(A)$.
   (b) $T(F^n) = C(A)$.
   (c) T is injective if and only if $N(A) = \{\mathbf{0}\}$.
   (d) T is surjective if and only if $C(A) = F^m$.

27. Prove the *Fundamental Theorem of Linear Algebra*, which says:
   Let $F = \mathbb{R}$ or $\mathbb{C}$ and let $<,>$ be standard (dot or Hermitian) inner product. Then for any $A \in M_{mn}(F)$,
   (a) $C(A) = [N(A^H)]^\perp$;
   (b) $R(A) = [N(A)]^\perp$.

28. Determine the premultiplier matrix that clears the column beneath the first pivot when performing Gaussian elimination on the matrix
$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 4 \end{pmatrix}.$$
   In other words, if you call the above matrix A, find a matrix E such that the matrix you get once the column under the first pivot is cleared is EA.

11.4 Matrix inverses

29. Use the Gauss-Jordan method to find the inverses of these matrices (if the inverse exists):
$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 3 & 10 & 3 & 8 \\ 3 & -2 & 8 & 7 \\ 2 & 1 & 4 & -5 \\ 5 & 11 & 7 & 3 \end{pmatrix}$$

30. Use as few calculations as possible (use theory, if possible) to determine whether or not each of these matrices are invertible (you need not find the inverse):
$$\begin{pmatrix} 5 & 0 & 0 \\ -3 & -7 & 0 \\ 8 & 5 & -1 \end{pmatrix} \quad \begin{pmatrix} 0 & 3 & -5 \\ 1 & 0 & 2 \\ -4 & -9 & 7 \end{pmatrix} \quad \begin{pmatrix} 1 & -5 & 4 \\ 0 & 3 & 4 \\ -3 & 6 & 0 \end{pmatrix} \quad \begin{pmatrix} -7 & 0 & 4 \\ 3 & 0 & -4 \\ 2 & 0 & 9 \end{pmatrix}$$

31. If A, B and C are $n \times n$ invertible matrices, does the matrix equation $C^{-1}(A + X)B^{-1} = I$ have a solution X? If so, find it.

32. Suppose A is a 4×3 matrix and B is a 3×4 matrix. Is it possible for the matrix AB to be invertible? Make a conjecture that generalizes this; prove your conjecture.

12 Coordinate systems and changes of basis

1. Consider the bases $\mathcal{B} = \{(1,-2,1),(2,3,2),(1,1,0)\}$ and $\mathcal{B}' = \{(1,0,2),(1,0,3),(2,2,1)\}$ of $\mathbb{R}^3$.

   (a) Find the $\mathcal{B}-$coordinate vector of $\mathbf{x} = (1,2,3)$.

   (b) If the $\mathcal{B}-$coordinate vector of $\mathbf{x}$ is $(2,1,-1)$, find $\mathbf{x}$. Express the computation you did here as matrix multiplication.

   (c) Find the $\mathcal{B}'-$coordinates of the vector whose $\mathcal{B}-$coordinates are $(1,-1,2)$.

2. Suppose $T : \mathbb{R}^3 \to \mathbb{R}^2$ has standard matrix

$$\begin{pmatrix} 1 & -3 & 6 \\ 2 & -1 & 5 \end{pmatrix}.$$

   Find the matrix of T relative to the basis $\{(1,1,1),(1,2,-1),(1,0,-1)\}$ of $\mathbb{R}^3$ and the basis $\{(5,2),(7,3)\}$ of $\mathbb{R}^2$.

3. Let $T : V \to V$ be a linear transformation (where V is finite-dimensional) be such that the matrix of T with respect to some basis of V (chosen for both the domain and range) is the identity matrix. Prove that the matrix of T with respect to every basis of V is the identity matrix (so long as you choose the same basis of V for both the domain and range).

4. Let $T : V_1 \to V_2$ be a linear transformation between finite-dimensional vector spaces V_1 and V_2. Show that if A and A' are two different matrices of T (taken with respect to different bases of V_1 and V_2, then A and A' have the same rank.

5. Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ and let $T : M_2(\mathbb{R}) \to M_2(\mathbb{R})$ be defined by $T(M) = AM$. Find the matrix of T relative to the basis $\{\Delta_{11}, \Delta_{12}, \Delta_{21}, \Delta_{22}\}$ of $M_2(\mathbb{R})$.

6. Let $T : \mathbb{P}^3 \to \mathbb{P}^3$ be defined by $(T(f))(x) = f'(x) - f(1)$. Find $T(x^2+2)$. Next, find the matrix of T relative to the basis $\{1, x, x^2, x^3\}$ of $\mathbb{P}^3$. Is T invertible? Why or why not?

7. Let W be the subspace of $C^\infty(\mathbb{R},\mathbb{R})$ which is spanned by $e^{2x}, e^{2x}\sin x$ and $e^{2x}\cos x$. Let $T : W \to W$ be defined by $T(f) = f'$. Find the matrix of T relative to the basis $\{e^{2x}, e^{2x}\cos x, e^{2x}\sin x$. Is T invertible? Why or why not?

8. Define $T : \mathbb{P}^2 \to \mathbb{R}^3$ by $T(f) = (f(-1), f(0), f(1))$. Show that T is linear; find the matrix of T relative to the basis $\{1, x, x^2\}$ of $\mathbb{P}^2$ and the standard basis of $\mathbb{R}^3$. Is T invertible? Why or why not?

9. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ have standard matrix $\begin{pmatrix} -3/5 & 4/5 \\ 4/5 & 3/5 \end{pmatrix}$. Find the matrix of T relative to the basis $\{(1,2),(-2,1)\}$ of $\mathbb{R}^2$.

13 Determinants

13.1 Computations with permutations

1. Let $\sigma \in S_8$ be defined by

$$\begin{array}{c|cccccccc} x & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ \hline \sigma(x) & 4 & 7 & 3 & 6 & 2 & 1 & 8 & 5 \end{array}.$$

   (a) Find $\sigma^2(1)$.

   (b) Write σ in cycle notation.

2. Let $\sigma = (124)(37)$ and $\tau = (25467)$ be permutations in S_7. Compute σ^{-1}, $\sigma\tau$, $\tau\sigma$, $\sigma\tau^{-1}$ and $\tau\sigma\tau$; write your answers in cycle notation.

3. Show that every cycle $(x_1 x_2 ... x_k)$ can be written as a product of transpositions. How many transpositions do you multiply to obtain a cycle with k elements in it?

4. Show that if you can write a cycle as a product of k transpositions, then you can write it as a product of $k+2$ transpositions (and therefore by induction, you can write it as a product of $k+2j$ transpositions for any positive integer j).

5. Fix n. Let $P : \mathbb{R}^n \to \mathbb{R}$ be defined by $P(x_1, ..., x_n) = \prod_{i<j}(x_i - x_j)$. (For example, if $n = 3$, we have $P(x_1, x_2, x_3) = (x_1 - x_2)(x_1 - x_3)(x_2 - x_3)$.) Let $\sigma \in S_n$ and define the *sign* of σ to be

$$sgn(\sigma) = \frac{P(x_1, x_2, ..., x_n)}{P(x_{\sigma(1)}, x_{\sigma(2)}, ..., x_{\sigma(n)})}.$$

   For example, suppose $n = 3$ and $\sigma = (12)$. Then

$$sgn(\sigma) = \frac{P(x_1, x_2, x_3)}{P(x_{\sigma(1)}, x_{\sigma(2)}, ..., x_{\sigma(n)})} = \frac{P(x_1, x_2, x_3)}{P(x_2, x_1, x_3)} = \frac{(x_1 - x_2)(x_1 - x_3)(x_2 - x_3)}{(x_2 - x_1)(x_2 - x_3)(x_1 - x_3)} = -1.$$

   (a) Convince yourself that the example is indeed worked out properly. *Hint:* Think carefully about what $P(x_2, x_1, x_3)$ means. What is $P(x_3, x_2, x_4, x_1)$?

   (b) Show that the sign of any permutation is either 1 or -1. *Note:* from this fact we can define the *parity* of a permutation: $\sigma \in S_n$ is said to be *even* if $sgn(\sigma) = 1$ and *odd* if $sgn(\sigma) = -1$.

   (c) What is the sign of the identity permutation (defined by $\sigma(x) = x$ for all x)?

   (d) Show that the sign of any transposition is -1.

   (e) Show that $sgn(\sigma\tau) = sgn(\sigma) \cdot sgn(\tau)$ for any permutations $\sigma, \tau \in S_n$.

   (f) Show that $sgn(\sigma^{-1}) = sgn(\sigma)$ for all $\sigma \in S_n$.

   (g) Show that if σ is written as the product of k transpositions, then $sgn(\sigma) = (-1)^k$. *Hint:* use previous parts of this question.

   (h) Find the sign of a cycle $(x_1 x_2 ... x_n)$. Use this fact to determine a method for determining the sign of a permutation quickly without having to work out any P's as above.

(i) Find the sign of $\sigma = (12854)(367) \in S_9$.

(j) Find the signs of all the permutations in S_4. How many permutations are even? How many are odd? Make a conjecture about the numbers of even and odd permutations in S_n.

6. Let $A \in M_n(F)$ and let $\sigma \in S_n$. The expression $\prod_{j=1}^{n} a_{\sigma(j),j}$ is the product of n entries taken from A. Show that the list of numbers multiplied together in this expression has one number taken from each row, and one number taken from each column, of A. *Note:* The $\prod$ here means multiply the entries that follow it, in the same way that $\sum$ means sum.

7. Write down a 4×4 matrix A (choose your own entries) and let $\sigma \in S_4$ be $\sigma = (12)(34)$. Compute, for your matrix, $\prod_{j=1}^{n} a_{\sigma(j),j}$.

13.2 Properties of the determinant

8. A square matrix is called *lower triangular* if all the entries above its diagonal are zero and is called *upper triangular* if all the entries below its diagonal are zero. (Diagonal matrices are those which are both upper and lower triangular.) Show that for any (upper or lower) triangular matrix, the determinant of that matrix is equal to the product of its diagonal entries. *Hint:* problem 6 is useful here.

9. Find an explicit formula for the determinant of a 3×3 matrix (write it out from the definition).

10. To find the determinant of a 5×5 matrix, how many terms would have to be added/subtracted together to obtain the determinant (if one used the definition)?

11. Show that if two rows of a matrix are identical, then the determinant of that matrix is zero. *Hint:* Interchange the two rows.

12. (a) Show that the determinant of a matrix is nonzero if and only if it is row equivalent to the identity matrix (thus if and only if the matrix is invertible, by previous theory).

    (b) Explain how performing row reductions on a matrix gives a method of computing its determinant.

13. Let (a, b) and (u, v) be two vectors in $\mathbb{R}^2$. Construct a parallelogram whose sides are the vectors (a, b) and (u, v). Find the area of this parallelogram. What does this area computation have to do with determinants?

14. Suppose A is $n \times n$. How are the following quantities related to $\det(A)$?

$$\det(2A) \quad \det(-A) \quad \det(A^2) \quad \det(A^{-1}) \quad \det(A^T) \quad \det(\overline{A}) \quad \det(A^H)$$

15. Suppose A is a square matrix such that $\det(A^4) = 0$. Can A be invertible? Why or why not?

16. Suppose Q is an orthogonal matrix. What must be true about $\det(Q)$?

17. Suppose U is a unitary matrix. What must be true about $\det(U)$?

18. In this problem we derive *Cramer's Rule*, which gives a method of solving systems of n linear equations in n variables which have a unique solution (the method uses determinants).

    First, given $A \in M_n(F)$ and $\mathbf{b} \in F^n$, define the matrix $A_i(\mathbf{b})$ to be the $n \times n$ matrix which has the same entries of A except on its i^{th}column, and whose i^{th} column is $\mathbf{b}$. For example, if $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ and $\mathbf{b} = (5, 6)$, then $A_2(\mathbf{b}) = \begin{pmatrix} 1 & 5 \\ 3 & 6 \end{pmatrix}$.

    Let $A \in M_n(F)$ be invertible and let $\mathbf{b} \in F^n$ be given. Let $\mathbf{x} = A^{-1}\mathbf{b}$, i.e. $\mathbf{x}$ is the solution to the equation $A\mathbf{x} = \mathbf{b}$.

    (a) Show that $A\mathbf{x} = \mathbf{b}$ if and only if $A\, I_i(\mathbf{x}) = A_i(\mathbf{b})$.

    (b) Use part (a) to show that $(\det A)(\det I_i(\mathbf{x})) = \det(A_i(\mathbf{b}))$.

    (c) Calculate $\det(I_i(\mathbf{x}))$ in terms of the components of $\mathbf{x}$.

    (d) Use parts (b) and (c) to derive a formula for the components of $\mathbf{x}$ in terms of the determinants of A and $A_i(\mathbf{b})$. This formula is called *Cramer's Rule.*

    (e) Use Cramer's Rule to solve

    $$\begin{cases} x + 2y = 2 \\ -x + 4y = 1 \end{cases}.$$

19. Suppose A is an invertible $n \times n$ matrix. Let $B = A^{-1}$ and denote the respective entries of A and B by a_{ij} and b_{ij}, respectively.

    (a) Let $\mathbf{b}_j$ be the j^{th} column of B, i.e. $\mathbf{b}_j = (b_{1j}, ..., b_{nj})$. What is $A\mathbf{b}_j$?

    (b) Use part (a) and Cramer's Rule to find a formula for b_{ij} in terms of determinants.

13.3 Computing determinants

20. Find determinants of each of the following matrices:

$$\begin{pmatrix} 3 & 0 & 4 \\ 2 & 3 & 2 \\ 0 & 5 & -1 \end{pmatrix} \qquad \begin{pmatrix} -1 & 0 & 0 \\ 2 & 3 & 0 \\ 0 & 5 & -4 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 2 & 1 & 0 \\ -2 & 3 & 3 & -1 \\ 1 & 0 & 5 & 2 \end{pmatrix}$$

14 Eigentheory

14.1 Definitions and properties

1. Let V be a vector space over F and let $T : V \to V$ be a linear transformation from V to itself. We say a subspace $W \subseteq V$ is *invariant (under T)* if $T(W) \subseteq W$.

   (a) Given linear transformation $T : V \to V$, there are two "trivial" subspaces of V which are always invariant (no matter what T is). What are they?

   (b) Show that $\ker(T)$ and $T(V)$ are both invariant subspaces of V under T.

   (c) Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the transformation in problem 9 on page 40. Find all the subspaces of $\mathbb{R}^2$ which are invariant under T. *Hint:* the result of problem 9, page 40 may be helpful.

   (d) Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the transformation $T(x, y) = (-y, x)$. Find all invariant subspaces under this T.

2. In this problem we will characterize all one-dimensional invariant subspaces of linear transformations:

   (a) Suppose $\mathbf{w} \neq \mathbf{0}$ is a vector and $\lambda \in F$ is such that $T(\mathbf{w}) = \lambda\mathbf{w}$. Show that $W = Span(\mathbf{w})$ is an invariant subspace (of dimension 1) under T.

   (b) Suppose W is a one-dimensional subspace of V which is invariant under T. Show that there exists a scalar λ such that $T(\mathbf{w}) = \lambda\mathbf{w}$ for every $\mathbf{w} \in W$.

3. Let $T : V \to V$ be a linear transformation. We say that a scalar $\lambda \in F$ is an *eigenvalue* of T if there exists a **nonzero** vector $\mathbf{w} \in V$ s.t. $T(\mathbf{w}) = \lambda\mathbf{w}$. Given an eigenvalue λ, any vector $\mathbf{w} \in V$ satisfying $T(\mathbf{w}) = \lambda\mathbf{w}$ is called an *eigenvector (corresponding to λ)* of T.

   (a) Fix an eigenvalue λ of T. Prove that the set of eigenvectors corresponding to λ is a subspace of V, called the *eigenspace* corresponding to λ.

   (b) Suppose $\lambda_1, ..., \lambda_k$ are distinct eigenvalues of T ("distinct" means $\lambda_i \neq \lambda_j$ for all $i \neq j$) with corresponding eigenvectors $\mathbf{w}_1, ..., \mathbf{w}_k$. Prove that the set $\{\mathbf{w}_1, ..., \mathbf{w}_k\}$ is linearly independent.

      Hint: Suppose the set is linearly dependent. Then there is an initial vector $\mathbf{w}_m$ which depends on the previous vectors, i.e. you get an equation $\mathbf{w}_m = c_1\mathbf{w}_1 + ... + c_{m-1}\mathbf{w}_{m-1}$ for scalars $c_1, ..., c_{m-1}$. Apply T to both sides of this equation and derive a contradiction.

   (c) Use the result of part (b) to prove that if $\dim V = n < \infty$, then $T : V \to V$ has at most n distinct eigenvalues.

4. Let $A \in M_n(F)$. We say that a scalar $\lambda \in F$ is an *eigenvalue* of the matrix A if there exists a **nonzero** vector $\mathbf{w} \in F^n$ s.t. $A\mathbf{w} = \lambda\mathbf{w}$. Given an eigenvalue λ, any vector $\mathbf{w} \in F^n$ satisfying $A\mathbf{w} = \lambda\mathbf{w}$ is called an *eigenvector (corresponding to λ)* of A. (Put another way, eigenvalues/vectors of A are eigenvalues/vectors of the linear transformation $T(\mathbf{x}) = A\mathbf{x}$. Read question 3 for the definitions of eigenvalue and eigenvector for a linear transformation.)

(a) Prove that $\mathbf{w} \neq \mathbf{0}$ is an eigenvector of A corresponding to λ if and only if $\mathbf{w} \in N(A - \lambda I)$.

(b) Prove that λ is an eigenvalue of A if and only if $\det(A - \lambda I) = 0$.

(c) Given matrix $A \in M_n(F)$, the expression $\det(A - xI)$ is called the *characteristic polynomial* of A. Explain why this expression is a polynomial (in the variable x) of degree n (be sure to explain why the degree cannot be less than n).

(d) Prove that if n is odd, then every $n \times n$ real matrix has a real eigenvalue.

(e) Prove that every complex matrix has a complex eigenvalue. *Hint:* Apply the Fundamental Theorem of Algebra, which says that every nonconstant polynomial with complex coefficients has a complex root.

(f) Find all eigenvalues of the matrix

$$A = \begin{pmatrix} -4 & -4 & -4 \\ 6 & 4 & 6 \\ 2 & 4 & 2 \end{pmatrix}.$$

Hint: use part (b) to solve for λ.

5. (Read questions 3 and 4 for the definitions of eigenvalue and eigenvector.)

(a) Let V and W be vector spaces over F with $\dim V = \dim W$. Let ϕ be any isomorphism $\phi : V \to W$. Let $T : V \to V$ be linear and let $S : W \to W$ be defined by $S = \phi \circ T \circ \phi^{-1}$.

i. Draw a commutative diagram (with sets and arrows) illustrating the relationship between S and T.

ii. Show that if $\lambda \in F$ is an eigenvalue for T with eigenvector $\mathbf{w}$, then λ is also an eigenvalue for S. What is an eigenvector (in terms of $\mathbf{w}$) corresponding to λ for S?

(b) Let V be a real vector space of dimension $n < \infty$. Prove that if n is odd, every linear transformation $T : V \to V$ has a real eigenvalue. *Hint:* use part (a) of this problem and part (d) of question 4.

(c) Let V be a vector space over $\mathbb{C}$ of dimension $n < \infty$. Prove that every linear transformation $T : V \to V$ has a (complex) eigenvalue. *Hint:* Use part (a) of this problem and the result of part (e) of question 4.

(d) Show that if A and A' are two different $n \times n$ matrices representing the same linear transformation $T : V \to V$, then A and A' have the same eigenvalues.

(e) Show that if P is an invertible matrix, then A and PAP^{-1} have the same eigenvalues.

6. (Read questions 3 and 4 for the definitions of eigenvalue and eigenvector.) Let V be a finite-dimensional vector space over F and let $T : V \to V$ be a linear transformation. Let $\mathcal{B} = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ be a basis of V and let $A \in M_n(F)$ be the matrix of T relative to $\mathcal{B}$.

(a) Prove that the following three statements are equivalent:
   i. A is upper triangular;
   ii. $T(\mathbf{v}_k) \in Span(\mathbf{v}_1, ..., \mathbf{v}_k)$ for all $k \leq n$;
   iii. $Span(\mathbf{v}_1, ..., \mathbf{v}_k)$ is invariant under T, for all $k \leq n$.

(b) Suppose the matrix A of T relative to some basis of V is upper triangular. Show that the eigenvalues of T, and also the eigenvalues of A, are the diagonal entries of A. *Hint:* Use part (a) of question 4.

(c) Prove that the following are equivalent:
   i. A is diagonal;
   ii. Every vector in the basis $\mathcal{B}$ is an eigenvector of A.

(d) Prove that if a linear transformation $T : V \to V$ has n **distinct** eigenvalues (where $\dim V = n$), then there is a basis $\mathcal{B}$ of V such that the matrix of T associated to $\mathcal{B}$ is diagonal.

(e) Prove that if a matrix $A \in M_n(F)$ has n **distinct** eigenvalues, then there is an invertible matrix $P \in M_n(F)$ such that $P^{-1}AP$ is a diagonal matrix.

7. Prove that if V is a finite dimensional vector space over $\mathbb{C}$ and $T : V \to V$ is linear, then there is a basis of V such that T has an upper-triangular matrix relative to that basis. *Hint:* Use induction on $\dim V$. The case $\dim V = 1$ is trivial. Here is an outline of how to carry out the induction:
   (a) Use part (c) of question 5 to find an eigenvalue λ of T.
   (b) Define the linear transformation $S : V \to V$ by $S(\mathbf{x}) = T(\mathbf{x}) - \lambda\mathbf{x}$. Show $W = Im(S)$ is invariant under T, and show that the kernel of S consists of eigenvectors of T corresponding to λ.
   (c) Use the induction hypothesis to find a basis of W such that the matrix B of $T : W \to W$ relative to that basis is upper-triangular.
   (d) Finally, show that the upper-triangular matrix
   $$A = \left(\begin{array}{c|c} B & * \\ \hline 0 & \lambda I \end{array}\right)$$
   is the matrix of T relative to an appropriately chosen basis of V.

8. Prove that if V is a finite dimensional vector space over $\mathbb{C}$ and $T : V \to V$ is linear, then there is a orthonormal basis of V such that T has an upper-triangular matrix relative to that basis.

9. Show that the trace of a matrix is equal to the sum of its eigenvalues. *Hint:* use problems 7 and 6 (b).

10. Show that the determinant of a matrix is equal to the product of its eigenvalues. *Hint:* use problems 7 and 6 (b).

11. Let $A \in M_n(\mathbb{C})$ be a matrix with only real entries. Suppose A has a complex eigenvalue λ. Show $\overline{\lambda}$ is also an eigenvalue of A. What is the relationship between the eigenvectors corresponding to λ and the eigenvectors corresponding to $\overline{\lambda}$?

14.2 Diagonalization and applications

12. A matrix $A \in M_n(F)$ is said to be *diagonalizable* if there is an invertible matrix P such that the matrix $\Lambda = P^{-1}AP$ is diagonal. We know from the previous subsection (in particular, problem 6 (e)) that if A has n **distinct** eigenvalues, then A is diagonalizable.

    (a) Prove that A is diagonalizable if and only if F^n has a basis consisting only of eigenvalues of A (the "if" part of this should have been done in 6 (e)).

    (b) Prove that if A is diagonalizable, then the eigenvalues of A are the diagonal entries of Λ. *Hint:* use part (e) of question 5 above.

    (c) By inspection, what are the eigenvectors of Λ? Applying part (a) of question 5, what does that tell you about the eigenvectors of A (specifically, what do they have to do with P or P^{-1}?)

    (d) Show that $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ is not diagonalizable.

13. Let $A \in M_2(\mathbb{R})$ be a matrix with no real eigenvalues *Remark:* by problems 4 (e) and 11, A has two complex eigenvalues λ and $\overline{\lambda}$. Since λ is not real, we can write $\lambda = e^{i\theta}$ where $\theta \notin \{0, \pi\}$.

    (a) Show A is diagonalizable if we allow entries of P and Λ to be complex.

    (b) Let $\lambda \in \mathbb{C}$ be an eigenvalue of A and let $\mathbf{v} \in \mathbb{C}^2$ be an eigenvector corresponding to λ with $||\mathbf{v}|| = 1$. Let $P \in M_2(\mathbb{R}$ be $P = (\Re(\mathbf{v}) \;\; \Im(\mathbf{v}))$. Show that $P^{-1}AP \in Q_2$ (and is a rotation matrix).

14. Suppose $A \in M_n(F)$ is diagonalizable.

    (a) Prove that there exists an invertible matrix P and a diagonal matrix Λ such that $A = P\Lambda P^{-1}$. (This is easy, given the definition in problem 12.) Writing A this way is called *diagonalizing* A.

    (b) Show that $A^n = P\Lambda^n P^{-1}$ for $n = 0, 1, 2, \ldots$

    (c) If $\Lambda = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$, find Λ^2, Λ^{13} and Λ^{42}. Generalize this.

    (d) Suppose $A = \begin{pmatrix} 5 & 7 \\ 2 & 3 \end{pmatrix}$. Find A^8 by diagonalizing A.

15. Diagonalize the matrix $\begin{pmatrix} -1 & 4 & -2 \\ -3 & 4 & 0 \\ -3 & 1 & 3 \end{pmatrix}$. You may assume without calculation that the eigenvalues are $\lambda = 1$, $\lambda = 2$ and $\lambda = 3$.

16. Determine whether or not these matrices are diagonalizable:

$$\begin{pmatrix} 7 & 4 & 16 \\ 2 & 5 & 8 \\ -2 & -2 & -5 \end{pmatrix} \qquad \begin{pmatrix} 4 & 0 & 0 \\ 1 & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

17. Denote the owl population and wood rat population in a certain ecosystem at time k by $\mathbf{x}_k = (O_k, R_k)$. Suppose biologists determine that these populations evolve by the equations

$$\begin{cases} O_{k+1} = .5\, O_k + .4\, R_k \\ R_{k+1} = -p\, O_k + 1.1\, R_k \end{cases}$$

where $p > 0$ is some unknown parameter representing the rate of deaths of rats due to predation by owls.

(a) Show that this set of recursive relations is equivalent to the matrix equation $\mathbf{x}_{k+1} = A\mathbf{x}_k$ for some matrix A. Find A.

(b) Show that $\mathbf{x}_k = A^k \mathbf{x}_0$ where $\mathbf{x}_0 = (O_0, R_0)$ represents the initial numbers of owls and rats in the ecosystem.

(c) Explain how you would compute the number of rats in the ecosystem when $k = 100$.

18. A certain beetle has 3 life stages: egg, larva, adult. During each time period, each adult lays 15 eggs and then dies, 80% of the eggs hatch into larvae, and 20% of the larvae survive to adulthood.

Suppose initially a certain garden harbors 50 eggs, 20 larvae, and 10 adults. What will be the situation at the end of one time period? Two periods later? Three? A hundred time periods later? Describe how to approach this problem using the machinery of linear algebra (eigentheory). (But dont actually compute anything, as the algebra is not simple.)

19. Suppose $A \in M_n(\mathbb{R})$ is diagonalizable and all of its eigenvalues are nonnegative. Explain how to define a matrix $B \in M_n(\mathbb{R})$ such that $B^2 = A$.

20. In this problem we define the *exponential* of a matrix. Recall that for a number x, we can write the exponential of x as a power series

$$e^x = \exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \frac{x^4}{4!} + ...$$

which converges for all $x \in \mathbb{R}$. Given a $n \times n$ matrix A, we can define the *exponential* of A to be the $n \times n$ matrix

$$e^A = \exp(A) = \sum_{n=0}^{\infty} \frac{A^n}{n!} = I + A + \frac{1}{2}A^2 + \frac{1}{6}A^3 + \frac{1}{4!}A^4 + \frac{1}{5!}A^5 + ...$$

In fact, this series “converges” for all matrices A (but exactly what is meant by convergence here is beyond the scope of this class).

(a) Suppose Λ is a diagonal matrix. Give a formula for e^Λ.

(b) Prove that if $P \in M_n(F)$ is invertible, then $e^{PQP^{-1}} = Pe^Q P^{-1}$ for any $Q \in M_n(F)$.

(c) Let $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$. Compute e^A by diagonalizing A and using previous parts of this problem.

(d) Compute $\exp \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ using the definition of the matrix exponential.

(e) Give examples of 2×2 matrices A and B which show that in general, $e^{A+B} \neq e^A e^B$.

(f) It is true that if $AB = BA$, then $e^{A+B} = e^A e^B$ Give a sketch of why this is by expanding out all three exponentials with their power series definitions, and match terms on the two sides up to, say, third powers to see what is going on. Why is commutativity necessary?

21. Recall that the differential equation $\frac{dx}{dt} = rx$ has as its solutions $x = x_0 e^{rt}$ where the x_0 is an arbitrary constant (in fact, $x_0 = x(0)$).

(a) Now, let $A \in M_n(F)$ and let $\mathbf{x}_0 \in \mathbb{R}^n$ be given. Define $\mathbf{x} : \mathbb{R} \to \mathbb{R}^n$ by $\mathbf{x}(t) = e^{At}\mathbf{x}_0$. Write the formula for $\mathbf{x}$ out (as a power series) and differentiate it term-by-term to obtain

$$\mathbf{x}'(t) = A\,\mathbf{x}(t).$$

This gives a technique to solve systems of linear differential equations. Let's consider the system

$$\begin{cases} x_1'(t) = 4x_1(t) - 5x_2(t) \\ x_2'(t) = -2x_1(t) + x_2(t) \end{cases}$$

with initial condition $x_1(0) = 3, x_2(0) = -1$.

(b) Write $\mathbf{x}(t) = (x_1(t), x_2(t))$ and explain why the given system is equivalent to the matrix differential equation

$$\mathbf{x}'(t) = \begin{pmatrix} 4 & -5 \\ -2 & 1 \end{pmatrix} \mathbf{x}(t), \quad \mathbf{x}_0 = \begin{pmatrix} 3 \\ -1 \end{pmatrix}.$$

(c) Solve this system by using the formula $\mathbf{x}(t) = e^{At}\mathbf{x}_0$. (The fact that this is a solution follows from (a); that this is the only solution of this system follows from a theorem in the study of differential equations called the "existence/uniqueness theorem".)

15 Spectral theory

15.1 Complex spectral theory

Throughout this subsection, $<,>$ represents the usual inner product on $\mathbb{C}^n$ with associated norm $||\cdot||$.

1. Prove that every eigenvalue of a Hermitian matrix is real.

2. Prove that if $A \in M_n(\mathbb{C})$ is such that $< A\mathbf{v}, \mathbf{v} >= 0$ for all $\mathbf{v} \in \mathbb{C}^n$, then $A = 0$. *Hint:* First show

$$< A\mathbf{v}, \mathbf{w} > = \frac{1}{4}\left(< A(\mathbf{v}+\mathbf{w}), \mathbf{v}+\mathbf{w} > - < A(\mathbf{v}-\mathbf{w}), \mathbf{v}-\mathbf{w} >\right) + i\frac{1}{4}\left(< A(\mathbf{v}+i\mathbf{w}), \mathbf{v}+i\mathbf{w} > - < A(\mathbf{v}-i\mathbf{w}), \mathbf{v}-i\mathbf{w} >\right).$$

(Without the A, this is the content of the Polarization Identities we proved in problem 17 on page 16.) Then explain why the hypothesis of this problem implies $< A\mathbf{v}, \mathbf{w} >= 0$ for all $\mathbf{v}, \mathbf{w}$; we've already discussed in class why ($< A\mathbf{v}, \mathbf{w} >= 0$ for all $\mathbf{v}, \mathbf{w}$) implies $A = 0$.

3. Prove that a matrix $A \in M_n(\mathbb{C})$ is Hermitian if and only if $< A\mathbf{v}, \mathbf{v} > \in \mathbb{R}$ for all $\mathbf{v} \in \mathbb{C}$. *Hint:* use conjugate symmetry of inner products and the dual relations, together with problem 2 of this section.

4. A matrix $A \in M_n(\mathbb{C})$ is called *normal* if $AA^H = A^H A$.
   (a) Show that every Hermitian matrix is normal.
   (b) Show that every unitary matrix is normal.
   (c) Show that for any $A \in M_{mn}(\mathbb{C})$, $A^H A$ and AA^H are both normal.
   (d) Show that A is normal if and only if $||A\mathbf{v}|| = ||A^H\mathbf{v}||$ for every $\mathbf{v} \in \mathbb{C}^n$.
   (e) Show that if A is normal, so is $A - \lambda I$ for any scalar λ.
   (f) Show that if A is normal, then given any eigenvector $\mathbf{v}$ of A corresponding to eigenvalue λ, then $\mathbf{v}$ is also an eigenvalue of A^H corresponding to eigenvalue $\overline{\lambda}$.
   (g) Show that if A is normal, then eigenvectors of A corresponding to different eigenvalues are orthogonal.
   (h) Show that if A is normal and U is unitary, then UAU^H and $U^H AU$ are both normal.

15.2 Real spectral theory

Throughout this subsection, $<,>$ represents the usual inner product on $\mathbb{R}^n$ with associated norm $||\cdot||$.

5. Prove that every eigenvalue of a real symmetric matrix is real. *Hint:* real symmetric matrices are Hermitian matrices. Look at problem 1.

6. (Compare with problem 2)
   (a) Prove that if $A \in M_n(\mathbb{R})$ is such that $< A\mathbf{v}, \mathbf{v} >= 0$ for all $\mathbf{v} \in \mathbb{R}^n$, then it is not necessarily the case that $A = 0$. (Take A to be a rotation.)
   (b) Prove that if $A \in M_n(\mathbb{R})$ is symmetric and is such that $< A\mathbf{v}, \mathbf{v} >= 0$ for all $\mathbf{v} \in \mathbb{R}^n$, then $A = 0$.

7. A matrix $A \in M_n(\mathbb{R})$ is called *normal* if $AA^T = A^T A$.

   (a) Show that every symmetric matrix is normal.

   (b) Show that every orthogonal matrix is normal.

   (c) Show that for any $A \in M_{mn}(\mathbb{R})$, $A^T A$ and AA^T are both normal.

   (d) Show that A is normal if and only if $||A\mathbf{v}|| = ||A^T\mathbf{v}||$ for every $\mathbf{v} \in \mathbb{R}^n$.

   (e) Show that if A is normal, so is $A - \lambda I$ for any scalar λ.

   (f) Show that if A is normal, then given any eigenvector $\mathbf{v}$ of A corresponding to eigenvalue λ, then $\mathbf{v}$ is also an eigenvector of A^T corresponding to eigenvalue λ.

   (g) Show that if A is normal, then eigenvectors of A corresponding to different eigenvalues are orthogonal.

15.3 Positive matrices

8. Let $F = \mathbb{R}$ or $\mathbb{C}$. A Hermitian square matrix $A \in M_n(F)$ is called *positive* if $\mathbf{x}^H A\mathbf{x} \geq 0$ for all $\mathbf{x} \in F^n$. In this problem we prove that the following five items are equivalent:

   i. A is positive;
   ii. A is Hermitian and the eigenvalues of A are all nonnegative;
   iii. there is a positive matrix B such that $B^2 = A$ (we call B the square root of A and write $B = \sqrt{A}$);
   iv. there is a Hermitian matrix B such that $B^2 = A$;
   v. there is a matrix B such that $A = B^H B$.

   (a) Prove (i) implies (ii). *Hint:* let the $\mathbf{x}$ in the definition of positive be an eigenvalue of A.

   (b) Prove (ii) implies (iii). *Hint:* the spectral theorem applies to A.

   (c) It is trivial that (iii) implies (iv) since every positive matrix is Hermitian. Nothing to do here.

   (d) Prove (iv) implies (v). This is also pretty easy.

   (e) Prove (v) implies (i). *Hint:* rewrite the expression $\mathbf{x}^H A\mathbf{x}$ as an inner product, and use the dual relations.

9. Recall that a Hermitian matrix $A \in M_n(F)$ is called *positive definite* if it is positive and if $\mathbf{x}^H A\mathbf{x} = 0$ only when $\mathbf{x} = \mathbf{0}$.

   (a) Prove that A is positive definite if and only if all its eigenvalues are positive.

   (b) Prove that the trace and determinant of any positive definite matrix are positive.

   (c) Define the *upper left submatrices* of A to be

$$A_1 = (a_{11})_{1\times 1}, A_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}_{2\times 2}, A_3 = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}_{3\times 3}, \ldots$$

(in other words, you look at the upper-leftmost $k \times k$ entries of A and call that $k \times k$ matrix A_k). Prove that a Hermitian matrix A is positive definite if and only if all of its upper left submatrices $A_1, A_2, ..., A_n$ have positive determinant.
Note: This is how you actually check whether or not a Hermitian matrix is positive definite.
Hints: To prove the "only if" part, prove by contradiction: suppose that there is a vector $\mathbf{x}$ such that $\mathbf{x}^H A\mathbf{x} \leq 0$. Write $\mathbf{x} = (x_1, x_2, ..., x_k, 0, ..., 0)$ where $x_k \neq 0$ and show that $(x_1, ..., x_k)^H A_k(x_1, ..., x_k) \leq 0$. Explain why this means $\det A_k \leq 0$ (use a previous part of this problem).
To prove the "if" part, use the spectral theorem to diagonalize A by a unitary matrix U. Then write $\mathbf{x} = (x_1, ..., x_n)$, calculate $\mathbf{x}^H A\mathbf{x}$ directly and show why this must be positive.

15.4 Singular value decomposition

10. In class I will prove the *Polar Decomposition Theorem* which says: let $A \in M_n(F)$ where $F = \mathbb{R}$ or $\mathbb{C}$. Then $A = U\sqrt{A^H A}$ for some unitary matrix U. Why is this called the Polar Decomposition Theorem?

11. Find a singular value decomposition of the matrix
$$A = \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix}.$$

15.5 Pseudoinverses and least-squares solutions

12. Let $A \in M_{mn}(F)$ have rank r, and let its singular value decomposition be $A = U\Sigma V^H$. Let U_r be the matrix of the columns of U corresponding to the nonzero singular values of A and let V_r be the matrix of the columns of V corresponding to the nonzero singular values of A. In other words, let
$$\Sigma = \left(\begin{array}{c|c} D & 0 \\ \hline 0 & 0 \end{array}\right)_{m\times n}$$
where D is a diagonal $r \times r$ matrix with the nonzero singular values of A down the diagonal, and partition U and V as
$$\begin{aligned} U &= \left(\begin{array}{ccc|ccc} \mathbf{u}_1 & \cdots & \mathbf{u}_r & \mathbf{u}_{r+1} & \cdots & \mathbf{u}_m \end{array}\right)_{m\times m} & V &= \left(\begin{array}{ccc|ccc} \mathbf{v}_1 & \cdots & \mathbf{v}_r & \mathbf{v}_{r+1} & \cdots & \mathbf{v}_n \end{array}\right)_{n\times n} \\ &= \left(\begin{array}{c|c} U_r & U_{m-r} \end{array}\right) & &= \left(\begin{array}{c|c} V_r & V_{n-r} \end{array}\right) \end{aligned}$$
Show that the equation $A = U\Sigma V^H$ reduces to $A = U_v D V_r^H$ by considering the product of the partitioned matrices. *Remark:* this factorization $A = U_v D V_r^H$ is called the *reduced singular value decomposition* of A.

13. Let $A \in M_{mn}(F)$ and let $A = U_v D V_r^H$ be its reduced singular value decomposition. Define the *pseudoinverse* (a.k.a. *Moore-Penrose pseudoinverse*) of A to be the matrix
$$A^+ = V_r D^{-1} U_r^H.$$

(a) Show that if A is invertible, then $A^+ = A^{-1}$.

(b) Prove that for any $\mathbf{y} \in F^m$, $AA^+\mathbf{y}$ is the projection (with respect to the standard inner product) of $\mathbf{y}$ onto $C(A)$.

Hint: To do this, by uniqueness of orthogonal decomposition, it is sufficient to show $AA^+\mathbf{y} \in C(A)$ and $\mathbf{y} - AA^+\mathbf{y} \in [C(A)]^\perp$.

(c) Prove that for any $\mathbf{y} \in F^n$, $A^+A\mathbf{y}$ is the projection (with respect to the standard inner product) of $\mathbf{y}$ onto $R(A)$.

(d) Prove $A^+AA^+ = A^+$ and $AA^+A = A$.

(e) Suppose $\mathbf{b} \in F^m$ is not in the column space of A (so the system $A\mathbf{x} = \mathbf{b}$ has no solution). Let $\mathbf{x}^+ = A^+\mathbf{b}$. This $\mathbf{x}^+$ is called the *least-squares solution* to $A\mathbf{x} = \mathbf{b}$.

i. Show $\mathbf{x}^+ \in R(A)$.

ii. Show that $\mathbf{b}^+ = A\mathbf{x}^+$ is the projection of $\mathbf{b}$ onto $C(A)$. (As a consequence, we know that $dist(\mathbf{b}^+, \mathbf{b}) \leq dist(A\mathbf{x}, \mathbf{b})$ for any $\mathbf{x} \in F^n$ by problem 8.54 (a).)

iii. Show that $\mathbf{x}^+$ is a solution of $A\mathbf{x} = \mathbf{b}$.

iv. Show that for any solution $\mathbf{u}$ of $A\mathbf{x} = \mathbf{b}$, $||\mathbf{x}^+|| \leq ||\mathbf{u}||$, with equality holding only when $\mathbf{u} = \mathbf{x}^+$.

14. Compute the pseudoinverse of each of these matrices:

$$\begin{pmatrix} 4 & -2 \\ 2 & -1 \\ 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 7 & 1 \\ 0 & 0 \\ 5 & 5 \end{pmatrix}$$

15. Find the least-squares solution $\mathbf{x}^+$ and the corresponding vector $\mathbf{b}^+$ for the system of equations

$$\begin{cases} x + 5y = 4 \\ 3x + y = -2 \\ -2x + 4y = -3 \end{cases}$$

16. Find the equation of the least-squares line that best fits the points $(1, 0)$, $(2, 1)$, $(4, 2)$ and $(5, 3)$.

17. A certain experiment produces the data $(1, 7.9)$, $(2, 5.4)$ and $(3, -.9)$. If you theorize that your model for the experiment should be some function of the form $y = A\sin x + B\cos x$, describe a procedure that you would use to solve for A and B (actually, you would only solve for a least-squares solution for A and B).