1. If $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is a linear transformation, then

   (a) For any vector $\vec{x}$ in $\mathbb{R}^n$, and any vector $\vec{v}$ in $\ker(T)$, then

   $$T(\vec{x}+\vec{v}) = T(\vec{x}) + T(\vec{v}) = T(\vec{x}) + \vec{0} = T(\vec{x}),$$

   So $\vec{x}$ and $\vec{x}+\vec{v}$ have the same image in $\mathbb{R}^m$ under the linear transformation T.

   (b) Conversely, suppose that $\vec{x}_1$ and $\vec{x}_2$ are two vectors in $\mathbb{R}^n$ with the same image $\vec{b}$ in $\mathbb{R}^m$. Then if we set $\vec{v} = \vec{x}_2 - \vec{x}_1$, it's clear that $\vec{x}_1 + \vec{v} = \vec{x}_2$, and

   $$T(\vec{v}) = T(\vec{x}_2 - \vec{x}_1) = T(\vec{x}_2) - T(\vec{x}_1) = \vec{b} - \vec{b} = \vec{0},$$

   so $\vec{v}$ is in the kernel of T.

   (c) If $\vec{b}$ is a vector in $\mathbb{R}^m$ and $\vec{x}_1$ a vector in $\mathbb{R}^n$ with $T(\vec{x}_1) = \vec{b}$ (i.e., $\vec{b}$ is in the image of T, and $\vec{x}_1$ is any one of the vectors in $\mathbb{R}^n$ mapping to $\mathbb{R}^m$), then

   - For any vector $\vec{v}$ in $\ker(T)$, $\vec{x} = \vec{x}_1 + \vec{v}$ is a vector in $\mathbb{R}^n$ such that $T(\vec{x}) = \vec{b}$, by part (a), so all the vectors of the form $\vec{x}_1 + \vec{v}$ with $\vec{v}$ in $\ker(T)$ are solutions to $T(\vec{x}) = \vec{b}$.
   - For any vector $\vec{x}$ which is a solution to $T(\vec{x}) = \vec{b}$, part (b) shows that there is a vector $\vec{v}$ in $\ker(T)$ so that $\vec{x} = \vec{x}_1 + \vec{v}$. In other words, not only (from above) are vectors of the form $\vec{x}_1 + \vec{v}$ solutions to $T(\vec{x}) = \vec{b}$, but *every* solution to the equation $T(\vec{x}) = \vec{b}$ is of this form.

2. The linear transformation $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ is given by

   $$T(x, y) = (6x - 3y,\, 2x - y,\, 4x - 2y),$$

   or (equivalently) the matrix $\begin{bmatrix} 6 & -3 \\ 2 & -1 \\ 4 & -2 \end{bmatrix}$.

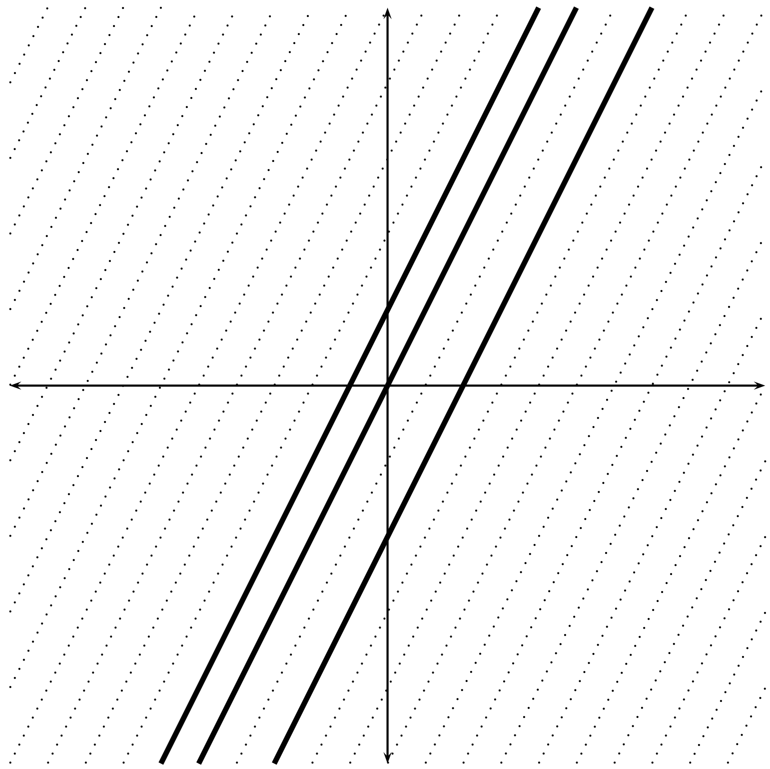
   (a) From the equations, the kernel is the solution to the equations $2x - y = 0$, $6x - 3y = 0$, and $4x - 2y = 0$ which of course are all multiples of the first equation $2x - y = 0$ or $y = 2x$.

   The kernel is therefore the line $y = 2x$ in $\mathbb{R}^2$, which we can also describe as being spanned by the vector $(1, 2)$.

(b) From the equations this is the same as solving $y - 2x = 1$, or $y = 2x + 1$.

(c) This time this is the same as solving $y - 2x = -2$ or $y = 2x - 2$.

(d) Here's a sketch: $\mathbb{R}^2$



The three lines are shown in black, along with some parallel lines, which are dotted.

(e) The lines shown are parallel. All points on the same line are mapped to the same point in $\mathbb{R}^3$ by T, the middle line is the kernel, mapping to $\vec{0}$.

What question (1) tells us is that two points $\vec{x}_2$ and $\vec{x}_1$ of $\mathbb{R}^2$ map to the same point of $\mathbb{R}^3$ if and only if their difference $\vec{x}_2 - \vec{x}_1$ is in the kernel.

Or, if we have any point $\vec{x}$ of $\mathbb{R}^2$, we find all other points of $\mathbb{R}^2$ mapping to the same point $T(\vec{x})$ in $\mathbb{R}^3$ by translating $\vec{x}$ by elements of the kernel, or equivalently translating the kernel by $\vec{x}$. This is the same thing as saying the the lines must all be parallel.

Geometrically, the map T is collapsing $\mathbb{R}^2$ along these parallel lines, with a one dimensional image in $\mathbb{R}^3$.

3. Suppose that A is a 3×2 matrix, and B a 2×3 matrix. The matrix A describes a linear map from $\mathbb{R}^2$ to $\mathbb{R}^3$, and the matrix B a linear map from $\mathbb{R}^3$ to $\mathbb{R}^2$.

(a) Since A is a map from $\mathbb{R}^2$ to $\mathbb{R}^3$, the dimension of $\text{im}(A)$ can be at most 2. Therefore since AB is the composition of the linear map given by B with the linear map given by A, $\dim(\text{im}(AB))$ can also be at most 2, but for a 3×3 matrix AB to be invertible, we have to have $\dim(\text{im}(AB)) = 3$.

An equivalent argument is to note that since B is a linear map from $\mathbb{R}^3$ to $\mathbb{R}^2$, its kernel $\ker(B)$ has to be at least one dimensional, i.e., there is some vector $\vec{v}$, $\vec{v} \neq \vec{0}$ with $B\vec{v} = \vec{0}$. But that means that $AB\vec{v} = \vec{0}$ too, so there are nonzero vectors in the kernel of AB, and that also means that AB can't be invertible. (These were the properties (vi) and (viii) of the last assignment.)

(b) An easy example is $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ with product

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

More generally (and the point of part (c)) is that as long as $\dim(\text{im}(A)) = 2$ (i.e., the linear map given by A is surjective) and $\dim(\text{im}(B)) = 2$ (i.e., that the linear map given by B is injective) the map BA will have 2 dimensional image, and hence (by one of the characterizations of invertible $n \times n$ matrices) be invertible.

(c) If BA is invertible, then its kernel is just the zero vector $\vec{0}$, so the kernel of A must be the zero vector too. (If $\vec{v}$ is in $\ker(A)$, then $\vec{v}$ is also in $\ker(BA)$, so $\ker(BA) = \vec{0}$ implies $\ker(A) = \vec{0}$.) So, if BA is invertible, we must have $\dim(\ker(A)) = 0$.

On the other hand, if BA is invertible, then $\dim(\text{im}(BA)) = 2$, so we must have $\dim(\text{im}(B)) = 2$, since anything in the image of BA is in the image of B. (For any vector $\vec{w}$ in $\mathbb{R}^2$, if there is a vector $\vec{v}$ in $\mathbb{R}^2$ with $\vec{w} = BA\vec{v}$, then $\vec{w}$ is the result of applying the matrix B to the vector $A\vec{w}$ in $\mathbb{R}^3$, and so $\vec{w}$ is in the image of B).

This shows that the dimension of $\text{im}(B)$ must be at least 2, but it can also be at most 2 since $\text{im}(B)$ is contained in $\mathbb{R}^2$. Therefore $\dim(\text{im}(B)) = 2$.

But now, by the "rank-nullity" theorem, $\dim(\ker(B)) = 3 - 2 = 1$.

4. If T_α is rotation counterclockwise by α, and T_θ rotation counterclockwise by angle θ in $\mathbb{R}^2$, then

(a) The composition $T_\alpha \circ T_\theta$ is just rotation counterclockwise by angle $\alpha + \theta$.

The composition $T_\theta \circ T_\alpha$ is also rotation counterclockwise by angle $\alpha + \theta$, so this is an example of two matrices where the order of multiplication doesn't matter. (This doesn't usually happen – even if A and B are both $n \times n$ matrices, the products AB and BA usually aren't the same.)

(b) Multiplying, the matrix for $T_\alpha \circ T_\theta$ is $\begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} =$

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

(c) The matrix for rotation by $\alpha + \theta$ is $\begin{bmatrix} \cos(\alpha+\theta) & -\sin(\alpha+\theta) \\ \sin(\alpha+\theta) & \cos(\alpha+\theta) \end{bmatrix}$.

(d) Since the matrices represent the same linear transformation, their entries must be the same. Looking at the first columns of these matrices, this means that we must have

$$\begin{aligned} \cos(\alpha+\theta) &= \cos(\alpha)\cos(\theta) - \sin(\alpha)\cos(\theta), \text{ and} \\ \sin(\alpha+\theta) &= \cos(\alpha)\sin(\theta) + \sin(\alpha)\cos(\theta). \end{aligned}$$

Looking at the second columns gives the same identities, but in the opposite order.

(e) If we compose the transformation T_θ three times: $T_\theta \circ T_\theta \circ T_\theta$, the resulting linear transformation is rotation by 3θ. Multiplying, we have

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}^3 = \begin{bmatrix} \cos^3(\theta) - 3\cos(\theta)\sin^2(\theta) & -(3\cos^2(\theta)\sin(\theta) - \sin^3(\theta)) \\ 3\cos^2(\theta)\sin(\theta) - \sin^3(\theta) & \cos^3(\theta) - 3\cos(\theta)\sin^2(\theta) \end{bmatrix}$$

Comparing with the entries for the matrix for rotation by 3θ, we get the identities

$$\begin{aligned} \cos(3\theta) &= \sin^3(\theta) - 3\cos(\theta)\sin^2(\theta), \text{ and} \\ \sin(3\theta) &= 3\cos^2(\theta)\sin(\theta) - \sin^3(\theta). \end{aligned}$$

5. Matrix squaring and images.

(a) A simple example is the matrix $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$.

Its image and kernel are both the x-axis, i.e., the linear subspace spanned by $(1, 0)$. One consequence (that we've seen before) is that A^2 must be the zero matrix.

(b) If A is an $n \times n$ matrix with $A^2 = 0$ (the zero matrix), then we have $\text{im}(A) \subseteq \ker(A)$, i.e., the image of A must be contained in the kernel of A.

(We've gone through this type of argument before: suppose that $\vec{w}$ is a vector in the image of A, i.e., that there is some vector $\vec{v}$ with $\vec{w} = A\vec{v}$. Since $A\vec{w} = A(A\vec{v}) = A^2\vec{v} = \vec{0}$ (since $A^2 = 0$), we see that $\vec{w}$ is in the kernel of A.)

Let k be the dimension of the kernel of A, and i the dimension of the image of A. By the "rank-nullity" theorem, we have $i + k = n$.

On the other hand, since $\text{im}(A)$ is contained in $\ker(A)$, we must have $i \leq k$.

This means that i can be at most half the size of n. There are many ways to combine the two equations above to get this, but one way is to note that $2i = i + i \leq i + k = n$, so $2i \leq n$ or $i \leq n/2$.

If $n = 10$, we get $i \leq 10/2 = 5$, and so $\dim(\text{im}(A)) \leq 5$.

(c) Suppose that A is an $n \times n$ matrix with $A^2 = A$.

If $\vec{w}$ is a vector in the image of A, there is a vector $\vec{v}$ with $\vec{w} = A\vec{v}$. Multiplying both sides of this equation by A on left (i.e., putting both of these vectors through A) we get

$$A\vec{w} = AA\vec{v} = A\vec{v} = \vec{w}.$$

The middle equality is because $A^2 = A$, and the last one because $A\vec{v} = \vec{w}$, since we chose chose $\vec{v}$ with this property.

In other words, for any vector $\vec{w}$ in the image of A, $A\vec{w} = \vec{w}$.

If $\vec{w}$ is also in the kernel of A, then we have $\vec{0} = A\vec{w} = \vec{w}$, so $\vec{w} = \vec{0}$. I.e., the only vector which $\ker(A)$ and $\text{im}(A)$ have in common is the zero vector.

There are nontrivial examples of matrices with this property. For example, if A_m is the matrix for "projection onto a line of slope m in $\mathbb{R}^2$":

$$A_m = \begin{bmatrix} \frac{1}{1+m^2} & \frac{m}{1+m^2} \\ \frac{m}{1+m^2} & \frac{m^2}{1+m^2} \end{bmatrix}$$

from homework 3, 4(a), has this property. In fact, if A is an $n \times n$ matrix with $A^2 = A$, then it turns out that A must be the matrix of projection onto some subspace of $\mathbb{R}^n$, something we will verify later in the course.