Math 43 Review Notes - Chapter 7

Linear Transformations

A linear transformation T is a function that takes vectors as its inputs and has vectors as its outputs. It must satisfy the following two properties:
(1) $T(\mathbf{v}+\mathbf{w}) = T(\mathbf{v}) + T(\mathbf{w})$ for any vectors $\mathbf{v}$ and $\mathbf{w}$.
(2) $T(c\mathbf{v}) = c\,T(\mathbf{v})$ for any vector $\mathbf{v}$ and any number c.

An important consequence of this definition is that $T(\mathbf{0}) = \mathbf{0}$. This can be seen by taking $c = 0$ in property (2).

Notation: Sometimes we use the notation $T : \mathbb{R}^n \to \mathbb{R}^m$ to mean that T takes vectors with n components and outputs vectors with m components. (remember, $\mathbb{R}^n$ stands for all vectors with n components, each of which is a real number.)

The columns of the $n \times n$ identity matrix I form what we call the *standard basis* for $\mathbb{R}^n$. For example, when $n = 2$ they are $(1, 0)$ and $(0, 1)$ and when $n = 3$ they are $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$.

Example: Determine which of the following are linear transformations.

(a) $T(\begin{bmatrix} x \\ y \\ z \end{bmatrix}) = \begin{bmatrix} 3x - y + z \\ 2x + 5y \end{bmatrix}$

(b) $T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} 3x - 4 \\ y \end{bmatrix}$

(c) $T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} \sqrt{x} \\ y \end{bmatrix}$

Solution:
(a) This is a linear transformation. To show it carefully, verify that properties (1) and (2) hold for generic vectors.
(b) This is *not* a linear transformation because

$$T(\begin{bmatrix} 0 \\ 0 \end{bmatrix}) = \begin{bmatrix} 3(0) - 4 \\ 0 \end{bmatrix} = \begin{bmatrix} -4 \\ 0 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

A linear transformation must always have $T(\mathbf{0}) = \mathbf{0}$.
(c) This is *not* a linear transformation because property (2) doesn't work.

$$T(2\begin{bmatrix} 1 \\ 0 \end{bmatrix}) = T(\begin{bmatrix} 2 \\ 0 \end{bmatrix}) = \begin{bmatrix} \sqrt{2} \\ 0 \end{bmatrix}, \qquad 2\,T(\begin{bmatrix} 1 \\ 0 \end{bmatrix}) = 2\begin{bmatrix} \sqrt{1} \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

As a rule of thumb, linear transformations are those functions whose outputs have components that look similar to the components in (a). Transformations having a square root term, a nonzero constant term like the 4 in (b), an x^2, a cosine function, or in general any nonlinear terms, will not be linear transformations.

Example: Suppose T is a linear transformation with $T(\begin{bmatrix} 1 \\ 2 \end{bmatrix}) = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$, and $T(\begin{bmatrix} 3 \\ 3 \end{bmatrix}) = \begin{bmatrix} 4 \\ 7 \end{bmatrix}$.
Find $T(\begin{bmatrix} 4 \\ 5 \end{bmatrix})$, $T(\begin{bmatrix} 9 \\ 9 \end{bmatrix})$, and $T(\begin{bmatrix} 14 \\ 16 \end{bmatrix})$.

Solution: Use properties (1) and (2).

$$T(\begin{bmatrix} 4 \\ 5 \end{bmatrix}) = T(\begin{bmatrix} 1 \\ 2 \end{bmatrix}) + T(\begin{bmatrix} 3 \\ 3 \end{bmatrix}) = \begin{bmatrix} 3 \\ 5 \end{bmatrix} + \begin{bmatrix} 4 \\ 7 \end{bmatrix} = \begin{bmatrix} 7 \\ 12 \end{bmatrix}.$$

$$T(\begin{bmatrix} 9 \\ 9 \end{bmatrix}) = T(3\begin{bmatrix} 3 \\ 3 \end{bmatrix}) = 3\,T(\begin{bmatrix} 3 \\ 3 \end{bmatrix}) = 3\begin{bmatrix} 4 \\ 7 \end{bmatrix} = \begin{bmatrix} 12 \\ 21 \end{bmatrix}.$$

$$T(\begin{bmatrix} 14 \\ 16 \end{bmatrix}) = T(2\begin{bmatrix} 1 \\ 2 \end{bmatrix} + 4\begin{bmatrix} 3 \\ 3 \end{bmatrix}) = 2\,T(\begin{bmatrix} 1 \\ 2 \end{bmatrix}) + 4\,T(\begin{bmatrix} 3 \\ 3 \end{bmatrix}) = 2\begin{bmatrix} 3 \\ 5 \end{bmatrix} + 4\begin{bmatrix} 4 \\ 7 \end{bmatrix} = \begin{bmatrix} 22 \\ 38 \end{bmatrix}.$$

Matrix of a Linear Transformation

One very important example of a linear transformation is multiplication by a matrix. For, example the transformation

$$T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+3y \\ 2x+4y \end{bmatrix}$$

is a linear transformation. The important fact is that *every linear transformation can be written in this way, as a matrix times a vector.*

To find the matrix of $T : \mathbb{R}^n \to \mathbb{R}^m$, replace each column of the $n \times n$ identity matrix I with T of that column.

Example: Find the matrix A associated to the following linear transformations.

(a) $T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} 2x-3y \\ x \end{bmatrix}$

(b) $T(\begin{bmatrix} x \\ y \\ z \end{bmatrix}) = \begin{bmatrix} x+y+z \\ 2x-3y-4z \end{bmatrix}$

Solution:

(a) $T(\begin{bmatrix} 1 \\ 0 \end{bmatrix}) = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad T(\begin{bmatrix} 0 \\ 1 \end{bmatrix}) = \begin{bmatrix} -3 \\ 0 \end{bmatrix}.$ Therefore $A = \begin{pmatrix} 2 & -3 \\ 1 & 0 \end{pmatrix}$.

(b) $T(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad T(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}) = \begin{bmatrix} 1 \\ -3 \end{bmatrix}, \quad T(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}) = \begin{bmatrix} 1 \\ -4 \end{bmatrix}.$ Therefore $A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & -3 & -4 \end{pmatrix}$.

Shortcut: The first column of A consists of the coefficients of the x terms, the second column has the coefficients of the y terms, etc.

Different Bases

The above example gives the matrix with respect to the standard basis. Sometimes we may need the matrix with respect to different bases. Suppose $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis for the input vectors and $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ is a basis for the output vectors. To find the matrix A, compute $T(\mathbf{v}_1)$, $T(\mathbf{v}_2)$, etc., then write each of the outputs in terms the $\mathbf{w}$'s. The coefficients of the $\mathbf{w}$'s give the entries of A.

Example: Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 1 \\ 3 \\ 9 \end{bmatrix}$, $\mathbf{w}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\mathbf{w}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, and let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be the linear transformation $T(\begin{bmatrix} x \\ y \\ z \end{bmatrix}) = \begin{bmatrix} x+y \\ y+z \end{bmatrix}$. Find the matrix of T with respect to the the bases $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ and $\{\mathbf{w}_1, \mathbf{w}_2\}$.

Solution: First compute

$$T(\mathbf{v}_1) = \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \quad T(\mathbf{v}_2) = \begin{bmatrix} 3 \\ 6 \end{bmatrix}, \quad T(\mathbf{v}_3) = \begin{bmatrix} 4 \\ 12 \end{bmatrix}.$$

It is easy to see that the first vector is $2\mathbf{w}_1$ and the second vector is $3\mathbf{w}_2$. To write the third vector in terms of $\mathbf{w}_1$ and $\mathbf{w}_2$ you can either try by trial and error, or solve the following equation:

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 4 \\ 12 \end{bmatrix}$$

The columns of the matrix here are the basis vectors $\mathbf{w}_1$ and $\mathbf{w}_2$. The solution is $a = -4$, $b = 8$. Thus the third vector is $-4\mathbf{w}_1 + 8\mathbf{w}_2$. The table on the left gives the coefficents of the $\mathbf{w}$'s. It helps us see what A is.

	$\mathbf{v}_1$	$\mathbf{v}_2$	$\mathbf{v}_3$
$\mathbf{w}_1$	2	0	-4
$\mathbf{w}_2$	0	3	-8

$$A = \begin{pmatrix} 2 & 0 & -4 \\ 0 & 3 & 8 \end{pmatrix}$$

Geometry of Linear Transformations

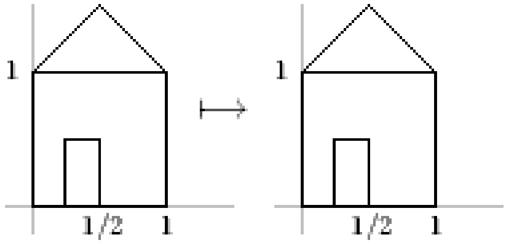
We can understand ordinary functions better by graphing them. We can't graph most linear transformations, but we can instead see how they transform geometric shapes.

The transformations are completely determined by what they do to the standard basis vectors. In two dimensions we consider what happens to the basis vector $(1,0)$ along the x-axis and the basis vector $(0,1)$ along the y-axis. These vectors may be stretched/shrunk and/or rotated by the linear transformation. The rotation gives new x and y directions for the transformed picture, while a stretching/shrinking of $(1,0)$ will stretch/shrink the original picture in the new x direction, and a stretching/shrinking of $(0,1)$ will stretch/shrink the picture in the new y direction.

Example: Determine how the shape of the house changes under the following linear transformations.

(a) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

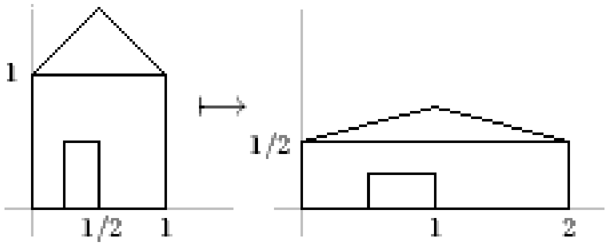
$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} 0 \\ 1 \end{bmatrix}$



The matrix is the identity matrix. The output is the same as the input.

(b) $\begin{pmatrix} 2 & 0 \\ 0 & .5 \end{pmatrix}$

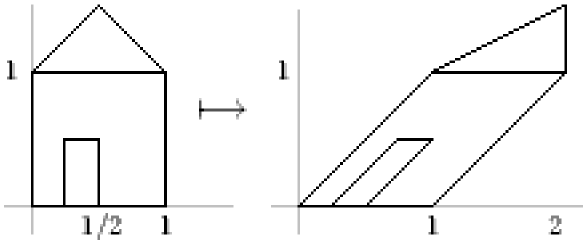
$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} 0 \\ .5 \end{bmatrix}$



Here the basis vector along the x-axis gets magnified by a factor of 2, while the basis vector along the y-axis gets shrunk to half its size. Thus the picture is stretched in the x direction and crunched in the y direction.

(c) $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$

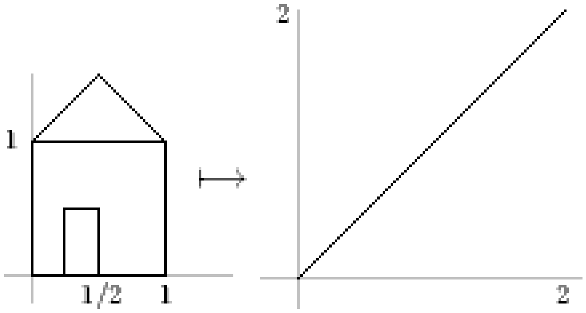
$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} 1 \\ 1 \end{bmatrix}$



Here the basis vector along the x-axis is unchanged, while the basis vector along the y-axis is rotated by 45° and stretched a bit (the length of $(0,1)$ is 1, while the length of $(1,1)$ is $\sqrt{2}$). The effect is called a shear.

(d) $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$

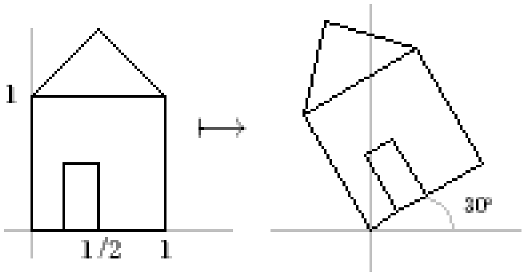
$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} 1 \\ 1 \end{bmatrix}$



Here both basis vectors are rotated into the line $y = x$. So the entire house is "smooshed" onto that line. The peak of the house gets sent to $(2, 2)$.

(e) $\begin{pmatrix} \cos 30^\circ & -\sin 30^\circ \\ \sin 30^\circ & \cos 30^\circ \end{pmatrix}$

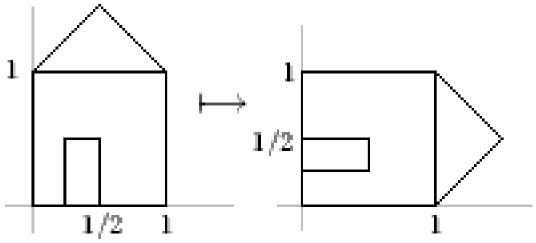
$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} \cos 30^\circ \\ \sin 30^\circ \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} -\sin 30^\circ \\ \cos 30^\circ \end{bmatrix}$



Here using some trigonometry we see that each basis vector is rotated counterclockwise by 30°. Therefore the entire house is rotated by 30°.

(f) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

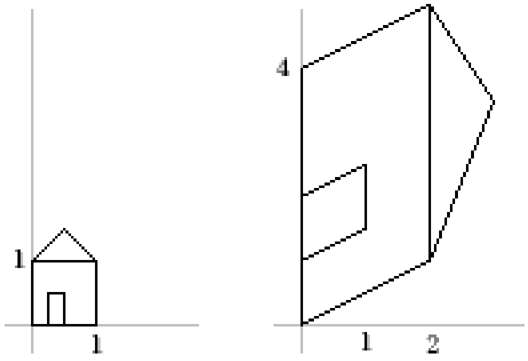
$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} 1 \\ 0 \end{bmatrix}$



Here both the basis vectors are swapped. Thus every vector has its x and y components swapped. The original y-axis becomes the new x-axis, and the original x-axis becomes the new y-axis.

(g) $\begin{pmatrix} 0 & 2 \\ 4 & 1 \end{pmatrix}$

$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 0 \\ 4 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} 2 \\ 1 \end{bmatrix}$



Here the new x-axis is the old y-axis and the new y-axis is the line through $(4, 1)$. Both directions are stretched.

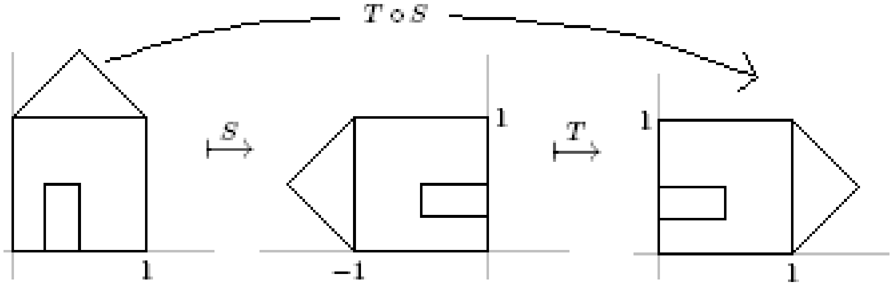
From these pictures we can see why these transformations are called linear: straight lines are transformed into straight lines or points. A nonlinear transformation might transform them into curves. Notice also that in all cases the transformed house touches the origin $(0, 0)$. This is because $T(\mathbf{0}) = \mathbf{0}$ for any linear transformation.

Compositions

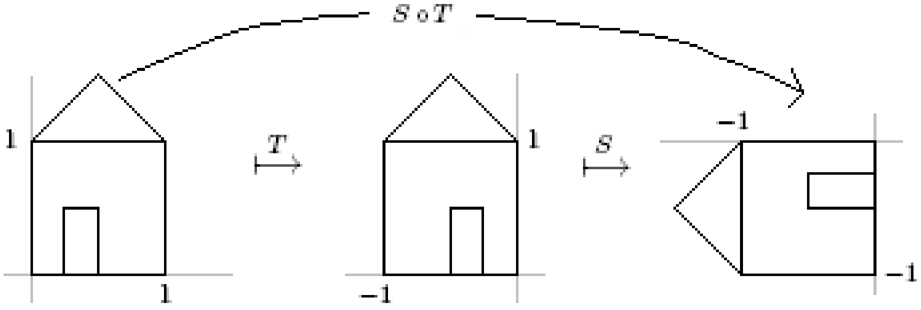
The *composition* of two linear transformations S and T, denoted $T \circ S$, is the linear transformation obtained by first doing S and then doing T. If the matrix of S is A and the matrix of T is B, then the linear transformation $T \circ S$ has matrix BA.

Example: Suppose S and T are linear transformations with matrices $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$.

The transformation S rotates the house by 90°, while T reflects the house about the y-axis. The linear transformation $T \circ S$ has matrix $BA = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$



The composite effect is a coordinate swap (see the second to last example on the previous page). Compare this to the linear transformation $S \circ T$ which has matrix $AB = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$



The composite effect is a reflection about the origin. Notice that this is different from $T \circ S$ (this has to do with the fact that order matters when multiplying matrices), so order matters in compositions.

Inverse Transformation

The *inverse transformation* T^{-1} of a linear transformation T undoes the effect of T. If the matrix of T is A, then the matrix of T^{-1} is given by A^{-1}.

Kernel

The *kernel* of a linear transformation consists of all the input vectors that give an output of $\mathbf{0}$.

For example, suppose $T(\begin{bmatrix} x \\ y \\ z \end{bmatrix}) = \begin{bmatrix} x+y \\ 3z \end{bmatrix}$.

In order for the output to be $\mathbf{0}$, we must have $x + y = 0$ and $3z = 0$, which becomes $y = -x$ and $z = 0$. Thus the kernel consists of all vectors of the form $(x, -x, 0)$, or equivalently, all multiples of $(1, -1, 0)$.

Another way to find the kernel is to find the matrix A of T. The kernel is then just the nullspace of A. This is because the nullspace of A is by definition all the vectors $\mathbf{v}$ for which $A\mathbf{v} = \mathbf{0}$, and since $T(\mathbf{v}) = A\mathbf{v}$, we see that the kernel of T and the nullspace of A are really the same thing.

In the example, the matrix of T is $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$. Find the nullspace, and see that it gives the same result.

Range

The *range* of a linear transformation consists of all the output vectors.

For example, suppose $T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} x \\ 2x \\ x+y \end{bmatrix}$.

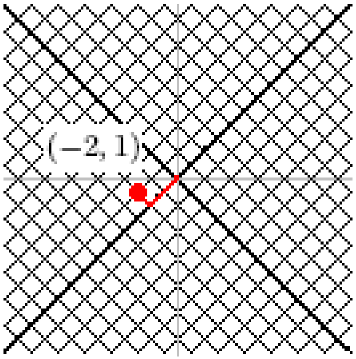
Write the output vector as $x \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + y \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. Thus the range consists of all linear combinations of the vectors $(1, 2, 1)$ and $(0, 0, 1)$.

Another way to find the range is to find the matrix A of T. The range is then just the column space of A. Remember that the column space consists of all vectors of the form $A\mathbf{v}$. (This is one of several ways to view the column space.) Since $T(\mathbf{v}) = A\mathbf{v}$, we see that the column space of A and the range of T are really the same thing.

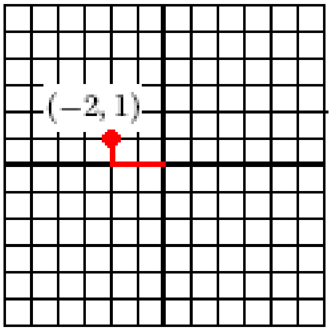
In the example, the matrix of T is $A = \begin{pmatrix} 1 & 0 \\ 2 & 0 \\ 1 & 1 \end{pmatrix}$. Find the column space, and see that it gives the same result.

Coordinate Systems

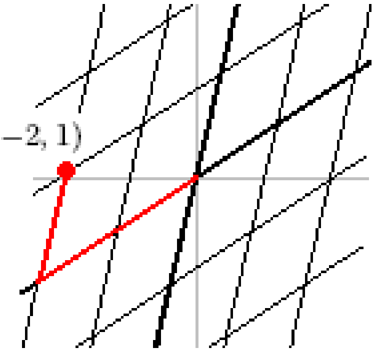
It is often useful to use coordinate systems different from the usual xy coordinates. Every coordinate system is specified by a basis. For example, in the two-dimensional examples below, the coordinate system in the middle is the usual xy coordinate system. Its basis is the standard basis. The left coordinate system is the usual one rotated by 45°, and the one on the right is a stranger one which nevertheless has its uses. In each, the point $(2, 1)$ is indicated.



Basis: $\begin{bmatrix} \cos 45° \\ \sin 45° \end{bmatrix}, \begin{bmatrix} -\sin 45° \\ \cos 45° \end{bmatrix}$



Basis: $\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}$



Basis: $\begin{bmatrix} 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \end{bmatrix}$

Change of Basis Matrix

Given a coordinate system and its basis, we make a matrix M whose columns are the basis vectors. For example, for the coordinate system on the right $M = \begin{pmatrix} 3 & 1 \\ 2 & 4 \end{pmatrix}$.

Continuing this example: given a point with coordinates $(2, 1)$ in the slanted coordinate system, to find its standard coordinates, multiply by M.

$$\begin{pmatrix} 3 & 1 \\ 2 & 4 \end{pmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}.$$

Given a point with standard coordinates $(2, 3)$, to find its coordinates in the slanted coordinate system, multiply by M^{-1}.

$$\frac{1}{10} \begin{pmatrix} 4 & -1 \\ -2 & 3 \end{pmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}.$$

Change of Basis and Linear Transformations

We want to know how the matrix of a linear transformation changes when we change the basis. The answer is simple: the matrix of the linear transformation in the new coordinates is given by $B = M^{-1}AM$ where A is the usual matrix with respect to standard coordinates, and M is the change of basis matrix.

To see why this works, notice that $B\mathbf{v} = M^{-1}A(M\mathbf{v})$. The term $M\mathbf{v}$ translates v into standard coordinates. Then we apply the linear transformation to it by multiplying by A. Finally, multiplying by M^{-1} translates back to the new coordinate system.

Example: Let T be the linear transformation $T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} x + 2y \\ 5x + 4y \end{bmatrix}$. Find the matrix of T with respect to the basis $\{(1, 2), (3, 5)\}$.

Solution: The change of basis matrix is $M = \begin{pmatrix} 1 & 3 \\ 2 & 5 \end{pmatrix}$ and the matrix of T with respect to the standard basis is $A = \begin{pmatrix} 1 & 2 \\ 5 & 4 \end{pmatrix}$. Therefore the matrix of T with respect to this basis is

$$B = M^{-1}AM = \begin{pmatrix} -5 & 3 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 5 & 4 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 2 & 5 \end{pmatrix} = \begin{pmatrix} 14 & 40 \\ -3 & -9 \end{pmatrix}$$

The matrix of the linear transformation in the above basis is rather awful. We would like to know what the best basis to use would be. In other words, what basis will make the matrix of the linear transformation as simple as possible? The simplest matrix we can hope for is a diagonal matrix.

Recall diagonalization gives $A = S\Lambda S^{-1}$, where Λ is a diagonal matrix. Solve this equation for Λ to get $\Lambda = S^{-1}AS$. This is in the form $B = M^{-1}AM$ where $B = \Lambda$ and $M = S$. Therefore to find a basis in which the linear transformation is given by a diagonal matrix, we use the basis which consists of the eigenvectors of A. Notice that for this to work, A must be diagonalizable.

Example: Let T be the linear transformation in the example above. Find a basis in which T is given by a diagonal matrix.

Solution: The eigenvalues of A are -1 and 6 with corresponding eigenvectors $(-1, 1)$ and $(2, 5)$. Therefore the desired basis is $\{(-1, 1), (2, 5)\}$.

Approximation of Dominant Eigenvalues and Eigenvectors

The *dominant eigenvalue* of a matrix is the eigenvalue with largest absolute value.* An eigenvector for a dominant eigenvalue is called a *dominant eigenvector.* The dominant eigenvalue and its eigenvectors are often important in applications.

We will use what is called the *power method* to approximate the dominant eigenvalue and eigenvectors.† Start with a unit vector x_0.‡ Let

$$\mathbf{x_1} = \frac{A\mathbf{x_0}}{\|A\mathbf{x_0}\|}\ ,\quad \mathbf{x_2} = \frac{A\mathbf{x_1}}{\|A\mathbf{x_1}\|}\ ,\quad \mathbf{x_3} = \frac{A\mathbf{x_2}}{\|A\mathbf{x_2}\|}\ ,\ \ldots$$

The vectors $\mathbf{x_1}$, $\mathbf{x_2}$, ... eventually get closer and closer to the unit dominant eigenvector and the terms $A\mathbf{x_1}\cdot\mathbf{x_1}$, $A\mathbf{x_2}\cdot\mathbf{x_2}$, ... eventually get closer and closer to the dominant eigenvalue. When the values of consecutive terms are very close together, the approximation is likely good, so this tells us when to stop.

The method above is best for use on a computer or calculator. If you need to do the computation by hand, finding the lengths of the vectors can be cumbersome. Therefore use the following modification for hand calculations:

$$\mathbf{y_1} = A\mathbf{x_0}\ ,\quad \mathbf{y_2} = A\mathbf{y_1}\ ,\quad \mathbf{y_3} = A\mathbf{y_2}\ ,\ \ldots$$

To convert from the $\mathbf{y}$'s to the $\mathbf{x}$'s turn the $\mathbf{y}$ into a unit vector. In general, $\mathbf{x_n} = \mathbf{y_n}/\|\mathbf{y_n}\|$. The last vector we get when we decide to stop computing is an approximation to a dominant eigenvector. To approximate the unit dominant eigenvector, divide the last vector by its length.

We can also approximate the dominant eigenvalue using the $\mathbf{y}$'s. Recall from above that $A\mathbf{x_n}\cdot\mathbf{x_n}$ is an approximation to the dominant eigenvalue. We have

$$A\mathbf{x_n}\cdot\mathbf{x_n} = \frac{\mathbf{y_{n+1}}\cdot\mathbf{y_n}}{\mathbf{y_n}\cdot\mathbf{y_n}}$$

Example: Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$. Let $\mathbf{x_0} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

$$\mathbf{y_1} = A\mathbf{x_0} = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

$$\mathbf{y_2} = A\mathbf{y_1} = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}\begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 7 \\ 9 \end{bmatrix}$$

$$\mathbf{y_3} = A\mathbf{y_2} = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}\begin{bmatrix} 7 \\ 9 \end{bmatrix} = \begin{bmatrix} 25 \\ 39 \end{bmatrix}$$

$$\mathbf{x_3} = \frac{\mathbf{y_3}}{\|\mathbf{y_3}\|} = \frac{1}{\sqrt{2146}}\begin{bmatrix} 25 \\ 39 \end{bmatrix} \approx \begin{bmatrix} .5397 \\ .8419 \end{bmatrix}$$

$$A\mathbf{x_2}\cdot\mathbf{x_2} = \frac{\mathbf{y_3}\cdot\mathbf{y_2}}{\mathbf{y_2}\cdot\mathbf{y_2}} = \frac{(25,39)\cdot(7,9)}{(7,9)\cdot(7,9)} = \frac{526}{130} \approx 4.05$$

The actual unit dominant eigenvector is $(2,3)/\sqrt{13} \approx (.5547, .8321)$. We see that $\mathbf{x_3}$ is already a fairly good approximation. The actual dominant eigenvalue is 4. We see that $A\mathbf{x_2}\cdot\mathbf{x_2}$ is a fairly good approximation. Note that the method for hand computation is not good for use on computers (or by hand if you're computing lots of terms) because the numbers quickly become huge. That's why we divided each vector by its length in the original form of the method.

---

*If the eigenvalue with largest absolute value occurs more than once, then it is not considered to be a dominant eigenvalue. For example, if $\det(A-\lambda I)$ works out to $(\lambda+5)(\lambda-1)$, then the eigenvalues are -5 and 1. Therefore -5 is the dominant eigenvalue since it has the largest absolute value. If on the other hand, $\det(A-\lambda I)$ works out to $(\lambda-2)^2(\lambda-1)$, then the eigenvalues are $2, 2, 1$. However there is no dominant eigenvalue, since the eigenvalue with largest absolute value occurs more than once (*i.e.*, it is a repeated root of the polynomial).

†This method works if the dominant eigenvalue is positive. A modification of the method will work if it isn't.

‡Warning: You can choose almost any unit vector as x_0. However the method may not work if x_0 happens to be orthogonal to the dominant eigenvectors.