Similar Matrices

This learning module covers the following ideas. When you make your lesson plan, it should explain and contain examples of the following:

1. Show that the null space (kernel), column space (image), and eigenspaces of a matrix (linear transformation) are vector subspaces. Be able to extend a basis for these spaces to a basis for the domain or codomain.

2. Explain how to describe vectors using different coordinate systems. Be able to change between different coordinate systems, and use these ideas to construct matrix representations of linear transformations.

3. For square matrices, explain why similar matrices $B = P^{-1}AP$ represent the same linear transformation (just using a different basis). Explain why the determinant, eigenvalues, rank, and nullity of a linear transformation are not dependent upon the basis chosen.

4. Explain how to diagonalize a matrix, as well as explain when it is or is not possible.

Following you will find preparation assignments and suggested homework, and then my best attempt at condensing the information we are learning into a concise set of notes.

1   Preparation and Suggested Homework

Preparation Problems	
Day 1	6.3, 6.6, 6.9, 6.12, 6.16
Day 2	9.18, 9.1, 9.6, 9.20, 11.13
Day 3	Handout TBA, 11.17, 11.18, 11.19

Homework list:

Concept	Sec.	Suggestions	Relevant Problems
Coordinates	6	1,3,4,6,9(use columns)	1-10,22-28
Change of Basis	6	12,15,16,31	12-16,30-33
Matrix Representations (different bases)	9	16,18,19,44	16-20,42-45
Matrix Representations (same basis)	9	1,2,5,6,9,11	1-12,27-34,36,39
Maple illustrations (on handout)	Handout	TBA	TBA
Diagonalization	11	9,11,13,14,17,18	9,11-22,57,60-63

Notice that the book does not use inverse matrices where it can. Instead, it does everything without ever referencing inverse matrices. So please take advantage of the fact that you can use inverse matrices to change coordinates. Also, remember that once you leave 2 by 2 matrices, please use your computer to find eigenvalues and inverse matrices. You should practice finding the eigenvectors of some of the 3 by 3 matrices by hand.

2   The Null Space, Column Space, and Eigen Spaces

We have already spent time studying two of the three most important subspaces related to a linear transformation. These three spaces are

Definition 1. Consider a linear transformation $T : V \to U$.

1. The kernel (null space) is all vectors in the domain V whose image is zero, $\{\vec{x}|T(\vec{x}) = \vec{0}\}$.

2. The image (column space) is all vectors in the range whose preimage is nonzero, $\{\vec{y}|T(\vec{x}) = \vec{y}$ for some $\vec{x} \in V\}$.

3. For each eigenvalue λ, the eigenspace corresponding to λ is the set of eigenvectors corresponding to λ together with the zero vector (remember that zero is not an eigenvector).

(Draft March 4, 2009)

Before we proceed, let's verify that each of these spaces is indeed a vector subspace. In what follows, assume that T is a linear transformation whose standard matrix representation is A, i.e. $T(\vec{x}) = A\vec{x}$.

1. We'll prove that the kernel is a vector subspace of the domain. We show it satisfies the 3 requirements that prove something is a subspace.

    (a) Every linear transformation satisfies $T(\vec{0}) = \vec{0}$, so the zero vector is in the kernel.

    (b) If $T(\vec{x}_1) = \vec{0}$ and $T(\vec{x}_2) = \vec{0}$, then $T(\vec{x}_1 + \vec{x}_2) = T(\vec{x}_1) + T(\vec{x}_2) = \vec{0} + \vec{0} = \vec{0}$, so $\vec{x}_1 + \vec{x}_2$ is in the kernel.

    (c) If $T(\vec{x}) = \vec{0}$, then $T(c\vec{x}) = cT(\vec{x}) = c\vec{0} = \vec{0}$, so $c\vec{x}$ is in the kernel.

2. We'll prove that the image is a vector subspace of the range. This proof is easiest if we utilize the fact that the span of any set of vectors is always a subspace. Since the image of T is the same as the column space of A, then immediately the image of T is the span of the columns of A, hence a vector subspace.

3. Now we'll show that for each eigenvalue λ of T, the set of eigenvectors together with zero is a vector subspace of the domain (or range, since they are the same if A is a square matrix). We'll do this by showing it satisfies the 3 requirements that prove something is a subspace.

    (a) We added zero to the space, so it contains zero.

    (b) If $\vec{x}_1$ and $\vec{x}_2$ are both eigenvectors (or zero), then $A\vec{x}_1 = \lambda\vec{x}_1$ and $A\vec{x}_2 = \lambda\vec{x}_2$. Then $A(\vec{x}_1 + \vec{x}_2) = A(\vec{x}_1) + A(\vec{x}_2) = \lambda\vec{x}_1 + \lambda\vec{x}_2 = \lambda(\vec{x}_1 + \vec{x}_2)$ so $\vec{x}_1 + \vec{x}_2$ is also an eigenvector (or zero).

    (c) If $\vec{x}$ is an eigenvector (or zero), then $A\vec{x} = \lambda\vec{x}$. We then compute $A(c\vec{x}) = cA(\vec{x}) = c\lambda\vec{x} = \lambda(c\vec{x})$ so $c\vec{x}$ is also an eigenvector (or zero) for any scalar c.

We have just shown that the kernel (null space), image (column space), and eigenspaces are vector spaces. You should take the time to repeat the proofs above (as I will ask you to repeat at least one of them in the future). The fact that these spaces are vector spaces means we can find a basis and use linear combinations of elements as often as we want.

2.1   Extending a basis for a subspace to the whole space

These three vector spaces play an important role in trying to understand linear transformations. In order to use these vector spaces in the work below, we need to learn how to extend a basis of a subspace to a basis of the whole space.

The kernel of $T : \mathbb{R}^n \to \mathbb{R}^m$ is a vector subspace of the domain $\mathbb{R}^n$, so we can select a basis for the kernel, namely $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k\}$. How do we pick $n-k$ other vectors $\{\vec{u}_{k+1}, \ldots, \vec{u}_n\}$ so that $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k, \vec{u}_{k+1}, \ldots, \vec{u}_n\}$ is a basis for the domain $\mathbb{R}^n$? There are many ways to do this, I'll just illustrate the simplest. Take your k basis vectors and augment by the n by n identity matrix. Then reduce the matrix, and select as your basis the n pivot columns of the resulting matrix. This will always include the first k vectors, and then you will just add in $n - k$ standard basis vectors.

Example 1. The reduced row echelon form of $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 1 \end{bmatrix}$ is $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. A basis for the null space is $\{(-2, 1, 0)\}$. Putting this vector in the first column of a matrix with 3 rows and augmenting by the identity gives us $\begin{bmatrix} -2 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{rref} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$. This means that a basis for the domain is the first, second and 4th vectors, namely $\{(-2, 1, 0), (1, 0, 0), (0, 0, 1)\}$.

The reduced row echelon form of $A = \begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix}$ is $\begin{bmatrix} 1 & \frac{4}{3} \\ 0 & 0 \end{bmatrix}$. A basis for the null space is $\{(-4/3, 1)\}$. Putting this vector in the first column of a matrix with 2 rows and augmenting by the identity gives us $\begin{bmatrix} -\frac{4}{3} & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \xrightarrow{rref} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & \frac{4}{3} \end{bmatrix}$. This means that a basis for the domain is the first, second columns, namely $\{(-4/3, 1), (1, 0)\}$. Are there other correct answers? Of course there are, as there are infinitely many ways to choose a basis.

The image of $T : \mathbb{R}^n \to \mathbb{R}^m$ is a vector subspace of the range $\mathbb{R}^m$, so we can select a basis for the image, namely $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_r\}$. How do we pick $m-r$ other vectors $\{\vec{u}_{k+1}, \ldots, \vec{u}_n\}$ so that $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_r, \vec{v}_{r+1}, \ldots, \vec{v}_n\}$ is a basis for the range $\mathbb{R}^m$? Again you can just take your r basis vectors and augment by the m by m identity matrix. Then reduce the matrix, and select as your basis the m pivot columns of the resulting matrix. This will always include the first r vectors, and then you will just add in standard basis vectors.

Example 2. The reduced row echelon form of $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 1 \end{bmatrix}$ is $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. A basis for the column space is $\{(1,2),(3,1))\}$. Since it already has 2 dimensions, there is no need to extend it to a basis for the range, we already have a basis.

The reduced row echelon form of $A = \begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix}$ is $\begin{bmatrix} 1 & \frac{4}{3} \\ 0 & 0 \end{bmatrix}$. A basis for the column space is $\{(3,6)\}$. Putting this vector in the first column of a matrix with 2 rows and augmenting by the identity gives us $\begin{bmatrix} 3 & 1 & 0 \\ 6 & 0 & 1 \end{bmatrix} \xrightarrow{rref} \begin{bmatrix} 1 & 0 & \frac{1}{6} \\ 0 & 1 & -\frac{1}{2} \end{bmatrix}$. This means that a basis for the range is the first and second columns, namely $\{(3,6),(1,0)\}$.

The reduced row echelon form of $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \\ -2 & -4 \end{bmatrix}$ is $\begin{bmatrix} 1 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$. A basis for the column space is $\{(1,3,-2)\}$. Putting this vector in the first column of a matrix with 3 rows and augmenting by the identity gives us $\begin{bmatrix} 1 & 1 & 0 & 0 \\ 3 & 0 & 1 & 0 \\ -2 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{rref} \begin{bmatrix} 1 & 0 & 0 & -\frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{3}{2} \end{bmatrix}$. This means that a basis for the range is the first and second columns, namely $\{(1,3,-2),(1,0,0),(0,1,0)\}$.

The last fact we will develop in this section relates to choosing a basis for the domain (and range) using eigenvectors of our matrix. In order to do this, we first need to know that eigenvectors corresponding to different eigenvectors are always linearly independent.

Theorem 1. *If $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n$ are n eigenvectors corresponding to n distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, then the vectors $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n$ are linearly independent.*

For a proof, see problem 11.44. The key reason we want to use this theorem is that if the domain and range V of a linear transformation $T : V \to V$ has dimension n, and there are n distinct eigenvalues, then we can form a basis for V by selecting an eigenvector corresponding to each eigenvalue. This creates a basis for V by using eigenvectors. A basis chosen in this fashion will always contain a basis for the null space (as vectors in the null space are eigenvectors corresponding to the eigenvalue 0). Choosing your basis vectors to be eigenvectors can greatly simplify how we view a linear transformation.

Example 3. The matrix $A = \begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix}$ has eigenvalues $5, -2$. Corresponding eigenvectors are $(1,1)^T$ and $(-4,3)^T$ (verify this by hand or by computer). A basis for the domain and range can be chosen to be $\{(1,1),(-4,3)\}$.

The matrix $A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ has eigenvalues $3, 3, 1$. Notice that 3 appears with multiplicity 2. To find the eigenvectors, we subtract 3 from the diagonal and reduce $\begin{bmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ to obtain $\begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. Since there are two free variables, we can find 2 linearly independent eigenvectors corresponding to 3, namely $(1,1,0)$ and $(0,0,1)$. An eigenvector corresponding to 1 is $(-1,1,0)$. These three vectors are linearly independent, and so a basis for the domain and range is $\{(1,1,0),(0,0,1),(-1,1,0)\}$.

Similar computations can be done for the matrix $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ which has eigenvalues $3, 3, 1$. To

find the eigenvectors corresponding to 3, we subtract 3 from the diagonal and reduce $\begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ to

obtain $\begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$. Now there is only one free variables, so we can find only 1 linearly independent

eigenvector corresponding to 3, namely $(1, 1, 0)$. There is also only one linearly independent eigenvector corresponding to 1, which is still $(-1, 1, 0)$. If you use software to find the eigenvectors, it will give you $(0, 0, 0)$ as a third (but remember that an eigenvector is never zero). In this case we can only obtain 2 linearly independent eigenvectors, and so a basis for the domain and range cannot be obtained just from eigenvectors alone. We could augment these two vectors by the identity (as done in the previous 2 examples) to obtain

$\begin{bmatrix} 1 & -1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{rref} \begin{bmatrix} 1 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$, which means a basis for the domain and range which contain

eigenvectors (and one other) is the first, second, and fifth columns, namely $\{(1, 1, 0), (-1, 1, 0), (0, 0, 1)\}$.

3 Changing between different coordinate systems

3.1 Coordinates relative to a basis

When we write the vector $\vec{u} = (2, 4, 7)$, we are giving the coordinates of the vector relative to the standard basis $E = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\} = \{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$. We could write $(2, 4, 7)$ as a linear combination of these basis vectors by writing $2\vec{e}_1 + 4\vec{e}_2 + 7\vec{e}_3$. The numbers $2, 4, 7$ are called the coordinates of $\vec{u}$ relative to the standard basis vectors, and they form a vector $[2, 4, 7]_E$ called the coordinate vector of $\vec{u}$ relative to E.

If instead we use the basis vectors $S = \{(1, 1, 1), (0, 1, 1), (0, 0, 1)\}$, then since $(2, 4, 7) = 2(1, 1, 1) + 2(0, 1, 1) + 3(0, 0, 1)$, we write $[\vec{u}]_S = [2, 2, 3]_S$ as the coordinate vector of $\vec{u}$ relative to S. We use brackets instead of parentheses when writing coordinate vectors, just to remind us that we are working with the coordinates of a vector instead of the vector. The subscript S at the end reminds us which basis vectors we are using.

Recall that if you want to find the coordinates of a vector relative to a basis, then enter your basis vectors in the columns of a matrix and then augment that matrix by any vectors whose coordinates you wish to find. Find the reduced row echelon form of this augmented matrix, and the coordinates of your vector will appear in the respective column. In this section we are basically giving a name (coordinates of $\vec{u}$ relative to S) to the idea of writing a column vector as a linear combination of the column vectors which precede it, something we have been doing since week 1.

Example 4. Given the basis $S = \{(1, 2), (3, 5)\}$, find the components of $(1, 0), (0, 1), (3, 6), (10, -7)$, and (a, b) relative to this basis. We just insert the basis vectors as the first two columns of our matrix, augment by the rest of the vectors, and then reduce, giving

$$\begin{bmatrix} 1 & 3 & 1 & 0 & 3 & 10 & a \\ 2 & 5 & 0 & 1 & 6 & -7 & b \end{bmatrix} \xrightarrow{rref} \begin{bmatrix} 1 & 0 & -5 & 3 & 3 & -71 & 3b - 5a \\ 0 & 1 & 2 & -1 & 0 & 27 & 2a - b \end{bmatrix}.$$

This means we have the coordinates of each vector relative to S as $(1, 0) = [-5, 2]_S, (0, 1) = [3, -1]_S, (3, 6) = [3, 0]_S, (10, -7) = [-71, 27]_S$, and $(a, b) = [3b - 5a, 2a - b]_S$. The reduced row echelon form gives us precisely the coordinates of our vectors relative to our basis.

Using the same basis as above, the coordinates $[3, -2]_S$ represent the vector $3(1, 2) - 2(3, 5) = (-3, -4)$.

The method above will always give you the coordinate of $\vec{u}$ relative to the basis $S = \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$ (where S is a basis for $\mathbb{R}^n$). However, there is another method of viewing the above row reduction in terms of inverse matrices, which we need to master in order to proceed. Much of the work in this unit will involve using inverse matrices to discover new ideas. If we want to write the vector $\vec{v}$ as a linear combination of the vectors in S, then we need to write $c_1\vec{u}_1 + c_2\vec{u}_2 + \cdots + c_n\vec{u}_n = \vec{v}$, which can be written in matrix form as $\begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_n \end{bmatrix} \begin{bmatrix} c_1 & c_2 & \cdots & c_n \end{bmatrix}^T = \vec{v}$ or $P[\vec{v}]_S = \vec{v}$, where P is the matrix whose columns are the basis vectors in S. Multiplying both sides on the left by P^{-1} gives us the formula $[\vec{v}]_S = P^{-1}\vec{v}$. So one way to find the coordinates of $\vec{v}$ relative to S is to multiply on the left of $\vec{v}$ by the inverse of P.

(Draft March 4, 2009)

Example 5. Given the basis $S = \{(1,2),(3,4)\}$, find the components of $(3,6), (10,-7)$, and (a,b) relative to this basis by using an inverse matrix. First, the inverse of $A = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$ is $A^{-1} = \frac{1}{-2}\begin{bmatrix} 4 & -3 \\ -2 & 1 \end{bmatrix}$. So the co-ordinates of $(3,6)$ relative to this basis are $A^{-1}[3,6]^T = [3,0]_S$. We also have $A^{-1}[10,-7]^T = [-61/2, 27/2]_S$ and $A^{-1}[a,b]^T = [\frac{1}{2}(3b-4a), a - \frac{b}{2}]_S$.

Finding inverse matrices can be time consuming by hand, so whenever you encounter a 3 by 3 matrix or larger, please use software to find it's inverse. For a 2 by 2 matrix, Cramer's rule provides us with a simple formula for finding the inverse (the proof is in the Applications unit handout): Just interchange the diagonal elements, change the sign on the other elements, and then divide by the determinant of the matrix as shown below:

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

Example 6. The inverse of $A = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$ is $A^{-1} = \frac{1}{-2}\begin{bmatrix} 4 & -3 \\ -2 & 1 \end{bmatrix}$. If $A = \begin{bmatrix} 13 & 3 \\ -7 & 2 \end{bmatrix}$ then $A^{-1} = \frac{1}{47}\begin{bmatrix} 2 & -3 \\ 7 & 11 \end{bmatrix}$.

3.2   Changing coordinates from one basis to another

If I am given the coordinates of $\vec{u}$ relative to S (called $[u]_S$), how do I find the coordinates of $\vec{u}$ relative to S' (called $[\vec{u}_{S'}]$)? Let A be the matrix whose columns are the basis vectors in S (relative to the standard basis E) and B be the matrix whose columns are the basis vectors of S' (relative to the standard basis E). Then we can write using the definition of coordinates the two equations $A[\vec{u}]_S = [\vec{u}]_E$ and $B[\vec{u}]_{S'} = [\vec{u}]_E$ which means we can equate the left hand sides of each equation to obtain the matrix equation $A[\vec{u}]_S = B[\vec{u}]_{S'}$. Multiplication on the left by B^{-1} gives $B^{-1}A[\vec{u}]_S = [\vec{u}]_{S'}$. The matrix $P = B^{-1}A$ is often called the change-of-basis matrix from S to S' because if you put in the coordinates $[\vec{u}]_S$ relative to S you get out the coordinates $[\vec{u}]_{S'}$ relative to S'. The change-of-coordinate matrixm Q from S' to S should invert this process, hence it is precisely $Q = P^{-1}$.

Example 7. Consider the two bases $S = \{(2,1),(3,-2)\}$ and $S' = \{(1,0),(2,-1)\}$ of the plane. (1) Find the coordinates of $[-1,3]_S$ relative to the basis S' by finding the change-of-coordinates matrix P from S to S', (2) find the coordinates of $[-1,3]_{S'}$ relative to the basis S by finding the change of coordinate matrix Q from S' to S, and (3) verify that $Q = P^{-1}$.

1. Let $A = \begin{bmatrix} 2 & 3 \\ 1 & -2 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$ (the matrices whose columns are the basis vectors). We know that $A[-1,3]^T = (7,-7) = \vec{u}$ is the actual vector we are trying to represent in terms of the basis S'. To find the coordinates relative to S', we write $A[-1,3]_S = B[\vec{u}]_{S'}$, or $B^{-1}A[-1,3]_S = [\vec{u}]_{S'}$, where $B^{-1} = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$. The change-of-coordinate matrix P from S to S' is $B^{-1}A = \begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix}$ which means that $[\vec{u}]_{S'} = P[-1,3]^T = [-7,7]_{S'}$ (notice that $\vec{u} = (7,-7)$ in the standard basis).

2. We know that $[1,3]_{S'} = B[-1,3]^T = (5,-3) = \vec{u}$ is the actual vector we are trying to represent in terms of the basis S. To find the coordinates relative to S', we write $B[-1,3]_{S'} = A[\vec{u}]_S$, or $A^{-1}B[-1,3]_{S'} = [\vec{u}]_S$, where $A^{-1} = \begin{bmatrix} \frac{2}{7} & \frac{3}{7} \\ \frac{1}{7} & -\frac{2}{7} \end{bmatrix}$. The change-of-coordinate matrix Q from S' to S is $A^{-1}B = \begin{bmatrix} \frac{2}{7} & \frac{1}{7} \\ \frac{1}{7} & \frac{4}{7} \end{bmatrix}$ which means that $[\vec{u}]_{S'} = Q[-1,3]^T = [\frac{1}{7}, \frac{11}{7}]_S$ (notice that $\vec{u} = (7,-7)$ in the standard basis).

3. Because $P = B^{-1}A$ and $Q = A^{-1}B$, we know that $P^{-1} = (B^{-1}A)^{-1} = A^{-1}(B^{-1})^{-1} = A^{-1}B = Q$ (remember that inverting matrices requires that we invert each matrix and reverse the order).

4   Matrix representations relative to bases

In the previous unit we always used the standard basis vectors to construct matrix representations of our linear transformations. For example, the linear transformation $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by $T(x,y) = (2x +$

$y, 3x+4y)$ we represented in matrix form by first computing $T(1,0) = (2,3), T(0,1) = (1,4)$ and then putting these vectors in the columns of our matrix $A = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}$. This matrix allows us to input coordinates relate to the standard basis, and get out coordinates relative to the standard basis.

With an arbitrary linear transformation $T : V \to U$ (where both V and U are finite dimensional) where the basis S and S' are chosen for V and U respectively, how do we construct the matrix representation of our linear transformation so that we can input coordinates relate to S and get out coordinates relative to S'. Notationally this matrix is written $[T]_{S,S'}$, and we seek a matrix so that $[T]_{S,S'}[\vec{v}]_S = [T(\vec{v})]_{S'}$. To answer this, first let B be the matrix whose columns are the basis vectors of S, and let C be the matrix whose columns are the basis vectors of S'. We then have immediately that $B[\vec{v}]_S = [\vec{v}]_E = \vec{v}$ and $C[T(\vec{v})]_{S'} = [T(\vec{v})]_E = T\vec{(v)}$ by the definition of coordinates. If we let A represent the standard matrix representation of the transformation, then we have $A\vec{v} = T(\vec{v})$. On the left we will replace $\vec{v}$ with $B[\vec{v}]_S$ and on the right we will replace $T(\vec{v})$ with $C[T(\vec{v})]_{S'}$ to obtain the equation

$$AB[\vec{v}]_S = C[T(\vec{v})]_{S'}.$$

Multiplying both sides on the left by C^{-1} gives

$$C^{-1}AB[\vec{v}]_S = [T(\vec{v})]_{S'}.$$

This last equation requires that we input the coordinates of a vector relative to the basis S, and then we get out the coordinates of the transformed vector relative to the basis S'. This is precisely what we were looking for, which means that the matrix representation relative to S and S' is $[T]_{S,S'} = C^{-1}AB$, where B's columns are the basis vectors of S and C's columns are the basis vectors of S'.

Example 8. Consider the linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^3$ defined by $T(x,y) = (x+y, x-y, 2x+3y)$. Find a matrix representation using the bases $S = \{(2,3), (-1,1)\}$ and $S' = \{(1,0,0), (1,1,0), (3,2,1)\}$. To start with, the standard matrix representation is $A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 2 & 3 \end{bmatrix}$. Let $B = \begin{bmatrix} 2 & -1 \\ 3 & 1 \end{bmatrix}$ and $C = \begin{bmatrix} 1 & 1 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$. Any vector $\vec{u}$ in the domain can be written as of coordinates relative to S using the product $B[\vec{u}]_S = \vec{u}$. Any vector in the image can be written as coordinates relative to S' using the product $C[T(\vec{u})]_{S'} = T(\vec{u})$. Since $A\vec{u} = T(\vec{u})$, we replace $\vec{u}$ with $B[\vec{u}]_S$ and $T(\vec{u})$ with $C[T(\vec{u})]_{S'}$ to obtain $AB[\vec{u}]_S = C[T(\vec{u})]_{S'}$ or $C^{-1}AB[\vec{u}]_S = [T(\vec{u})]_{S'}$. The inverse of C is (using software) $C^{-1} = \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$. So our matrix representation is

$$C^{-1}AB = C^{-1} = \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ -1 & -2 \\ 13 & 1 \end{bmatrix} = [T]_{S,S'}.$$

Graphically, the previous computations just give us different ways of viewing the domain and range. In Maple, we can draw grids to represent the basis vectors we have chosen for the domain and range. This just gives us ways of drawing grids on the domain and range which we can then use to count out our coordinates relative to the bases chosen. Check the Maple introduction for more information here.

4.1   Who cares - Why pick a different basis?

From the previous section, hopefully you discovered that there are many different ways of viewing the same transformation. Which view is the best? There is not really a correct answer here, but there are some nice bases you can choose to simplify computations. For example, if you start with the null space and extend a basis for the null space to a basis for the domain, then the matrix representation relative to this basis will contain columns of zeros (one for every basis vector from the null space). In other words, you can arrange your basis vectors to purposefully squash out of the matrix everything with gets mapped to zero. This can greatly simplify computations.

(Draft March 4, 2009)

Example 9. Consider the linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^2$ defined by $T(x, y, z) = (x+y-z, 2x+3y+4z)$.
The standard matrix representation is $A = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 3 & 4 \end{bmatrix}$ whose rref is $\begin{bmatrix} 1 & 0 & -7 \\ 0 & 1 & 6 \end{bmatrix}$. A basis for the null space
is $(7, -6, 1)$. We can extend this to a basis for the domain by augmenting by the identity and reducing,
$\begin{bmatrix} 7 & 1 & 0 & 0 \\ -6 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{rref} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -7 \\ 0 & 0 & 1 & 6 \end{bmatrix}$, which means that $S = \{(7, -6, 1), (1, 0, 0), (0, 1, 0)\}$ is a basis for
the domain. Using $E = \{(1, 0), (0, 1)\}$ as our basis for the range, we can compute the matrix representation
of T relative to the bases S and E by noting

$$ C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = C^{-1}, A = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 3 & 4 \end{bmatrix}, B = \begin{bmatrix} 7 & 1 & 0 \\ -6 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, C^{-1}AB = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 2 & 3 \end{bmatrix} = [T]_{S,E}. $$

Notice how this made an entire column of the matrix zero. If we rearrange the basis vectors of S so that
the null space is at the end, using $S = \{(1, 0, 0), (0, 1, 0), (7, -6, 1)\}$ instead, then our matrix is simply
$[T]_{S,E} = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 0 \end{bmatrix}$ (which just moves the columns of zero to the end of the matrix instead of the beginning.
This is done in practice just to make it easy to remember that we can ignore the last column of the matrix
(instead of the first).

After picking a basis for the domain which contains the kernel, let's carefully construct a basis for the
range to make our matrix representation extremely nice. If k is the dimension of the kernel and $S = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k, \vec{v}_{k+1}, \ldots, \vec{v}_n\}$ is a basis for the domain (whose first k vectors are in the kernel), then consider
the $n - k$ vectors $T(\vec{v}_{k+1}), \ldots, T(\vec{v}_n)$ which are nonzero vectors in the image. Since the image has to have
dimension $n - k$, these vectors are linearly independent and we can extend them to a basis S' for the range.
The matrix representation of T relative to the bases S and S' will always be a matrix which consists of all
zeros and $n - k$ 1's (what can be simpler than zeros and 1's). This representation of T illustrates how the
kernel and image can be use to obtain an extremely nice form for our matrix representation.

Example 10. Consider the linear transformation from the previous example $T : \mathbb{R}^3 \to \mathbb{R}^2$ defined by
$T(x, y, z) = (x + y - z, 2x + 3y + 4z)$. Using the bases $S = \{(1, 0, 0), (0, 1, 0), (7, -6, 1)\}$ and E (the standard
basis for the plane), the matrix representation of T relative to these bases is $[T]_{S,E} = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 0 \end{bmatrix}$. Choose
for your basis for the image the vectors $S' = \{(1, 2), (1, 3)\}$ (the nonzero columns of the previous matrix).
The matrix representation of T relative to the bases S and S' is then found using

$$ C = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}, C^{-1} = \frac{1}{1}\begin{bmatrix} 3 & -1 \\ -2 & 1 \end{bmatrix}, A = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 3 & 4 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 & 7 \\ 0 & 1 & -6 \\ 0 & 0 & 1 \end{bmatrix}, C^{-1}AB = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = [T]_{S,S'}. $$

Not only have we made the last column zero, but we have also placed the identity matrix in the upper left
corner of the matrix. This will happen in general, even with bigger matrices.

As a bigger example (do the computations on the computer to verify the work below), consider the
transformation $T(x, y, z, w) = (2x - y, y + z, 2x + z, x + w)$. The standard matrix representation is $A = \begin{bmatrix} 2 & -1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 2 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$ whose rref is $\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. A basis for the null space is then $(-1, -2, 2, 1)$. We can
extend this to a basis for the domain by augmenting by the identity and reducing,

$$ \begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ -2 & 0 & 1 & 0 & 0 \\ 2 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{rref} \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & -2 \end{bmatrix}, $$

which means that $S = \{(1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 0, 1), (-1, -2, 2, 1)\}$ is a basis for the domain (moving
the null space vector to the end). Using the standard basis for the range, we can compute the matrix

(Draft March 4, 2009)

representation of T relative to the bases S and E by noting

$$C = I = C^{-1}, A = \begin{bmatrix} 2 & -1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 2 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}, C^{-1}AB = \begin{bmatrix} 2 & -1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 2 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} = [T]_{S,E}.$$

This made the last column of the matrix zero. Now choose for your basis for the image the vectors $\{(2,0,2,1), (-1,1,0,0), (0,1,1,0)\}$ (the nonzero columns of the previous matrix). Extend this set to a basis for the image by augmenting by the identity and reducing

$$\begin{bmatrix} 2 & -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{rref} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & -1 & 2 \\ 0 & 0 & 1 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 & 1 & -1 & 0 \end{bmatrix},$$

which means that $S' = \{(2,0,2,1), (-1,1,0,0), (0,1,1,0), (1,0,0,0)\}$ is a basis for the range. The matrix representation of T relative to the bases S and S' is then found using

$$C = \begin{bmatrix} 2 & -1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 2 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, C^{-1} = \frac{1}{1}\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 1 & -2 \\ 1 & 1 & -1 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}, C^{-1}AB = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = [T]_{S,S'}.$$

Not only have we made the last column zero, but we have also placed the identity matrix in the upper left corner of the matrix. This is perhaps the simplest form you can obtain, but it requires using a different basis for the domain and range.

4.2   Similar matrices - When the domain and range are the same

If a linear transformation $T : V \to V$ has the same domain as range, then often we pick the same basis S for both the domain and range. In this case we write $P[\vec{u}]_S = [\vec{u}]_E$, where P is the matrix whose columns are the basis vectors in S. If A is the standard matrix representation of the linear transformation, we can write $A[\vec{u}]_E = [T(\vec{u})]_E$ which after replace the standard coordinates with S coordinates becomes $AP[\vec{u}]_S = P[T(\vec{u})]_S$ or $P^{-1}AP[\vec{u}]_S = [T(\vec{u})]_S$, hence the matrix representation of T relative to the basis S is $[T]_S = P^{-1}AP$. This new matrix is found by multiplying on the right by a matrix, and on the left by its inverse.

In general, we say A and B are similar matrices if there exists some invertible matrix P such that $B = P^{-1}AP$. Some important facts about similar matrices follow (together with brief justifications):

1. Every matrix A is similar to itself, because $A = I^{-1}AI$.

2. If A is similar to B, then B is similar to A. This is because if $B = P^{-1}AP$, then multiplying both sides on the right by P^{-1} and on the left by P gives $PBP^{-1} = A$, or $A = (P^{-1})^{-1}BP^{-1}$ (as $(P^{-1})^{-1} = P$).

3. If A is similar to B and B is similar to C, then A is similar to C. This is because if $B = P^{-1}AP$ and $C = Q^{-1}BQ$, then $C = Q^{-1}BQ = Q^{-1}P^{-1}APQ = R^{-1}AR$ where $R = PQ$ and recall that $R^{-1} = Q^{-1}P^{-1}$ (inverting a product requires inverting each factor and reversing the order of multiplication).

4. If A and B are similar, then they have the same determinant. This is because $|B| = |P^{-1}AP| = |P^{-1}||A||P| = |P|^{-1}|A||P| = |A|$.

5. If A and B are similar, then they have the same eigenvalues (not the same eigenvectors). This is because $|B - \lambda I| = |P^{-1}AP - \lambda I| = |P^{-1}AP - \lambda P^{-1}IP| = |P^{-1}(A - \lambda I)P| = |P^{-1}||(A - \lambda I)||P|$, which means $|B - \lambda I| = 0$ if and only if $|P^{-1}||(A - \lambda I)||P| = 0$, the latter equals zero if and only if $|(A - \lambda I)| = 0$.

6. If A and B are similar, then they have the same rank and nullity. The column space and null space of A and B may differ, but their dimensions will not change. The proof requires more work than we will have time to develop, but the key idea is that if a matrix squashing 2 dimensions (has nullity 2), then a similar matrix will squash 2 dimensions as well.

Similar matrices are just different ways of viewing the same linear transformation in terms of a different basis. The facts above show that the determinant, eigenvalues, rank, and nullity of a linear transformation do not depend on the basis chosen. They are important numbers which belong to any linear transformation (regardless of the basis) and can be discussed without referencing the basis vectors (or obtaining a matrix).

Example 11.   • Consider the matrix $A = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$. Let $P = \begin{bmatrix} 2 & 9 \\ 1 & 4 \end{bmatrix}$ which means $P^{-1} = \begin{bmatrix} -4 & 9 \\ 1 & -2 \end{bmatrix}$. The matrix representation of T relative to S (the column vectors of P) is the similar matrix

$$P^{-1}AP = \begin{bmatrix} -4 & 9 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 2 & 9 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 39 & 170 \\ -8 & -35 \end{bmatrix}.$$

• Consider the same matrix $A = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$. Pick a different matrix $P = \begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix}$ which means $P^{-1} = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ -\frac{1}{3} & \frac{1}{3} \end{bmatrix}$. The matrix representation of T relative to S (the column vectors of P) is the similar matrix

$$P^{-1}AP = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ -\frac{1}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 0 & -1 \end{bmatrix}.$$

Notice how this choice of basis resulted in a diagonal matrix. The next section shows how you can find the right matrix P so that A is similar to a diagonal matrix.

4.3   Picking the "Best Basis" - Diagonalizing a Matrix

In the previous section we discussed similar matrices. We said that A is similar to B if $B = P^{-1}AP$ for some invertible matrix A. This is equivalent to writing $AP = PB$, for some invertible P. In this section, we will discuss a key method of choosing P so that the matrix A is similar to a diagonal matrix D, or $AP = PD$ for a diagonal matrix D. In order to do this, we use eigenvectors.

Let A be an n by n matrix. Suppose A has n linearly independent eigenvectors. Let P be the matrix whose columns are the eigenvectors of A. Then $AP = PD$ where D is the diagonal matrix of eigenvalues (written in the same order as the order in which the eigenvectors are written in P). This means $P^{-1}AP = D$, or that A is similar to D. This process is called diagonalizing a matrix. Notice that if A is similar to D, then $AP = PD$ must hold, which means that P must be a matrix whose columns are linearly independent eigenvectors (so that P has an inverse). In other words, a matrix is diagonalizable if and only if there are n linearly independent eigenvectors.

Example 12.   • Consider the linear transformation $T(x,y) = (x + 4y, 2x + 3y)$ whose standard matrix representation is $A = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$. The eigenvalues are 5 and -1 (verify this by hand on 2 by 2 matrices, but use the computer to do anything bigger). Corresponding eigenvectors are $(1,1)$ and $(-2,1)$, so let $S = \{(1,1),(-2,1)\}$ be a basis for the plane. Form the matrix $P = \begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix}$ whose columns are the eigenvectors. The inverse of P is $P^{-1} = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ -\frac{1}{3} & \frac{1}{3} \end{bmatrix}$. The matrix representation of T relative to S is

$$P^{-1}AP = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ -\frac{1}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 0 & -1 \end{bmatrix},$$

a diagonal matrix. We have diagonalized A (or found a diagonal representation of the linear transformation T). Type this example into the Maple introduction to visual what we have just done.

- Let's now repeat with a 3 by 3 matrix. Consider the linear transformation defined by the matrix $A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$. The eigenvalues are 3,3,1 (use the computer). Corresponding eigenvectors are $(0,0,1),(1,1,0)$, and $(-1,1,0)$, so let $S = \{(0,0,1),(1,1,0),(-1,1,0)\}$ be a basis for our domain and range. Form the matrix $P = \begin{bmatrix} 0 & 1 & -1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix}$ whose columns are the eigenvectors. The inverse of P is $P^{-1} = \begin{bmatrix} 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$. The matrix representation of T relative to S is

$$P^{-1}AP = \begin{bmatrix} 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 0 & 1 & -1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

  a diagonal matrix. We have diagonalized A (or found a diagonal representation of the linear transformation T). Type this example into the Maple introduction to visual what we have just done.

- This last example shows that not every matrix can be diagonalized. Consider the linear transformation defined by the matrix $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$. The eigenvalues are 3,3,1 (use the computer). Corresponding eigenvectors are $(0,0,1)$ for 3 and $(-1,1,0)$ for 1. There are only 2 linearly independent eigenvectors (the computer will give you a $(0,0,0)$ vector, but that is not an eigenvector). In this case, there are not enough linearly independent eigenvectors to diagonalize this matrix. Jordan Form is a method of handling these kinds of matrices, but we will not endeavor to discuss Jordan form this semester.

Is every n by n matrix diagonalizable? In other words, does every n by n matrix have n linearly independent eigenvectors? The last example above showed that the answer is no. There are matrices which do not have this property. In order for a matrix to not have enough eigenvectors, an eigenvalue must appear with multiplicity greater than 1, but not have the same number of corresponding linearly independent eigenvectors.

Definition 2. The number of times an eigenvalues appears as a root of the characteristic equation is called the algebraic multiplicity of the eigenvalue. The dimension of the corresponding eigenspace is called the geometric multiplicity of the eigenvalue.

Theorem 2. *The algebraic multiplicity of each eigenvalue equals the geometric multiplicity of each eigenvalue if and only if A is diagonalizable. If all the eigenvalues are distinct, then A is diagonalizable.*

5   So Who Cares? Why is all of this useful?

When viewing linear transformations as ways of stretching the plane or space, using eigenvectors as your basis for both the domain and range provides us with an extremely nice view of the transformation. The transformation produces only a stretch in the outward direction of our basis vectors (which is why the matrix representation is diagonal). This is illustrated in Maple.

If we return to the vector field view, recall that you can see the eigenvectors in a vector field graph. This will allow you to determine the matrix P from the picture. If you measure how much the vector field stretches the eigenvectors radially, then you can physically measure the matrix D. This allows you to find the standard matrix representation of the vector field by using the product $D = P^{-1}AP$ and solving for A to obtain $A = PDP^{-1}$.

(Draft March 4, 2009)