Names: ____________________

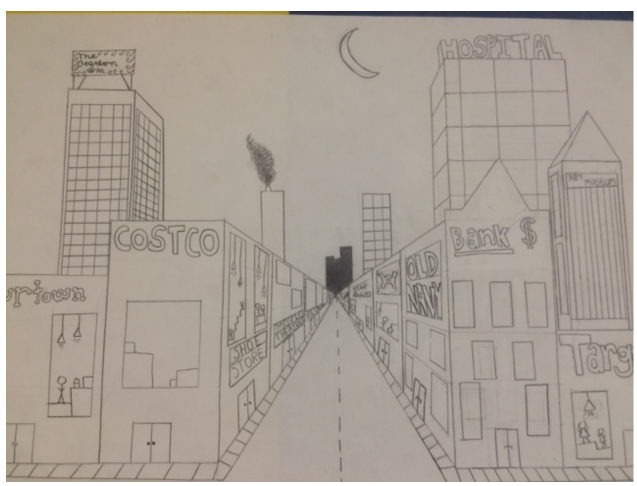
NetIDs: ____________________

Discussion section: D ________ Group number: ________

MATH 257 - WORKSHEET 3: MATRIX MULTIPLICATION

Today, your team will try to figure out: can **matrix multiplication** model the distortions of seeing a three-dimensional sign from an angle?

Original to distorted



Distorted sign viewed from the street

(1) **Matrix multiplication is simpler than it looks!** Let's see why.

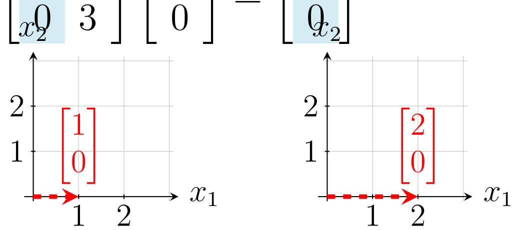
Let $\mathbf{e}_j$ represent the unit vector in the x_j direction. It's all zeros except the jth element, which is 1. We'll just talk about $\mathbb{R}^2$ to start, so $\mathbf{e}_1 = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$ and $\mathbf{e}_2 = \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$.

We'll also need this arbitrarily chosen matrix M:

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

(a) Multiplying M by $\mathbf{e}_1$ gives the *first* column of M. **Why?** As a group, discuss how matrix multiplication works to come up with a brief explanation, and **write** it below.

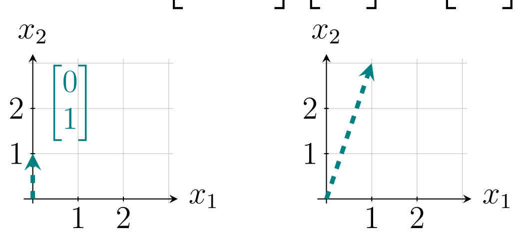
$$\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$



Using the row-column rule for matrix multiplication, each row is matched up against the column vector. Row by row, the 1 at the top of the column vector picks out the first number in the row. The other number is zeroed out, so the sum of the two is just the first number in the row. Putting all the rows' first numbers together, we get the matrix's first column.

(b) Similarly, multiplying M by $\mathbf{e}_2$ gives the *second* column of M. **Without doing any calculations, write in** $M\mathbf{e}_2$ below and **plot** it below right. As above, discuss and **explain** why this works.

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

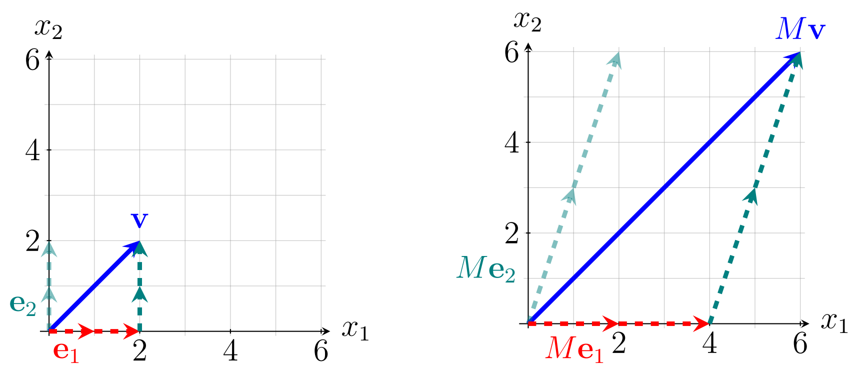


This is just like the previous situation, except the 1 is in the second element of the column vector, so it picks out the second number in the row. Putting all the rows' second numbers together, we get the matrix's second column.

(c) Consider the arbitrarily chosen vector $\mathbf{v} = \left[\begin{smallmatrix}2\\2\end{smallmatrix}\right]$, which is a a linear combination of $\mathbf{e}_1$ and $\mathbf{e}_2$. As it turns out, $M\mathbf{v}$ can be formed from a linear combination of $M\mathbf{e}_1$ and $M\mathbf{e}_2$ with *the same coefficients*.

Compute $M\mathbf{v}$ directly to confirm that this is true. As a group, discuss the matrix multiplication rule so you can **write a brief explanation** of why this trick works.

$$\mathbf{v} = \boxed{2}\,\mathbf{e}_1 + \boxed{2}\,\mathbf{e}_2 \qquad M\mathbf{v} = \boxed{2}\,M\mathbf{e}_1 + \boxed{2}\,M\mathbf{e}_2$$



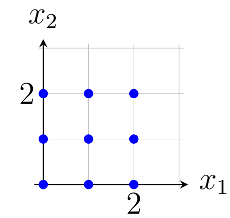
This property is essentially that matrix multiplication is *linear*. It arises from the fact that matrix multiplication consists of addition and multiplication. The coefficient can first be incorporated into the vector during matrix multiplication and later factored out again. In equations: $M\mathbf{v} = M(c_1\mathbf{e}_1 + c_2\mathbf{e}_2) = c_1 M\mathbf{e}_1 + c_2 M\mathbf{e}_2$. For $\mathbf{v} = \left[\begin{smallmatrix}2\\2\end{smallmatrix}\right]$, $c_1 = 2$ and $c_2 = 2$.

Let's look at that in detail in the matrix multiplication. When the first row of M is lined up vertically against $\mathbf{v}$, the result is 2 times the first element of the row plus 2 times the second element—which is the same as 2 times what we got from $\mathbf{e}_1$ plus 2 times what we got from $\mathbf{e}_2$. The same applies to the result from the second row. Since each row of the result is 2 times the result from $\mathbf{e}_1$ plus 2 times the result from $\mathbf{e}_2$, we can move the addition outside the vector and factor out the two 2s to find that the *entire* result is 2 times the result from $\mathbf{e}_1$ plus 2 times the result from $\mathbf{e}_2$.

(d) **Reflection**. Take a few minutes to pause the problem solving and discuss what you have learned from it.

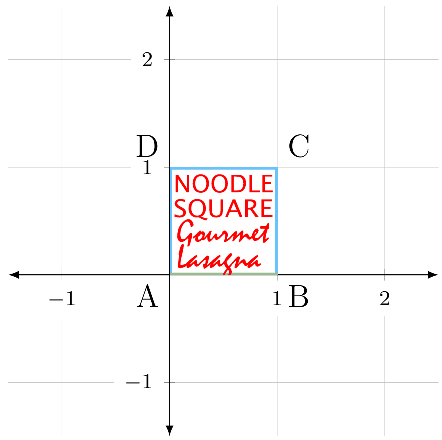
- How can you generalize the unit vector shortcut to $m \times n$ matrices?
- Given any vector $\mathbf{u} \in \mathbb{R}^2$, how can you quickly plot $M\mathbf{u}$ using what you learned above?
- Given a square grid G in $\mathbb{R}^2$ (see example below right), how can you quickly plot the transformed grid points $\{M\mathbf{g} : \mathbf{g} \in G\}$?

Summarize your discussion below in 3–4 bullet points.



- The jth unit vector in $\mathbb{R}^n$, $\mathbf{e}_j$, picks out the jth column of M, which is in $\mathbb{R}^m$.
- The columns of M tell you result of multiplying each unit vector by M. Since the coefficients stay the same, if $\mathbf{u} = \left[\begin{smallmatrix} c_1 \\ c_2 \end{smallmatrix}\right]$, we can plot $\mathbf{u}$ by adding tip-to-tail: c_1 times the first column of M plus c_2 times the second column of M. **Big takeaway: to multiply a matrix by a column vector, you can add the columns of the matrix multiplied by the corresponding components of the vector.**
- The grid points are equally spaced in the x_1 direction by some factor of $\mathbf{e}_1$, and in the x_2 direction by some factor of $\mathbf{e}_2$ (in the example, that factor is 1 for both directions). The transformed grid points will be equally spaced by the same factors of the *transformed* unit vectors $M\mathbf{e}_1$ and $M\mathbf{e}_2$. Each transformed grid point will keep the same "coordinates", but in terms of the *transformed* unit vectors instead of the original ones.

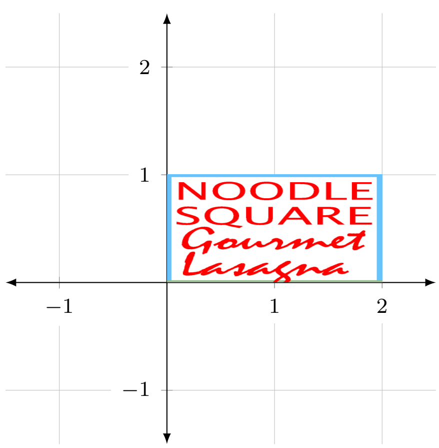
(2) Now that we've explored a single matrix in depth, let's characterize a few different matrices. **Using as little calculation as possible**, figure out what the unit square sign $ABCD$ plotted below will look like when transformed by each matrix. Draw the result on the adjacent empty plot. Keep track of each corner's final location by labeling it with the corresponding letter.



(a) $\begin{bmatrix} a & 0 \\ 0 & 1 \end{bmatrix}, a = 2$

Describe how this matrix distorts the sign. What does a represent?

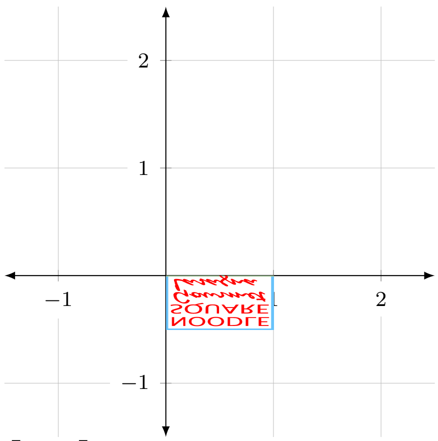
$\boxed{\text{This matrix stretches the sign by a factor of } a \text{ in the } x \text{ direction.}}$



(b) $\begin{bmatrix} 1 & 0 \\ 0 & b \end{bmatrix}, b = -0.5$

Describe how this matrix distorts the sign. What does b represent? How does the negative value affect the figure?

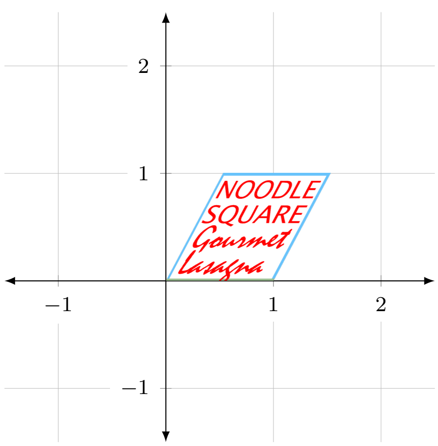
This matrix flips the sign over the x axis (the negative sign in b) and compresses it by a factor of 2 in the y direction (the 0.5 in b).



(c) $\begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix}, s = 0.5$

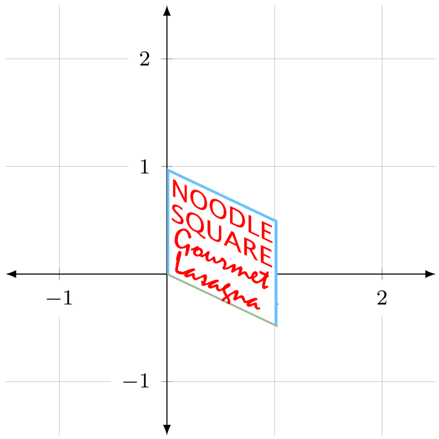
Describe how this matrix distorts the sign. What does s represent?

This matrix skews the sign in the x direction until the formerly vertical sides have slope $1/s = 2$.

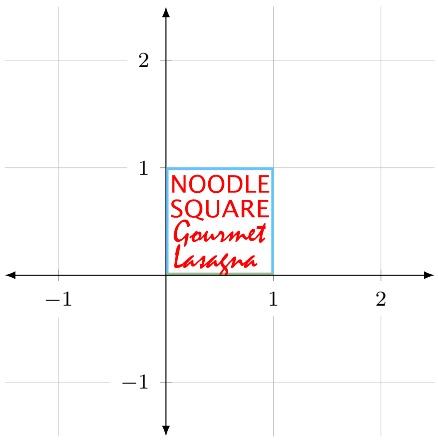


(d) $\begin{bmatrix} 1 & 0 \\ t & 1 \end{bmatrix}$, $t = -0.5$
Describe how this matrix distorts the sign. What does t represent?

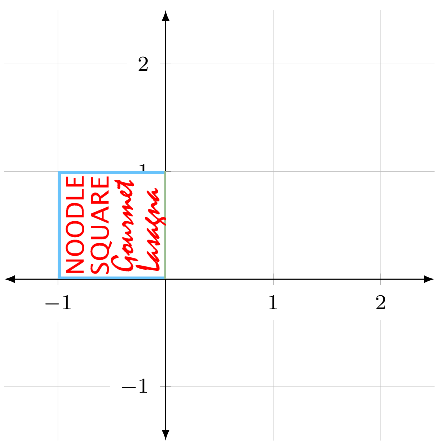
This matrix skews the sign in the $-y$ direction until the formerly horizontal sides have slope $t = -0.5$.



(3) The sign below, a unit square, has been distorted in four different ways. For each, discuss as a group: **can matrix multiplication create** this distortion? If yes, **state** the appropriate matrix. If no, **explain why not.**



(a) The sign has been rotated by $\pi/2$ radians counterclockwise around the origin.

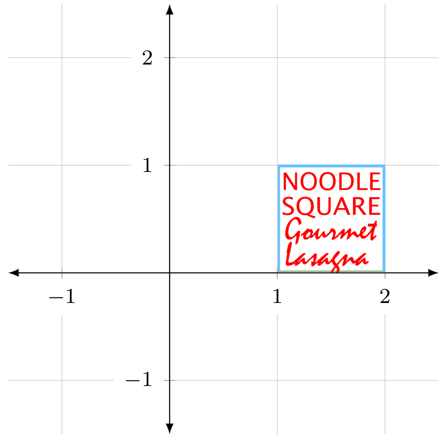


Possible. The x unit vector is rotated to $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and the y unit vector is rotated to $\begin{bmatrix} -1 \\ 0 \end{bmatrix}$, so we can put those in the columns of M:

$$M = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

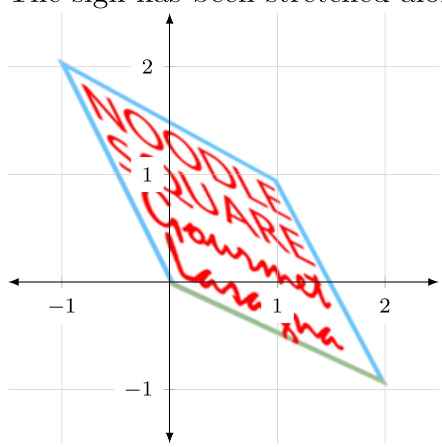
The whole grid rotates with them. You can check that the far corner is still the sum of the two closer corners.

(b) The sign has been translated one unit to the right.



Not possible. Look at the bottom-left corner, which started at $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$. It doesn't matter what matrix you left-multiply that by, you'll always get $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$—never $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. The bottom-left corner of the sign has to stay "attached" to the origin, so translation is not achievable by matrix multiplication.

(c) The sign has been stretched along a diagonal.

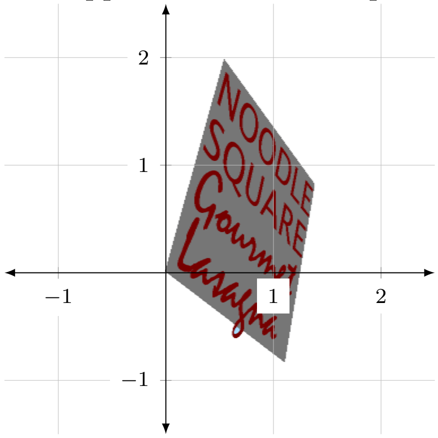


Possible. The x unit vector gets stretched to $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ and the y unit vector gets stretched to $\begin{bmatrix} -1 \\ 2 \end{bmatrix}$, so we can put those in the columns of M:

$$M = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

The whole grid stretches with them. You can check that the corner at $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is still the sum of the two other nonzero corners.

(d) The sign has been drawn in perspective, as if it were seen from an angle. Note that opposite sides are not parallel because they converge to a "vanishing point".



Not possible. Any matrix will transform the original unit square grid into a grid like you discussed in part (1)(d). The transformed grid points will be aligned in rows and columns parallel to the two transformed unit vectors. Since the top and bottom of our original sign both started out parallel to the original x unit vector, they have to end up parallel to the transformed x unit vector—*parallel*. This sign's top and bottom are not parallel, so a matrix multiplication can't produce it from our original sign.

(4) **Reflection**. Take a few minutes to pause the problem solving and discuss what you have learned from it.

- Is it possible to use matrix multiplication to create the distortions of a sign seen three–dimensionally from an angle?
- What can matrix multiplication do? What can't it do?
- This problem only considered matrix multiplication from $\mathbb{R}^2$ to $\mathbb{R}^2$. How can you generalize what you learned to higher dimensions?

Summarize your discussion below in 3–4 bullet points.

- No. As we saw before, signs seen three–dimensionally from an angle are most accurately drawn in perspective, which involves positioning edges that are parallel in reality along lines that intersect at a "vanishing point". Matrix multiplication can't transform parallel lines into lines that aren't parallel. (Although if the three-dimensional sign were seen from far away at a high zoom, the vanishing point can be so far off–screen that the distorted sides are essentially parallel. In that case, matrix multiplication actually does work.)
- Matrix multiplication can stretch and skew and reflect relative to the origin, but it has to keep parallel lines parallel. These possibilities and constraints can be captured with the following image: matrix multiplication can transform any parallelogram with a corner at the origin into any other parallelogram with the same corner at the origin.
- In three dimensions, you can imagine the same image as above except with parallelograms replaced by parallelepipeds. A parallelepiped is the three-dimensional version of a parallelogram; it's like a cube except with parallelogram-shaped faces instead of square faces. Four dimensions are hard to picture, but you can roughly imagine the parallelogram picture continuing into higher dimensions.

Nice work! Your team has explored and determined whether matrix multiplication can model the distortion of looking at a three-dimensional sign from an angle.
Now that you have developed a graphical picture of matrix multiplication in an applied context, practice using it on the following abstract problems.

(5) Let $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Using a graphical picture of what left–multiplying by A does, determine $A^n\mathbf{v}$ for any natural number n. Then determine just the matrix, A^n.

A keeps $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ the same and sends $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ to $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. This slides every point over 1 unit to the right for every unit it is above the x axis (and to the left below the x axis). Doing this to $\mathbf{v}$ once gives $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$, which doesn't change its height (or any vector's height) above the x axis, so doing it n times just gives $\begin{bmatrix} 1+n \\ 1 \end{bmatrix}$. We can conclude that A^n slides everything over n units to the right for every unit it is above the x axis, which is represented by the matrix $\begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}$.

(6) Suppose you are told there is a mystery 2×2-matrix A such that $A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$ and $A \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

(a) Can you determine A just from this information?

One strategy is to think about A in terms of its columns: suppose that $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 \end{bmatrix}$, where both $\mathbf{a}_1$, $\mathbf{a}_2$ are column vectors in $\mathbb{R}^2$.

Then, as we saw in this worksheet, we can each express matrix multiplication as a linear combination of the columns of A. Then set it equal to what we are given:

$$A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \mathbf{a}_1 + \mathbf{a}_2 = \begin{bmatrix} -2 \\ 2 \end{bmatrix}, \qquad A \begin{bmatrix} -1 \\ 1 \end{bmatrix} = -\mathbf{a}_1 + \mathbf{a}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Now we just need to solve for $\mathbf{a}_1$ and $\mathbf{a}_2$. Summing these two equations together, we get

$$2\mathbf{a}_2 = (\mathbf{a}_1 + \mathbf{a}_2) + (-\mathbf{a}_1 + \mathbf{a}_2) = \begin{bmatrix} -2 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix},$$

i.e. $\mathbf{a}_2 = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$. Then we have $\mathbf{a}_1 = \begin{bmatrix} -2 \\ 2 \end{bmatrix} - \mathbf{a}_2 = \begin{bmatrix} -\frac{3}{2} \\ \frac{3}{2} \end{bmatrix}$. Therefore, we have $A = \begin{bmatrix} -\frac{3}{2} & -\frac{1}{2} \\ \frac{3}{2} & \frac{1}{2} \end{bmatrix}$.

(b) Is always enough to know $A\mathbf{v}$ and $A\mathbf{w}$ for two distinct vectors $\mathbf{v}$ and $\mathbf{w}$ in order to determine A?

No. Assume that $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$. Then similarly, we should obtain that

$$\mathbf{a}_1 + \mathbf{a}_2 = \mathbf{b}, \qquad 2\mathbf{a}_1 + 2\mathbf{a}_2 = 2\mathbf{b},$$

for some constant vector $\mathbf{b}$. However, one of equations is obviously redundant so it is not enough to determine A. In general, we need $\mathbf{v}$ and $\mathbf{w}$ to be linearly independent to avoid redundant equations.