Math 211P
Brussel
Fall 2008

Handout 1

This handout outlines the basics you'll need on vectors, matrices, and determinants.

Words you must know: *orthogonal, position vector, orthonormal basis, unit vector, dot product, orthogonal transformation, determinant, length/area/volume distortion.*

You must be able to: *find the length of a vector, interpret scalar multiples of vectors, explain how a matrix transforms a space, compute the dot product of two vectors, tell whether two vectors are orthogonal, tell if a set of vectors is an orthonormal basis, compute the action of a matrix on a vector, tell if a matrix is orthogonal, compute the determinant of a matrix, compute the distortion caused by a matrix.*

We'll use these concepts all semester long.

Basics on Vectors

$\mathbb{R} = \mathbb{R}^1$ is the set of real numbers, or "1-space". $\mathbb{R}^2$ and $\mathbb{R}^3$ are the symbols for "2-space" and "3-space", ordinary two and three-dimensional space. We will think of $\mathbb{R}^2$ and $\mathbb{R}^3$ in two ways: as a set of points and as a set of vectors. Let's concentrate first on $\mathbb{R}^3$, since, as far as vectors and points are concerned, $\mathbb{R}^2$ is just $\mathbb{R}^3$ with zero third coordinate. For more detail, read the text §12.1.

A *point* P in $\mathbb{R}^3$ is a location. A *vector* $\vec{v}$ is an arrow. A vector is completely determined by its length and direction. In particular, *vectors have no location.* This means if two vectors drawn in different places on coordinate axes have the same length and direction, they are the same vector. Vectors can be used to describe anything that is completely determined by its magnitude (the length) and direction. Thus we have vectors for spatial displacement (*displacement vectors*), for velocity, and for force.

To make measurements and mark points, we install a 3-dimensional grid, or *coordinate system.* Our first coordinate system is the usual x, y, and z-axes. We'll see other systems later.

Every point is determined by its x, y, and z *coordinates.* For example, $P = (3, 2, -5)$ is the point at $+3$ on the x-axis, $+2$ on the y-axis, and -5 on the z-axis. By the Pythagorean theorem, the distance of P from the origin is $\sqrt{3^2 + 2^2 + (-5)^2} = \sqrt{38}$.

A *unit vector* is any vector whose length is 1. Every vector can be written in terms of the unit vectors $\vec{\imath}$, $\vec{\jmath}$, and $\vec{k}$, which point in the positive x, y, and z-directions, respectively. For example, $\vec{v} = 3\vec{\imath}+2\vec{\jmath}-5\vec{k}$ is the arrow that runs $+3$ in the x-direction, $+2$ in the y-direction, and -5 in the z-direction. We also use the notation $\vec{v} = \langle 3, 2, -5\rangle$. The numbers 3,2, and -5 are $\vec{v}$'s *(scalar) components.* By the Pythagorean theorem, $\vec{v}$'s length is $\sqrt{3^2 + 2^2 + (-5)^2} = \sqrt{38}$.

What's the difference between a point and a displacement vector? A point's components are distances measured from the origin. A displacement vector's components are just displacements, from nowhere in particular. It's possible to make the two ideas coincide by always drawing displacement vectors with their tails at the origin. We then call them *position vectors* and use the same notation as for points. We say that the origin is the vectors' *basepoint.*

The set of all vectors in $\mathbb{R}^3$ is called a *vector space*, and denoted also by $\mathbb{R}^3$. The three unit vectors $\vec{\imath}$, $\vec{\jmath}$, and $\vec{k}$ form an *orthonormal basis* for this vector space, meaning

(a) they are orthogonal, or perpendicular, to each other (ortho);
(b) they are unit vectors (normal); and
(c) they assign unique components to every vector (basis).

When based at the origin, they form a little cube of volume 1. To do all of this for $\mathbb{R}^2$, ignore the third component: The two vectors $\vec{\imath}$ and $\vec{\jmath}$ form an orthonormal basis for $\mathbb{R}^2$. When based at the origin, they form a nice little square $\square$ of area 1. For $\mathbb{R}^1$ the vector $\vec{\imath}$ is a basis, and at the origin it makes a unit interval ·—·.

The *sum* $\vec{u}+\vec{v}$ of two vectors is the vector you get by connecting them head to tail. The *scalar multiple* $a\vec{v}$ of a vector is the vector $\vec{v}$ "scaled" by a. That is, it's parallel to $\vec{v}$, but the length is multiplied by a. Note it points in the same direction if $a > 0$, the opposite direction if $a < 0$. The *zero vector* $\vec{0}$ is the vector of length 0.

Notation: $\mathbb{R}^3$ will mean either the point space or the vector space, depending on context. Similarly in two dimensions.

Basics on Matrices

A *matrix* is a rectangular array of numbers or symbols. An $m \times n$ matrix is one with m rows and n columns. Here are four examples:

$$\underset{3\times 3}{\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}} \qquad \underset{3\times 2}{\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{pmatrix}} \qquad \underset{2\times 2}{\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}} \qquad \underset{1\times 3}{\begin{pmatrix} a & b & c \end{pmatrix}}$$

Each matrix is made up of adjacent *column vectors*, that is, vectors written in column notation. The column vectors of the first matrix are $\langle a_1, a_2, a_3\rangle, \langle b_1, b_2, b_3\rangle$, and $\langle c_1, c_2, c_3\rangle$.

In linear algebra you'll learn how an $m \times n$ matrix A sends vectors or points $\vec{v}$ in $\mathbb{R}^n$ to vectors or points $\vec{w}$ in $\mathbb{R}^m$. We write

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

We have the relations $A(\vec{u}+\vec{v}) = A\vec{u} + A\vec{v}$ and $A(c\vec{v}) = c(A\vec{v})$, for $c \in \mathbb{R}$. Since $\mathbb{R}^n$ goes to $\mathbb{R}^m$, we say A *maps* $\mathbb{R}^n$ into $\mathbb{R}^m$, and write $A : \mathbb{R}^n \to \mathbb{R}^m$. A *square* matrix takes $\mathbb{R}^n$ to itself, and we then say it *transforms* $\mathbb{R}^n$. So a 3×3 matrix transforms $\mathbb{R}^3$ by moving its points (or vectors) around. Following are the basics on matrices acting on vectors or points.

Dot product. A 1×2 matrix maps a vector or point in $\mathbb{R}^2$ to a vector or point in $\mathbb{R}^1$, i.e., a number. Here's how it works for the *matrix* $(a_1 \;\; a_2)$ and the *vector or point* $(x, y) \in \mathbb{R}^2$:

$$\begin{pmatrix} a_1 & a_2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = a_1x + a_2y$$

In the 1×3 case,

$$\begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = a_1x + a_2y + a_3z$$

In this case we view these matrices as vectors and write $\langle a_1, a_2, a_3\rangle \cdot \langle x, y, z\rangle = a_1x + a_2y + a_3z$. This important operation on vectors is called the *dot product.* We'll see why it's so useful in class. Weirdly, *the dot product of two vectors is zero if and only if they are perpendicular.* We'll justify this later.

Transformations of space. A 1×1 matrix is just a number, (a). The matrix (a) transform $\mathbb{R}^1$ (the x-axis) by scalar multiplication; it either stretches or shrinks by the factor a. Thus the unit interval is transformed into an interval of length $|a|$. A vector $x\vec{\imath}$ is transformed to $ax\vec{\imath}$.

The rule for computing the effect of a 2×2 (square) matrix on $(x, y) \in \mathbb{R}^2$:

$$\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = (a_1x + b_1y,\; a_2x + b_2y)$$

Here (x, y) could be either a point or a vector. Note the first component on the right side is the dot product of the matrix's first row vector and (x, y), the second is dot product of the second row vector and (x, y).

The origin $(0,0)$ stays fixed: $A\vec{0} = \vec{0}$. The matrix maps it to itself. Now look what happens to the displacement vectors $\vec{\imath} = \langle 1,0\rangle$ and $\vec{\jmath} = \langle 0,1\rangle$: They go to the column vectors $A\vec{\imath} = a_1\vec{\imath} + b_1\vec{\jmath}$ and $A\vec{\jmath} = a_2\vec{\imath} + b_2\vec{\jmath}$. So the little unit square $\square$ formed by the edges $\vec{\imath}$ and $\vec{\jmath}$ transforms into the parallelogram $\lozenge$ formed by the edges $a_1\vec{\imath} + b_1\vec{\jmath}$ and $a_2\vec{\imath} + b_2\vec{\jmath}$.

It is important to see here that A sends the standard unit vectors $\vec{\imath}$ and $\vec{\jmath}$ to A's first and second column vectors, respectively. Now imagine the whole plane $\mathbb{R}^2$ is "tiled" into unit squares, like graph paper. The matrix does the same thing to each square, and the whole space is affected just like the little square is. This is how the matrix transforms space. Note if the column vectors are parallel then A *flattens* $\mathbb{R}^2$ into a line.

A similar rule holds for 3×2 matrices on $\mathbb{R}^2$ and 3×3 matrices on $\mathbb{R}^3$. In the 3×2 case the rule is

$$\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = (a_1x + b_1y,\ a_2x + b_2y,\ a_3x + b_3y)$$

In this case the matrix takes the unit square to the (possibly degenerate) parallelogram in $\mathbb{R}^3$ formed by $\langle a_1, a_2, a_3\rangle$ and $\langle b_1, b_2, b_3\rangle$, the columns of the matrix. As $A : \mathbb{R}^2 \to \mathbb{R}^3$, it takes the plane $\mathbb{R}^2$ to the *span* of the column vectors in $\mathbb{R}^3$, which could be either a plane, if the vectors point in different directions, or a line, if they are parallel, or even a point, if both vectors are zero.

Similarly in the 3×3 case the three components are the dot products of (x, y, z) with the first, second, and third rows, in order. A 3×3 matrix transforms the nice little cube formed by the edges $\vec{\imath}$, $\vec{\jmath}$, and $\vec{k}$ into a (possibly degenerate) parallelepiped, formed by the matrix's column vectors. By extension we can see what happens to all of $\mathbb{R}^3$ by tiling it into little unit cubes.

This is an important result. **To see how a matrix transforms space, look at its column vectors.** Make sure you see how to do this.

Example. From now on we view 2×2 (or 3×2 or 3×3) matrices as operators that transform space in a way that we can understand by looking at the column vectors. The transformation could be a rotation or reflection, a stretching, a shrinking, or some combination. For example, look at the matrices

$$i.\ \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \qquad ii.\ \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad iii.\ \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \qquad iv.\ \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

Each matrix transforms the unit square into a different parallelogram. By looking at the column vectors, check the following: The first matrix stretches the square into the rectangle with edges $\langle 2,0\rangle$ and $\langle 0,1\rangle$. The second rigidly rotates it clockwise $90°$. The third bends it into the parallelogram formed by $\langle 1,1\rangle$ and $\langle 0,1\rangle$, with acute angle $45°$ and side lengths $\sqrt{2}$ and 1. The fourth completely flattens it into a line segment.

Extending to all of $\mathbb{R}^2$: The first matrix stretches $\mathbb{R}^2$ by a factor of 2 in the x-direction. The second rigidly rotates $\mathbb{R}^2$ clockwise $90°$. The third rotates the x-axis counterclockwise by $45°$ while leaving the y-axis fixed. Its first column $\langle 1,1\rangle$ has length $\sqrt{2}$, not 1, so this matrix also stretches the x-axis by a factor of $\sqrt{2}$ before it rotates it. The fourth rotates the y-axis clockwise into the x-axis, smashing everything into a line.

Orthogonal transformations. 2×2 and 3×3 matrices transform $\mathbb{R}^2$ and $\mathbb{R}^3$ while fixing the origin. A square matrix is an *orthogonal transformation* if it transforms $\mathbb{R}^2$ and $\mathbb{R}^3$ *rigidly*, without any stretching or shrinking (while fixing the origin). Like a rotation or a reflection. Since it is rigid, an orthogonal transformation preserves the distances between points and the angles between lines, and also the lengths of vectors and the angles between vectors. In particular, the little unit square in $\mathbb{R}^2$ or unit cube in $\mathbb{R}^3$ *keeps its shape.*

Important concept: Since it is rigid, *an orthogonal transformation takes one orthonormal basis to another.* So *to check that a matrix represents an orthogonal transformation, show the column vectors are* *mutually perpendicular, of length 1.* As before, you picture the motion by looking at the three columns. Oh, right. How do you check that the columns are mutually perpendicular? Use the dot product.

Here are two examples of orthogonal transformations in $\mathbb{R}^3$.

A. The rigid motion that reverses the direction of the z-axis but fixes the x and y-axes. Since this matrix by definition takes $\vec{\imath}$, $\vec{\jmath}$, $\vec{k}$ to $\vec{\imath}, \vec{\jmath}, -\vec{k}$, the column vectors are $\vec{\imath}$, $\vec{\jmath}$, $-\vec{k}$, so the matrix is

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Check

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = (a, b, -c)$$

So it does what it's supposed to.

B. The rigid motion that rotates the xy-plane 45 degrees counterclockwise about the z-axis is given by the matrix

$$N = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Check that it takes $\vec{\imath}, \vec{\jmath}, \vec{k}$ to $\frac{1}{\sqrt{2}}\vec{\imath} + \frac{1}{\sqrt{2}}\vec{\jmath}$, $-\frac{1}{\sqrt{2}}\vec{\imath} + \frac{1}{\sqrt{2}}\vec{\jmath}$, and $\vec{k}$, respectively. If you draw them you'll see they're mutually perpendicular unit vectors (do it!), as desired. Computationally, you see this by first checking that they all have length 1, using the length formula. Then use the dot product:

$$\langle 1/\sqrt{2}, 1/\sqrt{2}, 0\rangle \cdot \langle -1/\sqrt{2}, 1/\sqrt{2}, 0\rangle = -1/2 + 1/2 + 0 = 0$$
$$\langle 1/\sqrt{2}, 1/\sqrt{2}, 0\rangle \cdot \langle 0, 0, 1\rangle = 0$$
$$\langle -1/\sqrt{2}, 1/\sqrt{2}, 0\rangle \cdot \langle 0, 0, 1\rangle = 0$$

Determinants

Each 2×2 matrix transforms $\mathbb{R}^2$ somehow, and in particular turns a square, located anywhere in $\mathbb{R}^2$, into a parallelogram, maybe even a *degenerate* flat one. The area must change by some factor, and we call this *area distortion* produced by the matrix. Thus in $\mathbb{R}^2$ the area distortion is the area of the transformed unit square. A 3×2 will similarly cause an area distortion as $\mathbb{R}^2$ is transformed into a plane in $\mathbb{R}^3$. In $\mathbb{R}^3$, the volume of the transformed unit cube (a parallelepiped) is called the *volume distortion.* In $\mathbb{R}^1$, the length of the transformed unit interval is called the *length distortion.*

Computing the distortion caused by a square matrix is easy. In $\mathbb{R}^1$, the unit interval is stretched/shrunk into a new interval of length $|a_1|$, the absolute value of a_1. So the length distortion is $|a_1|$. In $\mathbb{R}^2$ the area of the new parallelogram determined by a 2×2 matrix is the absolute value of the matrix's *determinant.* The determinant of a 2×2 matrix is

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1 b_2 - a_2 b_1$$

So the area distortion is $|a_1 b_2 - a_2 b_1|$ (absolute value). Let's check this on the four matrices in the example. The first has area 2. The second didn't change shape, and the area stays 1. The third was squeezed into a parallelogram with acute angle 45°, but one side got longer, and the result is, no area change. The fourth was degenerate, so the area is 0. Seems right.

In $\mathbb{R}^3$ the same result holds. A 3×3 matrix transforms the the unit cube into a parallelepiped whose volume is the absolute value of the matrix's determinant. The determinant of a 3×3 matrix is

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1 \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}$$

This is called expansion on the first row. We get the same answer if we expand on the first column:

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}$$

In fact we can expand on any row or column, but if the row or column is even, we must multiply by -1. Write it out if you don't believe it.

For example we ask for the volume distortion produced by the matrix

$$\begin{pmatrix} 1 & 0 & -1 \\ 1 & 1 & 2 \\ 1 & 0 & 1 \end{pmatrix}.$$

It takes the unit cube into the parallelepiped formed by the edges $\langle 1,1,1\rangle$, $\langle 0,1,0\rangle$, and $\langle -1,2,1\rangle$, all coming from the origin. Try to picture it. I can't. What's the new volume? It's $1(1) - 0(1-2) + (-1)(-1) = 2$. The volume is 2. So the volume distortion is 2.

The area distortion caused by a 3×2 matrix is a little more complicated. The rule is that the distortion caused by the 3×2 matrix

$$\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{pmatrix}$$

is the square root of the sum of the squares of determinants of all of the 2×2 submatrices,

$$\left(\begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}^2 + \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix}^2 + \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}^2 \right)^{1/2}.$$

A matrix transforms the whole space, not just the unit interval/square/cube. Thus it transforms *every* geometric figure into another figure. Each figure is distorted, so its size is distorted. To compute the new figure's size you just multiply by the matrix's determinant.

Remember this result: **The determinant of a square matrix measures length/area/volume distortion.** For a 3×2 matrix it's the square root of the sum of the squares of all of the 2×2 subdeterminants. Feel free to marvel at how measuring the distortion could possibly be this easy. *Why* does it work? Let's not discuss that question now.

Here's an example. The unit circle has area π. To squish the circle into an ellipse that hits the $+x$-axis at 3 and the $+y$-axis at 2, use the matrix $\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$. The area of this ellipse is the area of the circle times the determinant of the matrix, which is 6. So the area of the ellipse is 6π. Generalizing, we see that the area of an ellipse whose x and y-intercepts are $\pm a$ and $\pm b$ is πab. Try the same thing in $\mathbb{R}^3$, squishing the sphere into an ellipsoid. You should get a nice formula for the volume of an ellipsoid.

By the way, since rigid motions don't distort area or volume, *an orthogonal transformation always has determinant* ± 1. But as the third matrix in the example shows, determinant ± 1 does not mean the transformation is orthogonal. Volume can be preserved even though the box's shape is distorted.

Let's review.

- The length of a vector $\langle a, b, c \rangle$ is $(a^2 + b^2 + c^2)^{1/2}$.
- A unit vector has length 1.
- Two vectors $\vec{u}$ and $\vec{v}$ are perpendicular if and only if $\vec{u} \cdot \vec{v} = 0$.
- To see how a matrix transforms space, look at its column vectors.
- A matrix is orthogonal if its columns form an orthonormal basis.
- The determinant of a square matrix measures the distortion in (length, area, volume) of the resulting transformation of space.
- The square root of the sum of the squares of the 2×2 subdeterminants of a 3×2 matrix measures the area distortion of the resulting transformation of space.

The dot product couldn't be more important. Square matrix multiplication will be used to analyze distortions and rotations of curves, surfaces, and solids, like in the ellipse example above. The determinant will be used to compute length, area and volume distortion when integrating over parameterized curves and surfaces and when integrating in different coordinate systems, like the polar, cylindrical, and spherical systems.