Math 43 Review Notes

[Disclaimer: This is not a complete list of everything you need to know, just some of the topics that gave people difficulty.]

Dot Product

If $\mathbf{v} = (v_1, v_2, v_3)$ and $\mathbf{w} = (w_1, w_2, w_3)$, then the dot product of $\mathbf{v}$ and $\mathbf{w}$ is given by

$$\mathbf{v}\cdot\mathbf{w} = v_1w_1 + v_2w_2 + v_3w_3$$

For example, if $\mathbf{v} = (4, 7, 6)$ and $\mathbf{w} = (2, 3, 9)$, then $\mathbf{v}\cdot\mathbf{w} = 8 + 21 + 54 = 83$. Notice that the answer is always a *number*.

Row Operations

There are three types of row operations:

(1) Adding a multiple of one row to another.
(2) Switching two rows.
(3) Multiplying a row by a constant.

There's a modification of (1) that you can use to avoid fractions. Namely, if you were going to do the operation $row2 - (3/7)row1$, you could instead use $7row2 - 3row1$. This is just the first operation multiplied by the denominator 7.

Any of these operations can be used to solve $A\mathbf{x} = \mathbf{b}$, and in finding inverses, but to find the LU factorization *only* use operations of type 1 (and not the modification).

Echelon and Reduced Row Echelon (rref) forms

A *pivot* is the first nonzero entry in a row which has no other pivots directly above it.

You can identify the echelon form of a matrix by the following properties:

(1) Below each pivot are zeros.
(2) Each pivot is to right of the one above it.
(3) Rows of all zeros (if any at all) must come at the end.

Echelon form is like a lower triangular form for matrices which aren't necessarily square. The reduced row echelon form of a matrix is an echelon form, but now each pivot must be a 1 and there must be zeroes above the pivots as well as below them.

$$\text{Echelon Forms}\quad \begin{pmatrix} \mathbf{x} & x & x \\ 0 & \mathbf{x} & x \\ 0 & 0 & \mathbf{x} \end{pmatrix} \begin{pmatrix} \mathbf{x} & x & x & x \\ 0 & \mathbf{x} & x & x \\ 0 & 0 & 0 & \mathbf{x} \end{pmatrix} \begin{pmatrix} \mathbf{x} & x & x & x \\ 0 & \mathbf{x} & x & x \\ 0 & 0 & \mathbf{x} & x \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & \mathbf{x} & x \\ 0 & 0 & x \\ 0 & 0 & \mathbf{x} \\ 0 & 0 & 0 \end{pmatrix}$$

$$\text{RREF of above}\quad \begin{pmatrix} \mathbf{1} & 0 & 0 \\ 0 & \mathbf{1} & 0 \\ 0 & 0 & \mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{1} & 0 & x & 0 \\ 0 & \mathbf{1} & x & 0 \\ 0 & 0 & 0 & \mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{1} & 0 & 0 & x \\ 0 & \mathbf{1} & 0 & x \\ 0 & 0 & \mathbf{1} & x \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & \mathbf{1} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \mathbf{1} \\ 0 & 0 & 0 \end{pmatrix}$$

Elimination Matrices

For an operation of the form $rowi - m \cdot rowj$ you get the elimination matrix E_{ij} from the identity matrix by replacing entry (i, j) (that's row i, column j) with $-m$. If you do row operations in the order we've done in class (use the pivot in the first row to get zeroes below it, then use the pivot in the second row to get zeroes below it, etc.), then E_{ij} will have its $-m$ in the same location as the entry that you were trying to zero out. Note that if you did the operation $rowi + m \cdot rowj$, E_{ij} would have a $+m$ instead of $-m$.

The matrix E_{ij} has the effect that E_{ij} times *any* matrix A subtracts m times row j from row i of A.

LU Factorization

If you do a sequence of row operations to reduce a matrix A into lower triangular form U, it can be written as something like $E_{32}E_{31}E_{21}A = U$. Solving this for A gives $A = (E_{21}E_{31}E_{32})^{-1}U$. That big inverse is what we call L, and we find it similarly to finding the elimination matrices. To find L, start with the identity matrix. For each operation of the form $rowi - m \cdot rowj$ replace the (i,j) entry with $+m$ (note signs are the opposite of what they are for elimination matrices since L is an inverse of elimination matrices.) Again, if you do the operations in the order we've done in class, then whatever position you're trying to get a zero in, the corresponding entry in L gets replaced by m.

For example let

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 4 \\ 3 & -5 & 14 \end{pmatrix}$$

Reduce A to lower triangular form, indicating the elimination matrices and find the LU Factorization of A.

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 4 \\ 3 & -5 & -5 \end{pmatrix} \overset{row2-2row1}{\to} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 3 & -5 & -5 \end{pmatrix} \overset{row3-3row1}{\to} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 0 & -8 & -8 \end{pmatrix} \overset{row3+4row2}{\to} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & 4 \end{pmatrix}$$

$$E_{21} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad E_{31} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix} \quad E_{32} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 4 & 1 \end{pmatrix} \quad L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & -4 & 1 \end{pmatrix} \quad U = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & 4 \end{pmatrix}$$

Using LU Factorization To Solve $A\mathbf{x} = \mathbf{b}$

We can write $A\mathbf{x} = \mathbf{b}$ as $LU\mathbf{x} = \mathbf{b}$. Letting $U\mathbf{x} = \mathbf{c}$ we see that we can solve $A\mathbf{x} = \mathbf{b}$ in two steps:

(1) Solve $L\mathbf{c} = \mathbf{b}$.
(2) Solve $U\mathbf{x} = \mathbf{c}$.

Each step is easy, only requiring back substitution and no row operations. For example, use the LU Factorization in the above example to solve $A\mathbf{x} = \mathbf{b}$ with $\mathbf{b} = (2, 5, 6)$.

(1) First solve $L\mathbf{c} = \mathbf{b}$.

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & -4 & 1 \end{pmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}$$

Write this as three equations

$c_1 = 2$
$2c_1 + c_2 = 5$
$3c_1 - 4c_2 + c_3 = 6$

Solve to get $c_1 = 2$, $c_2 = 1$, $c_3 = 4$.

(2) Then solve $U\mathbf{x} = \mathbf{c}$.

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & 4 \end{pmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix}$$

Write this as three equations

$x_1 + x_2 + x_3 = 2$
$2x_2 + 3x_3 = 1$
$4x_3 = 4$

Solve to get $x_1 = 2$, $x_2 = -1$, $x_3 = 1$.

PA=LU Factorization

If you try to to the LU factorization and find that you can't do it without a row exchange, then use the $PA = LU$ factorization. You start with row operations just like in the LU factorization, but as soon as you see you have to flip two rows, flip them instead in the *original matrix* and *start over again* with row operations on the this "new" A to get L and U. Put your row flips into the matrix P. You get P by starting with the identity matrix and flipping the same rows of it that you flipped during your row operations (i.e., if you flipped rows 2 and 3 during row operations, then P is the identity matrix with rows 2 and 3 flipped).

For example let

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 4 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Find the PA=LU Factorization of A.

$$\begin{pmatrix} 1 & 2 & 0 \\ 2 & 4 & 1 \\ 1 & 1 & 1 \end{pmatrix} \overset{row2-2row1}{\rightarrow} \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

At this point, stop and notice that the (2,2) entry is 0. This should be a pivot. We need this to be nonzero in order to make the (3,2) entry zero, so we have to do a row exchange. Exchange rows 2 and 3 of A and start over.

$$\begin{pmatrix} 1 & 2 & 0 \\ 1 & 1 & 1 \\ 2 & 4 & 1 \end{pmatrix} \overset{row2-row1}{\rightarrow} \begin{pmatrix} 1 & 2 & 0 \\ 0 & -1 & 1 \\ 2 & 4 & 1 \end{pmatrix} \overset{row3-2row1}{\rightarrow} \begin{pmatrix} 1 & 2 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

The last matrix is in lower triangular form. It is our U. We have

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} \quad U = \begin{pmatrix} 1 & 2 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

Transposes and Symmetric Matrices

To get A^T, row i of A becomes column i of A^T. Remember the rules for transposes:

(1) $(A^T)^T = A$
(2) $(A \pm B)^T = A^T \pm B^T$
(3) $(AB)^T = B^T A^T$
(4) $(A^{-1})^T = (A^T)^{-1}$

A symmetric matrix is a matrix whose entries are mirror images of each other on either side of the diagonal. Mathematically, they are defined as matrices with $A = A^T$.

$$\begin{pmatrix} a & u & v & w \\ u & b & x & y \\ v & x & c & z \\ w & y & z & d \end{pmatrix}$$

Example: If A and B are symmetric, show ABA is symmetric.

Answer: $(ABA)^T = A^T B^T A^T = ABA$ (using property (3) and the fact that $A = A^T$, $B = B^T$). We have shown $(ABA)^T = ABA$, so it is symmetric.

Example: If A and B are symmetric is $AB(A+B)$ symmetric?

Answer: $[AB(A+B)]^T = (A+B)^T B^T A^T = (A^T + B^T) B^T A^T = (A+B)BA$. Matrix multiplication is not commutative so this is not necessarily the same as $AB(A+B)$. So it is not necessarily symmetric. For example, check that $AB(A+B)$ and $(A+B)BA$ are not the same with

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 2 & 3 \\ 0 & 0 \end{pmatrix}$$

Vector Spaces and Subspaces

A vector space, roughly speaking, is a set where you can define addition and multiplication by a number in such a way that basic algebraic rules hold (like $c(\mathbf{v}+\mathbf{w}) = c\mathbf{v}+c\mathbf{w}$, and a few others). Examples of vector spaces are all vectors with 2 components, or all vectors with 3 components, etc. Another example is the set of all 2x2 matrices (or 3x3 matrices, etc).

A subspace of a vector space is a set of some of the objects from the vector space which have to satisfy the following two properties:

(1) if $\mathbf{v}$ and $\mathbf{w}$ are in the subspace, then $\mathbf{v}+\mathbf{w}$ has to be in it, too.
(2) if $\mathbf{v}$ is in the subspace and c is a number, then $c\mathbf{v}$ must be in the subspace.

Usually the vectors in the set to be checked all have something in common or look a certain way. To show it really is a subspace you have to check that both properties hold, in other words, check that $\mathbf{v}+\mathbf{w}$ and $c\mathbf{v}$ have that same thing in common. To show something is a subspace you have to use generic vectors, specific examples are not enough. To show something is *not* a subspace however, specific examples are enough, just find a specific example where one of the properties fails.

Example: Show that all vectors (b_1, b_2, b_3) with $b1 + b2 + b3 = 0$ is a subspace of $\mathbb{R}^3$.

Answer: Examples of vectors in the subspace are $(-1, 0, 1)$ or $(3, -1, -2)$. These are vectors whose entries add up to 0. This is what all vectors in the subspace have in common.

First check property 1. Let $\mathbf{c} = (c_1, c_2, c_3)$ and $\mathbf{d} = (d_1, d_2, d_3)$ be any two vectors where $c_1 + c_2 + c_3 = 0$ and $d_1 + d_2 + d_3 = 0$. Then $\mathbf{c}+\mathbf{d} = (c_1 + d_1, c_2 + d_2, c_3 + d_3)$. For this to be in the subspace the entries in $\mathbf{c}+\mathbf{d}$ must add up to 0. This is true since

$$(c_1 + d_1) + (c_2 + d_2) + (c_3 + d_3) = (c_1 + c_2 + c_3) + (d_1 + d_2 + d_3) = 0 + 0 = 0.$$

Net check property 2. Let $\mathbf{c} = (c_1, c_2, c_3)$ be a vector with $c_1 + c_2 + c_3 = 0$ and let n be any number. Then $n\mathbf{c} = (nc_1, nc_2, nc_3)$ is in the subspace since

$$nc_1 + nc_2 + nc_3 = n(c_1 + c_2 + c_3) = n \cdot 0 = 0.$$

Since both properties hold, it is a subspace.

Example: Is the set of all vectors (b_1, b_2, b_3) with all of the entries whole numbers a subspace of $\mathbb{R}^3$?

Answer: No. You can check that the first property works since adding whole numbers gives whole numbers, but property 2 doesn't work. For example, pick the vector $\mathbf{b} = (1, 1, 1)$ and let $n = .5$. Then $n\mathbf{b} = (.5, .5, .5)$ is not in the subspace since its entries are not whole numbers.

Nullspace

The nullspace of a matrix A, denoted $N(A)$ consists of all vectors which are solutions to the equation $A\mathbf{x} = \mathbf{0}$.

Example: Find the nullspace of the following matrix:

$$A = \begin{pmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{pmatrix}$$

Answer: First reduce A to rref to get

$$A = \begin{pmatrix} \mathbf{1} & -2 & 0 & -1 & -3 \\ 0 & 0 & \mathbf{1} & 2 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Keep in mind that finding the nullspace is the same as solving $A\mathbf{x} = \mathbf{0}$, so rewrite the above matrix in equation form.

$x_1 - 2x_2 - x_4 - 3x_5 = 0$
$x_3 + 2x_4 - 2x_5 = 0$

Solve these for the pivot variables x_1 and x_3.

$x_1 = 2x_2 + x_4 + 3x_5$
$x_3 = -2x_4 + 2x_5$

Now write

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2x_2 + x_4 + 3x_5 \\ x_2 \\ -2x_4 + 2x_5 \\ x_4 \\ x_5 \end{bmatrix} = x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}$$

The vectors at the last step are what the book calls "special solutions". There is one for each free variable. The null space consists of all linear combinations of these vectors.

Solving $A\mathbf{x} = \mathbf{b}$

Solving $A\mathbf{x} = \mathbf{b}$ consists solving $A\mathbf{x} = \mathbf{0}$ (i.e. finding the nullspace) plus one more step. First start by row reducing the *augmented matrix*. For example let $b = (-5, -5, 12)$ and let A be the matrix in the nullspace example above. Row reducing the augmented matrix to rref, we get:

$$\left(\begin{array}{ccccc|c} \mathbf{1} & -2 & 0 & -1 & -3 & 1 \\ 0 & 0 & \mathbf{1} & 2 & -2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Notice that if the last entry in the last row were not 0, then there would be no solution, as writing this in equation form, the last equation would be $0 =$ some nonzero number, which is nonsense.

Now to find the solution: First find the nullspace (which we did above) and then find a particular solution. A particular solution is the solution you get by setting the free variables (in this case x_2, x_4, x_5) all equal to 0. What it works out to is just set the pivot variables equal to the entries in the last column in the augmented matrix, and set the free variables equal to zero. Here we get $x_1 = 1, x_3 = 2, x_2 = x_4 = x_5 = 0$. This give the particular solution $(1, 0, 2, 0, 0)$. Adding this to the solution to $A\mathbf{x} = \mathbf{0}$ (the nullspace) translates it into a solution to $A\mathbf{x} = \mathbf{b}$. Thus the complete solution is:

$$\mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}$$

Rank, etc.
A few definitions:

(1) The *rank* of a matrix is the number of pivots.
(2) A matrix has *full column rank* if every column has a pivot.
(3) A matrix has *full row rank* if every row has a pivot.
(4) A *pivot column* is a column which contains a pivot.

To determine each of the above quantities, reduce the matrix to echelon form, from there it is easy.

Column Space

The column space of a matrix A, denoted $C(A)$, consists of all vectors $\mathbf{b}$ which are solutions of $A\mathbf{x} = \mathbf{b}$. Another way to think of it is as all vectors which are linear combinations of the columns of A. If A is simple enough you can determine $C(A)$ just by looking at it, but otherwise reduce A into echelon form to determine the pivot columns. Then $C(A)$ is the set of linear combinations of the pivot columns of A (not the echelon form, but A itself). Another way to determine $C(A)$ is to reduce the augmented matrix with the last column containing entries $b_1, b_2, \ldots$ to an echelon form. Here is an example:

$$A = \begin{pmatrix} 1 & 3 & 3 & 2 \\ 1 & 3 & 4 & 6 \\ 2 & 6 & 9 & 16 \end{pmatrix}$$

An echelon form of A found by row reduction is

$$\begin{pmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

From here we see that columns 1 and 3 of A are pivot columns. Thus $C(A)$ is all linear combinations of these columns. In other words $C(A)$ consists of all vectors which can be written in the following form:

$$C(A) = a \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} + b \begin{bmatrix} 3 \\ 4 \\ 9 \end{bmatrix} \quad \text{where } a \text{ and } b \text{ are any numbers.}$$

The other way we can find $C(A)$ is as follows: First, row reduce the following augmented matrix

$$A = \left(\begin{array}{cccc|c} 1 & 3 & 3 & 2 & b_1 \\ 1 & 3 & 4 & 6 & b_2 \\ 2 & 6 & 9 & 16 & b_3 \end{array}\right) \overset{row2-row1}{\rightarrow} \left(\begin{array}{cccc|c} 1 & 3 & 3 & 2 & b_1 \\ 0 & 0 & 1 & 4 & b_2 - b_1 \\ 2 & 6 & 9 & 16 & b_3 \end{array}\right) \overset{row3-2row1}{\rightarrow} \left(\begin{array}{cccc|c} 1 & 3 & 3 & 2 & b_1 \\ 0 & 0 & 1 & 4 & b_2 - b_1 \\ 0 & 0 & 3 & 12 & b_3 - 2b_1 \end{array}\right)$$

$$\overset{row3-3row2}{\rightarrow} \left(\begin{array}{cccc|c} 1 & 3 & 3 & 2 & b_1 \\ 0 & 0 & 1 & 4 & b_2 - b_1 \\ 0 & 0 & 0 & 0 & b_3 - 2b_1 - 3(b_2 - b_1) \end{array}\right)$$

For this to have a solution, the bottom right entry in the augmented matrix must be 0. This means that $b_3 = 3b_2 - b_1$ (setting the entry equal to 0 and solving for b_3). Thus any vector in the column space is of the form

$$\begin{bmatrix} b_1 \\ b_2 \\ 3b_2 - b_1 \end{bmatrix}$$

Notice that any vector from the first way of finding $C(A)$ can be written in this form.

Full Row Rank and Full Column Rank

If a matrix has full column rank (i.e. a pivot in every column), then the following are true:
(1) There are no free variables
(2) The null space of the matrix consists only of the vector of all zeros.
(3) $A\mathbf{x} = \mathbf{b}$ has either no solution or exactly one solution.

If a matrix has full row rank (i.e. a pivot in every row), then the following are true:
(1) The number of free variables is equal to the number of columns minus the number of rows.
(2) $A\mathbf{x} = \mathbf{b}$ always has at least one solution.

More About $A\mathbf{x} = \mathbf{b}$

(1) If the rref of a matrix has zero rows, then it's possible that there is no solution. If the rref has no zero rows, then there is certainly a solution. If there are free variables, then either there's no solution or there are infinitely many solutions (and nothing in between). If there are no free variables, then there can't be infinitely many solutions. There's either no solution or one solution in this case.

(2) If A has full row rank and full column rank, then A is a square matrix (same number of rows as columns) and the rref of A is the identity matrix. Thus $A\mathbf{x} = \mathbf{b}$ has exactly one solution, namely $A^{-1}\mathbf{b}$, and you can find it by row reducing A and back substitution (note that there's no nullspace to worry about). This is the most important case.

(3) If A has full row rank and more columns than rows, then there are no zero rows in the rref, hence there is at least one solution. However, since there are more columns than rows, there are some columns without pivots, hence there are free variables. Thus there are infinitely many solutions.

(4) If A has full column rank with more rows than columns, then there are zero rows in the rref, thus it is possible there is no solution. Since every column has a pivot, there are no free variables, hence there can't be an infinite number of solutions. Thus there is either no solution or one solution.

(5) If A has neither full row rank nor full column rank, then the rref has zero rows and there are free variables. Thus there is either no solution or infinitely many solutions.

(6) Remember, there can never be exactly 2 or exactly 3 solutions. Zero, one, or infinitely many solutions are the only possibilities. (It's a good exercise to try to show why exactly 2 solutions is impossible.)

Linear Independence

A set of vectors is *linearly independent* if none of the vectors can be written as a linear combination of the others. Otherwise we say the vectors are *linearly dependent.* Mathematically this is expressed by saying the vectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ are linearly independent if $c_1\mathbf{v_1} + c_2\mathbf{v_2} + \ldots + c_n\mathbf{v_n} = \mathbf{0}$ only when $c_1 = c_2 = \ldots = c_n = 0$. This is the same as saying the matrix with these vectors as its columns has full column rank (do you know why?).

If a vector in the set can be written as a linear combination of others, for many applications it is somehow redundant. A linearly independent set has no such redundancies, which is good.

If any of the following happens, then the vectors are linearly *dependent* .

(1) One vector is a multiple of another, or there is an obvious way to combine some of the columns to get another one of the columns.

(2) One of the vectors is the zero vector.

(3) There are more vectors than entries in each vector.

If the vectors pass the above three conditions, then create a matrix with the vectors as its columns. Row reduce the matrix to echelon form. If every column has a pivot, then the vectors are linearly independent. Otherwise they are linearly dependent.

Examples:

$$\text{Let } v_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad v_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \quad v_3 = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} \quad v_4 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad v_5 = \begin{bmatrix} 1 \\ 4 \\ 9 \end{bmatrix} \quad v_6 = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \quad v_7 = \begin{bmatrix} 12 \\ 5 \\ -2 \end{bmatrix}$$

(1) $\{v_2, v_3\}$ are linearly dependent since v_3 is twice v_2.

(2) $\{v_1, v_3, v_4\}$ are linearly dependent since v_1 is the zero vector.

(3) $\{v_2, v_4\}$ are linearly independent. To see this, make a matrix whose columns are v_2 and v_4 and row reduce it to an echelon form. The echelon form has pivots in every column.

$$\begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} \mathbf{1} & 1 \\ 0 & \mathbf{-1} \\ 0 & 0 \end{pmatrix}$$

(4) $\{v_2, v_4, v_5, v_6\}$ are linearly dependent because there are 4 vectors, but only 3 entries in each.

(5) $\{v_2, v_4, v_6\}$ are linearly dependent. To see this, make a matrix whose columns are v_2, v_4, and v_7 and row reduce it to an echelon form. The echelon does not have pivots in every column.

$$\begin{pmatrix} 1 & 1 & 12 \\ 2 & 1 & 5 \\ 3 & 1 & -2 \end{pmatrix} \rightarrow \begin{pmatrix} \mathbf{1} & 1 & 12 \\ 0 & \mathbf{-1} & -19 \\ 0 & 0 & 0 \end{pmatrix}$$

Math 43 Review Notes - Chapter 4

Orthogonal Subspaces

Recall that two vectors $\mathbf{v}$ and $\mathbf{w}$ are orthogonal (i.e. perpendicular) if their dot product $\mathbf{v}\cdot\mathbf{w}$ is 0. If we think of $\mathbf{v}$ and $\mathbf{w}$ as arrays with only one column, then we can also write the dot product of $\mathbf{v}$ and $\mathbf{w}$ as a multiplication of matrices, namely $\mathbf{v}^T\mathbf{w}$. You might see this from time to time in the book.

We say that two *subspaces* V and W of a vector space are orthogonal if every vector in V is orthogonal to every vector in W. To determine if two subspaces are orthogonal, it is enough just to check that each vector in a basis for V is orthogonal to each vector in a basis for W. Here is an important example of orthogonal subspaces:

Example: The nullspace and row space of a matrix A are orthogonal subspaces. Remember, vectors in the nullspace of a matrix A are vectors $\mathbf{x}$ for which $A\mathbf{x} = \mathbf{0}$, and vectors in the row space are linear combinations of the rows of A. A small example is enough to understand why the subspaces are orthogonal. Let

$$A = \begin{pmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \end{pmatrix}.$$

Let $\mathbf{x} = (x_1, x_2, x_3)$ be in the nullspace of A. Then $A\mathbf{x} = \mathbf{0}$. Write this out in equation form:

$$\begin{array}{ccccccc} 2x_1 & + & 3x_2 & + & 4x_3 & = & 0 \\ 5x_1 & + & 6x_2 & + & 7x_3 & = & 0 \end{array}$$

Looking closely at these, we see the first equation is just the dot product of row 1 and $\mathbf{x}$, and the second is the dot product of row 2 and $\mathbf{x}$. Since these are zero, the equations are saying that row 1 and row 2 are orthogonal to $\mathbf{x}$. So any combination of the rows is also orthogonal to $\mathbf{x}$. Therefore we have shown that the row space and nullspace are orthogonal.

Orthogonal Complement

The **orthogonal complement** of a subspace V consists of *all* vectors which are orthogonal to each vector in V. It is denoted by $V^\perp$. Like before, to find $V^\perp$ it is enough to find all vectors which are orthogonal each vector in a basis for V.

Example: Let V be the subspace which consists of all vectors $\mathbf{x} = (x_1, x_2, x_3)$ which satisfy $x_1 + 3x_2 + 2x_3 = 0$ (a plane in $\mathbb{R}^3$). Find $V^\perp$.

Solution: First find a basis for V. Solve the equation defining V to get $x_1 = -3x_2 - 2x_3$. Then any $\mathbf{x}$ in V can be written as

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -3x_2 - 2x_3 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$$

So $(-3, 0, 1)$ and $(-2, 1, 0)$ are a basis for V. (You can check for yourself that they are linearly indendent.) Any vector $\mathbf{b} = (b_1, b_2, b_3)$ in $V^\perp$ is orthogonal to both of these vectors. In other words:

$$\begin{array}{l} (b_1,b_2,b_3)\cdot(-3,0,1) = 0 \\ (b_1,b_2,b_3)\cdot(-2,1,0) = 0 \end{array} \longrightarrow \begin{array}{l} -3b_1 + 0 + b_3 = 0 \\ -2b_1 + b_2 + 0 = 0 \end{array} \longrightarrow \begin{array}{l} b_3 = 3b_1 \\ b_2 = 2b_1 \end{array}$$

Therefore, plugging these into $\mathbf{b}$ we get

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ 2b_1 \\ 3b_1 \end{bmatrix} = b_1 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

So, $V^\perp$ consists of all multiples of the vector $(1, 2, 3)$. ($V^\perp$ is a line in $\mathbb{R}^3$, perpendicular to the plane $x_1 + 3x_2 + 2x_3 = 0$.)

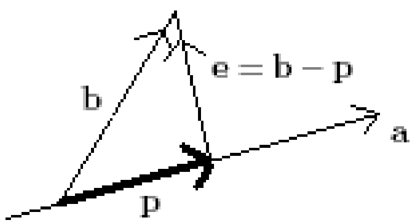
Example: Not only are the row space and nullspace orthogonal, but in fact they are orthogonal complements of each other. Try to see why.

Make sure you understand the difference between the definitions. The first definition is about a relationship between two subspaces: V and W are orthogonal if every vector in V is orthogonal to every vector in W. Notice also that the only vector that can be in both V and W is $\mathbf{0}$, since it is the only vector whose dot product with itself is 0 (i.e. it is the only vector which is perpendicular to itself).

The second definition gives a new subspace $V^{\perp}$, the orthogonal complement, which contains *every single vector* which is orthogonal to all the vectors in V.

Projections

First we will look at projecting a vector $\mathbf{b}$ onto another vector $\mathbf{a}$. One way to think about a projection is as the"shadow" that $\mathbf{b}$ casts on $\mathbf{a}$. Another way is to think of it as asking how much of $\mathbf{b}$ is in the direction of $\mathbf{a}$.



The projection of $\mathbf{b}$ onto $\mathbf{a}$ is given by

$$\mathbf{p} = \frac{\mathbf{a} \cdot \mathbf{b}}{\mathbf{a} \cdot \mathbf{a}} \mathbf{a}$$

The fractional term gives the percentage or fraction of $\mathbf{a}$ that the shadow of $\mathbf{b}$ takes up; in the book it is called $\widehat{x}$.

Example: Project $\mathbf{b} = (2, 4, 1)$ onto $\mathbf{a} = (3, 2, 5)$ and $\mathbf{c} = (3, -1, -2)$.

Solution: Projecting $\mathbf{b}$ onto $\mathbf{a}$ we get

$$\mathbf{p} = \frac{(2,4,1) \cdot (3,2,5)}{(3,2,5) \cdot (3,2,5)} \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix} = \frac{19}{38} \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 3/2 \\ 1 \\ 5/2 \end{bmatrix}$$

Projecting $\mathbf{b}$ onto $\mathbf{c}$ we get

$$\mathbf{p} = \frac{(2,4,1) \cdot (3,-1,-2)}{(3,-1,-2) \cdot (3,-1,-2)} \begin{bmatrix} 3 \\ -1 \\ -2 \end{bmatrix} = \frac{0}{14} \begin{bmatrix} 3 \\ -1 \\ -2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The vectors $\mathbf{b}$ and $\mathbf{c}$ are perpendicular (since their dot product is 0), hence the projection of $\mathbf{b}$ onto $\mathbf{c}$ is just $\mathbf{0}$, as we would expect.

There's one other way to think about projections. Instead of just projecting $\mathbf{b}$ onto $\mathbf{a}$, think about it as is projecting $\mathbf{b}$ onto the line through $\mathbf{a}$. Then $\mathbf{p}$ is the closest vector to $\mathbf{b}$ which lies on the line. The vector $\mathbf{e}$ in the picture is the "error" between $\mathbf{b}$ and $\mathbf{p}$. It is just $\mathbf{b} - \mathbf{p}$.

Since a line is just a special kind of subspace, this thinking gives us a way to project a vector $\mathbf{b}$ onto any subspace V. The projection will be the vector in V which is closest to $\mathbf{b}$. We will consider projecting onto the column space of a matrix A. In other words, we are looking for the vector which is a linear combination of the columns of A which is closest to $\mathbf{b}$. To do this we use the following:

$$P = A(A^T A)^{-1} A^T$$
$$\mathbf{p} = P\mathbf{b}$$

The matrix P is called the projection matrix. It will project any vector onto the column space of A.

Example: Let $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{pmatrix}$. Find the projection matrix P, and the projection of $\mathbf{b} = (1, 2, 3)$ onto the column space of A.

Solution: Compute:

$$A^T A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

$$(A^T A)^{-1} = \frac{1}{3} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

$$A(A^T A)^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{pmatrix} \frac{1}{3} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2 & -1 \\ -1 & 2 \\ -1 & -1 \end{pmatrix}$$

$$P = A(A^T A)^{-1} A^T = \begin{pmatrix} 2 & -1 \\ -1 & 2 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$$

$$\mathbf{p} = P\mathbf{b} = \frac{1}{3} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

Orthogonal Bases

Recall that vectors are orthogonal if their dot product is 0. If in addition to being orthogonal the vectors are all unit vectors (i.e. have length 1), then we say the vectors are **orthonormal**. A convenient way to say this is that the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v_n}$ are orthonormal if $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ unless $i = j$, in which case it equals 1.

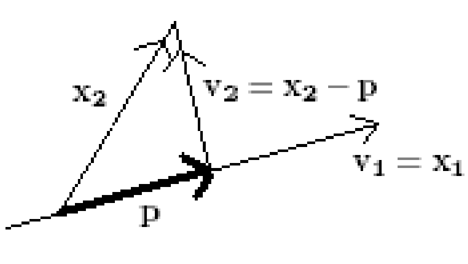
Example: Let $\mathbf{v}_1 = \frac{1}{\sqrt{11}} \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}$, $\mathbf{v}_3 = \frac{1}{\sqrt{66}} \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}$.

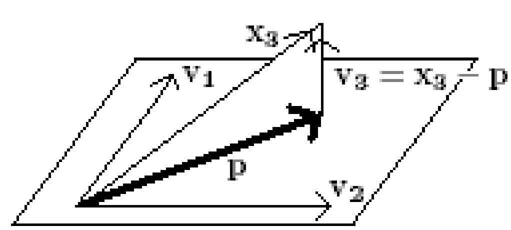
It is easy to check that $\mathbf{v}_1 \cdot \mathbf{v}_1 = \mathbf{v}_2 \cdot \mathbf{v}_2 = \mathbf{v}_3 \cdot \mathbf{v}_3 = 1$, and $\mathbf{v}_1 \cdot \mathbf{v}_2 = \mathbf{v}_1 \cdot \mathbf{v}_3 = \mathbf{v}_2 \cdot \mathbf{v}_3 = 0$. So the vectors are orthonormal.

The Gram-Schmidt Process

We say that a basis is an orthogonal basis if its vectors are orthogonal, and is an orthonormal basis if its vectors are orthonormal. For many applications orthogonal and orthonormal bases are much easier to work with than other bases. (Remember that a space has many different bases.) The Gram-Schmidt process gives a way of converting a basis into an orthogonal or orthonormal basis. You start with n linearly independent vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x_n}$ and Gram-Schmidt will create orthogonal vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v_n}$ which span the same space as the $\mathbf{x}'s$. In other words, Gram-Schmidt takes a basis for a subspace and turns it into an orthogonal basis for the same subspace.

The idea relies on projections. Start by letting $\mathbf{v}_1 = \mathbf{x}_1$. Then let $\mathbf{p}$ be the projection of $\mathbf{x}_2$ onto $\mathbf{v}_1$. Notice that $\mathbf{x}_2 - \mathbf{p}$ is orthogonal to $\mathbf{v}_1$. So let this be $\mathbf{v}_2$. Notice that $\mathbf{v}_1$ and $\mathbf{v}_2$ span the same space as $\mathbf{x}_1$ and $\mathbf{x}_2$. Now consider the projection $\mathbf{p}$ of $\mathbf{x}_3$ onto the subspace spanned by $\mathbf{v}_1$ and $\mathbf{v}_2$. Like before, $\mathbf{x}_3 - \mathbf{p}$ is orthogonal to both of $\mathbf{v}_1$ and $\mathbf{v}_2$, and since $\mathbf{v}_1$ and $\mathbf{v}_2$ are orthogonal, it has a simple formula (no messy matrix calculations). Continue projecting to get the rest of the $\mathbf{v}$'s.





Here is the process:

$$\mathbf{v}_1 = \mathbf{x}_1$$

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\,\mathbf{v}_1$$

$$\mathbf{v}_3 = \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\,\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\,\mathbf{v}_2$$

$$\mathbf{v}_4 = \mathbf{x}_4 - \frac{\mathbf{x}_4 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\,\mathbf{v}_1 - \frac{\mathbf{x}_4 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\,\mathbf{v}_2 - \frac{\mathbf{x}_4 \cdot \mathbf{v}_3}{\mathbf{v}_3 \cdot \mathbf{v}_3}\,\mathbf{v}_3$$

and so on ...

To make this set of vectors orthonormal, the vectors need to be unit vectors. Remember how to do that, we divide each vector by its length, i.e. our orthonormal vectors are

$$\frac{\mathbf{v}_1}{\| \mathbf{v}_1 \|}, \frac{\mathbf{v}_2}{\| \mathbf{v}_2 \|}, \frac{\mathbf{v}_3}{\| \mathbf{v}_3 \|}, \ldots$$

.

Example: Find a set of orthonormal vectors which span the same space as $\mathbf{x}_1 = (1, 1, 1)$, $\mathbf{x}_2 = (1, 2, 3)$, and $\mathbf{x}_3 = (4, 3, 8)$.

Solution: Check for yourself that the given vectors are linearly independent. Now use Gram-Schmidt to find the corresponding orthogonal vectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$.

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\mathbf{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \frac{1+2+3}{1+1+1}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

$$\mathbf{v}_3 = \begin{bmatrix} 4 \\ 3 \\ 8 \end{bmatrix} - \frac{4+3+8}{1+1+1}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \frac{-4+0+8}{1+0+1}\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

To make $\mathbf{v}_1$,$\mathbf{v}_2$, and $\mathbf{v}_3$ orthonormal, divide each by its length to get

$$\frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{6}}\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}.$$

Update: (12/13/05) *An easier way to find $V^{\perp}$.*

This method uses the fact that the row space and nullspace of a matrix are orthogonal. To find $V^{\perp}$, first find a basis for V. Next create a matrix A from the basis vectors by letting them be the *rows* of A. Then $V^{\perp}$ is the nullspace of A. In fact, this is really very similar to the method on the first page.

Math 43 Review Notes - Chapters 3, 5, & 6

A **basis** for a vector space is a collection of vectors that satisfy:

1. The vectors are linearly independent. (No redundant vectors)
2. Every element in the vector space is a linear combination of the basis vectors.

These say that a basis is a set from which you can obtain every vector in the vector space, and that this set is as small as possible. Many properties of a vector space can be proved just by proving them for the basis vectors. A typical basis has only a few vectors, whereas the vector space probably has infinitely many, so this is a big help.

Example: The vectors $(1, 0)$ and $(0, 1)$ are a basis for $\mathbb{R}^2$, the set of all vectors with two components. It is easy to see that they are linearly independent, and any vector with two components can be written in terms of them. For example,

$$\begin{bmatrix} 7 \\ -5 \end{bmatrix} = 7 \begin{bmatrix} 1 \\ 0 \end{bmatrix} - 5 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \text{ and in general, } \begin{bmatrix} a \\ b \end{bmatrix} = a \begin{bmatrix} 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Similarly, $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$ are a basis for $\mathbb{R}^3$. Can you find a similar basis for $\mathbb{R}^n$, the set of all vectors with n components?

A space has many different bases. For example, given a basis, just multiply each basis vector by a constant to get a new basis. In the above example, for instance, $(2, 0)$ and $(0, 2)$ are also a basis for $\mathbb{R}^2$. Note however that all the bases for a space have to have the same number of vectors. This number is called the **dimension** of the space.

Example: Find a basis for the subspace of $\mathbb{R}^3$ consisting of all vectors whose components add up to 0.

Solution: Let $\mathbf{b} = (b_1, b_2, b_3)$ be any vector in the subspace. Then the sum of its components is 0, i.e., $b_1 + b_2 + b_3 = 0$. Solve this for b_1 to get $b_1 = -b_2 - b_3$. Now write

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} -b_2 - b_3 \\ b_2 \\ b_3 \end{bmatrix} = b_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + b_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

We have just shown that any vector in the subspace can be written as a linear combination of the vectors $(-1, 1, 0)$ and $(-1, 0, 1)$, and you can easily check that they are linearly independent, thus these two vectors are a basis for the subspace. Since there are two vectors in the basis, the dimension is of the subspace is 2.

Often we will have a set of vectors that we know satisfies property (2) of a basis, i.e. any element in the space is a combination of the vectors. We call such a set a **spanning set**, or we say it **spans** the vector space. Suppose we are given a set of vectors and we want to find a basis for the set they span. They automatically satisfy property (2) of a basis, but not necessarily property (1), since there may be redundant vectors. To get rid of the redundant vectors we can make a matrix with the vectors as its columns and row reduce to find the pivot columns. The pivot columns correspond to the vectors we keep for our basis, while the other columns correspond to the redundant vectors which we will throw away.

Example: Find a basis for the set spanned by the vectors $(1, 2, 3)$, $(1, 4, 4)$, and $(-1, 2, -1)$.

Solution: As we said above, property (2) is automatically satisfied here, so we just have to throw out the redundant vectors. Following the above discussion,

$$\begin{pmatrix} 1 & 1 & -1 \\ 2 & 4 & 2 \\ 3 & 4 & -1 \end{pmatrix} \longrightarrow \begin{pmatrix} \mathbf{1} & 1 & -1 \\ 0 & \mathbf{2} & 4 \\ 0 & 0 & 0 \end{pmatrix}$$

The first two columns are the pivot columns, hence the corresponding vectors $(1, 2, 3)$ and $(1, 4, 4)$ are a basis. Note that depending on which row operations you do, you might get a different answer. But that's ok, as there are many possible bases for a space.

Example: Do the vectors $(1,1,1)$, $(1,2,3)$, and $(1,4,9)$ form a basis for $\mathbb{R}^3$?.

Solution: We have to verify properties (1) and (2) here. We can do this in one big step. Row reduce the matrix with the vectors as its columns. If there is a pivot in every column, then the vectors are linearly independent. If there is a pivot in every row, then the vectors span $\mathbb{R}^3$. (This is true since having a pivot in every row means that $A\mathbf{x} = \mathbf{b}$ has a solution for any $\mathbf{b}$, which means that any vector $\mathbf{b}$ in $\mathbb{R}^3$ can be written as a linear combination of the columns of A.)

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix} \longrightarrow \begin{pmatrix} \mathbf{1} & 1 & 1 \\ 0 & \mathbf{1} & 3 \\ 0 & 0 & \mathbf{2} \end{pmatrix}$$

There is a pivot in every row and every column, so the vectors do form a basis.

Bases for Nullspace, Row Space and Column Space

Let A be a matrix with echelon form U and rref R.

Nullspace

The nullspace of A consists of all vectors $\mathbf{x}$ for which $A\mathbf{x} = \mathbf{0}$. To find it, reduce A to rref R.

$$A = \begin{pmatrix} 1 & 2 & -3 & -9 & 4 \\ 1 & 2 & 0 & 3 & 6 \\ 2 & 4 & -3 & -6 & 12 \end{pmatrix} \longrightarrow R = \begin{pmatrix} \mathbf{1} & 2 & 0 & 3 & 0 \\ 0 & 0 & \mathbf{1} & 4 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{1} \end{pmatrix}$$

Write $R\mathbf{x} = \mathbf{0}$ in equation form and solve for the pivot variables.

$$\begin{aligned} \mathbf{x_1} + 2x_2 + 3x_4 &= 0 \\ \mathbf{x_3} + 4x_4 &= 0 \\ \mathbf{x_5} &= 0 \end{aligned} \qquad \longrightarrow \qquad \begin{aligned} \mathbf{x_1} &= -2x_2 - 3x_4 \\ \mathbf{x_3} &= -4x_4 \\ \mathbf{x_5} &= 0 \end{aligned}$$

Use this to find an expression for a typical vector in the nullspace.

$$\mathbf{x} = \begin{bmatrix} \mathbf{x_1} \\ x_2 \\ \mathbf{x_3} \\ x_4 \\ \mathbf{x_5} \end{bmatrix} = \begin{bmatrix} -2x_2 - 3x_4 \\ x_2 \\ -4x_4 \\ x_4 \\ 0 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -3 \\ 0 \\ -4 \\ 1 \\ 0 \end{bmatrix}$$

So any vector in the nullspace can be written as a combination of the vectors $(-2,1,0,0,0)$ and $(-3,0,-4,1,0)$. It turns out that the vectors one gets at this point are always linearly independent. So these two vectors are a basis. In general, then the vectors one gets at the last step are a basis. Notice that there is one vector for each free variable, so the dimension of the nullspace is equal to the number of free variables (i.e. the number of non-pivot columns). In this example the dimension is 2.

Column Space

The column space of a matrix A consists of all vectors which are linear combinations of its columns. Another way to think of it is as all vectors $\mathbf{b}$ for which $A\mathbf{x} = \mathbf{b}$ has a solution. The columns of A are a spanning set, so to get a basis we need to throw out any redundant vectors. To do this we row reduce A to an echelon form to see which columns have pivots (any echelon form will do, rref is ok, but not necessary). The pivot columns of A are a basis, and the dimension of the column space is the number of pivots. For example,

$$A = \begin{pmatrix} 1 & 2 & -3 & -9 & 4 \\ 1 & 2 & 0 & 3 & 6 \\ 2 & 4 & -3 & -6 & 12 \end{pmatrix} \longrightarrow U = \begin{pmatrix} \mathbf{1} & 2 & -3 & -9 & 4 \\ 0 & 0 & \mathbf{3} & 12 & 2 \\ 0 & 0 & 0 & 0 & \mathbf{2} \end{pmatrix}$$

Columns 1,3, and 5 are the pivot columns, so they are a basis, i.e $(1,1,2)$, $(-3,0,-3)$, and $(4,6,12)$ are a basis. Make sure to use the columns of A itself. The dimension in this case is 3.

Row Space

The row space of a matrix A consists of all vectors which are linear combinations of its rows. Another way to think of it is as the column space of A^T. The rows of A are a spanning set, so to get a basis we need to throw out any redundant vectors. To do this we row reduce A to an echelon form to see which rows have pivots (as before, any echelon form will do). The pivot rows of A (or U or R) form a basis. The dimension of the row space is the number of pivots. For example,

$$A = \begin{pmatrix} 1 & 2 & -3 & -9 & 4 \\ 1 & 2 & 0 & 3 & 6 \\ 2 & 4 & -3 & -6 & 12 \end{pmatrix} \longrightarrow U = \begin{pmatrix} \mathbf{1} & 2 & -3 & -9 & 4 \\ 0 & 0 & \mathbf{3} & 12 & 2 \\ 0 & 0 & 0 & 0 & \mathbf{2} \end{pmatrix}$$

Rows 1, 2, and 3 are pivot rows, so they are a basis, i.e., $(1, 2, -3, -9, 4)$, $(1, 2, 0, 3, 6)$, and $(2, 4, -3, -6, -12)$ are a basis. Notice here that you could instead use the pivot rows of U or R as a (nicer) basis. The dimension in this case is 3.

Remarks

Row operations don't change the nullspace or the row space; that's why we can use the pivot rows of A, U, or R as a basis for the row space. However, row operations do change the column space, so it's important that we only use the pivot columns of A itself as a basis for the column space of A.

Notice also that the dimensions of the row space and the column space are always the same number r, and r plus the dimension of the nullspace equals the number of columns of A.

Determinants

The determinant of a matrix is a single number that contains a lot of information about the matrix. The determinant of A is denoted either by $\det(A)$ or $|A|$ and it only applies to square matrices (matrices with the same number of rows as columns). We have three different ways of finding determinants. The best way is often to use a combination of the three.

Method 1: Row reduction

(1) The determinant of a triangular matrix (all zeroes above or below the diagonal) is the product of the entries on the diagonal.

(2) Switching two rows switches the sign of the determinant.

(3) Multiplying a row by a number multiplies the determinant by that number, so you'll have to *divide* your answer by that number to cancel it out.

(4) Adding or subtracting a multiple of one row from another doesn't affect the determinant.

(5) Row operations where the row you replace is also multiplied by a number do change the determinant (say you replace row 1 with $3\,\text{row } 1 - 2\,\text{row } 2$, for example). If you replace row j with $m\,\text{row } j$ - $n\,\text{row } k$, then this multiplies the determinant by m, so you have to divide your answer by m to cancel out its effect.

Combining these gives us a method for finding determinants. Row reduce the matrix to a triangular form, and then multiply the entries on the diagonal. Make sure to take into account signs from row switches, and be sure to divide by any multiples you introduced if you do the row operations in (3) or (5).

Example: Find the determinant of $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$.

Solution: Switch rows 1 and 3. The resulting matrix is triangular. The minus sign comes from the row switch.

$$\begin{vmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix} = -\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = -(1 \cdot 1 \cdot 1) = -1$$

Example: Find the determinant of $\begin{pmatrix} 2 & 1 & 3 \\ 4 & 3 & 6 \\ 2 & 3 & 8 \end{pmatrix}$.

Solution: Row reduce to a triangular matrix.

$$\begin{vmatrix} 2 & 1 & 3 \\ 4 & 3 & 6 \\ 2 & 3 & 8 \end{vmatrix} \overset{r2-2r1}{=} \begin{vmatrix} 2 & 1 & 3 \\ 0 & 1 & 0 \\ 2 & 3 & 8 \end{vmatrix} \overset{r3-r1}{=} \begin{vmatrix} 2 & 1 & 3 \\ 0 & 1 & 0 \\ 0 & 2 & 5 \end{vmatrix} \overset{r3-2r2}{=} \begin{vmatrix} 2 & 1 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{vmatrix} = 2 \cdot 1 \cdot 5 = 10$$

The row operations are of the type in (4) that doesn't affect the determinant.

Example: Find the determinant of $\begin{pmatrix} 3 & 2 & 1 \\ 7 & 5 & 2 \\ 0 & 0 & 4 \end{pmatrix}$.

Solution: Row reduce to a triangular matrix.

$$\begin{vmatrix} 3 & 2 & 1 \\ 7 & 5 & 2 \\ 0 & 0 & 4 \end{vmatrix} \overset{7r1}{=} \frac{1}{7} \begin{vmatrix} 21 & 14 & 7 \\ 7 & 5 & 2 \\ 0 & 0 & 4 \end{vmatrix} \overset{3r2}{=} \frac{1}{3} \cdot \frac{1}{7} \begin{vmatrix} 21 & 14 & 7 \\ 21 & 15 & 6 \\ 0 & 0 & 4 \end{vmatrix} \overset{r2-r1}{=} \frac{1}{3} \cdot \frac{1}{7} \begin{vmatrix} 21 & 14 & 7 \\ 0 & 1 & -1 \\ 0 & 0 & 4 \end{vmatrix} = \frac{1}{3} \cdot \frac{1}{7} \cdot 21 \cdot 1 \cdot 4 = 4$$

Another way to do this using the row operation in (5) would be

$$\begin{vmatrix} 3 & 2 & 1 \\ 7 & 5 & 2 \\ 0 & 0 & 4 \end{vmatrix} \overset{3r2-7r1\to r2}{=} \frac{1}{3} \begin{vmatrix} 3 & 2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 4 \end{vmatrix} = \frac{1}{3} \cdot 3 \cdot 1 \cdot 4 = 4$$

If you don't mind the fractions, then you could also do this problem just by subtracting (7/3)row 1 from row 2 to reduce to triangular form. This has the benefit of not having any multiples to remember to cancel out.

Method 2: Big Formula

The determinant can be given by a formula just in terms of the entries of the matrix. It is most useful when the matrix is 3×3 or smaller, since for an $n \times n$ matrix the formula has $n!$ terms. (For instance, when $n = 6$, that's 720 terms.)

(1) A 1×1 matrix only has one entry. The determinant is equal to that entry.

(2) For a 2×2 matrix, the formula is $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$.

(3) For a 3×3 matrix the formula is

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$$

This is already a lot to remember, but there is a shortcut *which only works in the* 3×3 *case.*

Example: (*Shortcut for* 3×3) Compute $\begin{vmatrix} 2 & 4 & 1 \\ 3 & 0 & 7 \\ 2 & 5 & 6 \end{vmatrix}$

First copy the indicated entries onto the right side of the matrix. Multiply the entries along each indicated diagonal, and add up the three values you get. This gives the first three terms in the big formula.

$$\left|\begin{matrix} 2 & 4 & 1 \\ \mathbf{3} & 0 & 7 \\ \mathbf{2} & \mathbf{5} & 6 \end{matrix}\right| \begin{matrix} & \\ \mathbf{3} & \\ \mathbf{2} & \mathbf{5} \end{matrix} \quad \longrightarrow (2)(0)(6) + (4)(7)(2) + (1)(3)(5) = 71$$

Next copy the other indicated entries to the left side of the matrix. Multiply the entries along each indicated diagonal, take the negative of each product, and add up the three values you get. This gives the last three terms in the big formula.

$$\begin{matrix} & \\ & \mathbf{7} \\ \mathbf{5} & \mathbf{6} \end{matrix} \left|\begin{matrix} 2 & 4 & 1 \\ 3 & 0 & \mathbf{7} \\ 2 & \mathbf{5} & \mathbf{6} \end{matrix}\right| \quad \longrightarrow -(1)(0)(2) - (4)(3)(6) - (2)(7)(5) = -142$$

The answer is then 71-142=-71.

(4) There is a formula which works for bigger matrices. If you look at the 3×3 formula, you'll see that for every possible permutation (x, y, z) of $(1, 2, 3)$, there's a term $a_{1x}a_{2y}a_{3z}$ (For example, $(3, 2, 1)$ and $(2, 1, 3)$ are possible permutations; there's 6 total.) Whether we add or subtract a given term is determined by how many steps it takes us to get from $(1, 2, 3)$ to (x, y, z). An even number of steps corresponds to adding, and an odd number corresponds to subtracting. (For example, it takes 3 steps to get from $(1, 2, 3)$ to $(3, 1, 2)$ — $(1, 2, 3) \to (3, 2, 1) \to (3, 1, 2)$.)

To get the 4×4 formula, use permutations of $(1, 2, 3, 4)$ instead of $(1, 2, 3)$. There should be 24 terms, one for each possible permutation. (For instance, $-a_{13}a_{24}a_{32}a_{41}$ is the term corresponding to the permutation $(3, 4, 2, 1)$.) Bigger matrices work similarly. However, with so many terms, the big formula is usually not practical for matrices larger than 3×3.

Method 3: Cofactors

Let M_{ij} be the matrix left over after crossing out row i and column j. The term $C_{ij} = (-1)^{i+j} \det(M_{ij})$ is called a cofactor. We can use cofactors to compute the determinant. The idea is to pick any row or any column of the matrix, and for each entry in that row or column, multiply the entry and its cofactor, then add up all of the products you computed. In terms of formulas,

$$\begin{array}{ll} \det A = a_{i1}C_{i1} + a_{i2}C_{i2} + \ldots + a_{in}C_{in} & \text{if you expand across row } i, \\ \det A = a_{1j}C_{1j} + a_{2j}C_{2j} + \ldots + a_{nj}C_{nj} & \text{if you expand across column } j. \end{array}$$

The $(-1)^{i+j}$ in C_{ij} gives either a plus or minus sign. One way to get the signs right is to remember the sign matrix $\begin{pmatrix} + - + \\ - + - \\ + - + \end{pmatrix}$, or for 4×4, $\begin{pmatrix} + - + - \\ - + - + \\ + - + - \\ - + - + \end{pmatrix}$. For any size matrix, start with a $+$ in the first entry, and the signs alternate from there.

The above formulas may not be very enlightening; the method is best demonstrated with some examples.

Example: Compute the determinant of $\begin{pmatrix} 2 & 3 & 4 \\ 1 & 2 & 5 \\ 0 & 0 & 6 \end{pmatrix}$ three ways,

(a) by expanding across row 1,
(b) by expanding down column 2,
(c) by expanding across row 3.

Solution:

(a) The formula says

$$\det A = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13}$$

We know $a_{11} = 2$, $a_{12} = 3$, $a_{13} = 4$, so we just have to find the cofactors.

$$A = \begin{pmatrix} \mathbf{2} & \mathbf{3} & \mathbf{4} \\ 1 & 2 & 5 \\ 0 & 0 & 6 \end{pmatrix} \quad \text{Signs} = \begin{pmatrix} + - + \\ - + - \\ + - + \end{pmatrix} \quad M_{11} = \begin{pmatrix} 2 & 3 & 4 \\ 1 & 2 & 5 \\ 0 & 0 & 6 \end{pmatrix} \quad M_{12} = \begin{pmatrix} 2 & 3 & 4 \\ 1 & 2 & 5 \\ 0 & 0 & 6 \end{pmatrix} \quad M_{13} = \begin{pmatrix} 2 & 3 & 4 \\ 1 & 2 & 5 \\ 0 & 0 & 6 \end{pmatrix}$$

$$\det A = 2\begin{vmatrix} 2 & 5 \\ 0 & 6 \end{vmatrix} - 3\begin{vmatrix} 1 & 5 \\ 0 & 6 \end{vmatrix} + 4\begin{vmatrix} 1 & 2 \\ 0 & 0 \end{vmatrix} = 2(12 - 0) - 3(6 - 0) + 4(0 - 0) = 6$$

Use the formula (2) from method 2 for evaluating each 2×2 determinant.

Streamlining things a bit, we see each term consists of two parts, first an entry in row 1, added or subtracted according to the corresponding entry in the sign matrix (or by the formula $(-1)^{i+j}$), and second, the determinant of the matrix we get by crossing out the row and column containing the entry.

(b)

$$\begin{pmatrix} 2 & \mathbf{3} & 4 \\ 1 & \mathbf{2} & 5 \\ 0 & \mathbf{0} & 6 \end{pmatrix} \quad \begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix} \quad \begin{pmatrix} 2 & 3 & 4 \\ 1 & 2 & 5 \\ 0 & 0 & 6 \end{pmatrix} \quad \begin{pmatrix} 2 & 3 & 4 \\ 1 & 2 & 5 \\ 0 & 0 & 6 \end{pmatrix} \quad \begin{pmatrix} 2 & 3 & 4 \\ 1 & 2 & 5 \\ 0 & 0 & 6 \end{pmatrix}$$

$$\det A = -3\begin{vmatrix} 1 & 5 \\ 0 & 6 \end{vmatrix} + 2\begin{vmatrix} 2 & 4 \\ 0 & 6 \end{vmatrix} - 0\begin{vmatrix} 2 & 4 \\ 1 & 5 \end{vmatrix} = -3(6-0) + 2(12-0) - 0(10-4) = 6$$

(c)

$$\begin{pmatrix} 2 & 3 & 4 \\ 1 & 2 & 5 \\ \mathbf{0} & \mathbf{0} & \mathbf{6} \end{pmatrix} \quad \begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix} \quad \begin{pmatrix} 2 & 3 & 4 \\ 1 & 2 & 5 \\ 0 & 0 & 6 \end{pmatrix} \quad \begin{pmatrix} 2 & 3 & 4 \\ 1 & 2 & 5 \\ 0 & 0 & 6 \end{pmatrix} \quad \begin{pmatrix} 2 & 3 & 4 \\ 1 & 2 & 5 \\ 0 & 0 & 6 \end{pmatrix}$$

$$\det A = 0\begin{vmatrix} 3 & 4 \\ 2 & 5 \end{vmatrix} - 0\begin{vmatrix} 2 & 4 \\ 1 & 5 \end{vmatrix} + 6\begin{vmatrix} 2 & 3 \\ 1 & 2 \end{vmatrix} = 0(15-8) - 0(10-4) + 6(4-3) = 6$$

We didn't have to write the last two terms in (c) since they were multiplied by 0. As you can see, the easiest of the three to compute is (c). When evaluating a determinant by cofactors, you get to pick the row or column to expand across. Usually it is easiest to expand across the one with the most zeros.

The Fastest Way:

Probably the fastest way to compute determinants, especially 4×4 determinants, is to combine the methods. Use row operations to create a row or a column with a lot of zeroes, and then use cofactor expansion down that row or column.

Example: Find the determinant of $\begin{pmatrix} 1 & 4 & 2 & 3 \\ 3 & 8 & 7 & 9 \\ 2 & 0 & 1 & 3 \\ 5 & 4 & 3 & 5 \end{pmatrix}$.

Solution: Subtract 2 row 1 from row 2 and subtract row 1 from row 3. These operations don't change the determinant. Then expand down column 2. We get

$$\begin{vmatrix} 1 & 4 & 2 & 3 \\ 3 & 8 & 7 & 9 \\ 2 & 0 & 1 & 3 \\ 5 & 4 & 3 & 5 \end{vmatrix} \overset{\substack{r2-2r1 \\ r4-r1}}{=} \begin{vmatrix} 1 & 4 & 2 & 3 \\ 1 & 0 & 3 & 3 \\ 2 & 0 & 1 & 3 \\ 4 & 0 & 1 & 2 \end{vmatrix} \overset{Expand}{=} -4\begin{vmatrix} 1 & 3 & 3 \\ 2 & 1 & 3 \\ 4 & 1 & 2 \end{vmatrix}$$

Now compute the 3×3 determinant using any method. The answer is $-4(17) = -68$.

Properties of Determinants

(1) $\det A = 0 \Leftrightarrow A$ has no inverse. In other words: if $\det A = 0$, then you know A has no inverse, and conversely, if you know A has no inverse, then it must be true that $\det A = 0$.

(2) $\det AB = (\det A)(\det B)$.

(3) $\det A^{-1} = 1/\det A$. Try to prove this in one line using property (2) and the fact that $AA^{-1} = I$.

(4) $\det A^T = \det A$. This rule implies that column operations affect the determinant in the same way that row operations do. (Column operations are like row operations, except that you perform them on the columns instead of the rows.)

Cramer's Rule

Cramer's Rule is a method for solving $A\mathbf{x} = \mathbf{b}$ using determinants. Since it uses determinants, it is really only practical for hand computations on small matrices, however it is useful for proving things, as well as for hand computations when row operations would involve a lot of fractions.

Method:

(1) Compute $\det A$. If you get 0, stop. Cramer's rule won't work.

(2) Let B_1 be the matrix you get by replacing column 1 of A with $\mathbf{b}$. Let B_2 be the matrix you get by replacing column 2 of A with $\mathbf{b}$, etc.

(3) Then the solution is given by

$$x_1 = \frac{\det B_1}{\det A}, \qquad x_2 = \frac{\det B_2}{\det A}, \qquad x_3 = \frac{\det B_3}{\det A}, \quad \ldots$$

Example: Use Cramer's rule to solve $A\mathbf{x} = \mathbf{b}$ where $A = \begin{pmatrix} 2 & 1 & 2 \\ 5 & 7 & 3 \\ 3 & 1 & 2 \end{pmatrix}$ and $\mathbf{b} = \begin{bmatrix} 1 \\ 3 \\ 6 \end{bmatrix}$.

Solution: Compute the following:

$$\det A = \begin{vmatrix} 2 & 1 & 2 \\ 5 & 7 & 3 \\ 3 & 1 & 2 \end{vmatrix} = -11 \qquad \det B_1 = \begin{vmatrix} \mathbf{1} & 1 & 2 \\ \mathbf{3} & 7 & 3 \\ \mathbf{6} & 1 & 2 \end{vmatrix} = -55$$

$$\det B_2 = \begin{vmatrix} 2 & \mathbf{1} & 2 \\ 5 & \mathbf{3} & 3 \\ 3 & \mathbf{6} & 2 \end{vmatrix} = 17 \qquad \det B_3 = \begin{vmatrix} 2 & 1 & \mathbf{1} \\ 5 & 7 & \mathbf{3} \\ 3 & 1 & \mathbf{6} \end{vmatrix} = 41$$

$$\text{Therefore} \quad x_1 = 5, \qquad x_2 = -\frac{17}{11}, \qquad x_3 = -\frac{41}{11}.$$

Inverses by Cofactors:

We can regard finding A^{-1} as solving the equation $AA^{-1} = I$ for A^{-1}. We can convert this into a system of equations of the form $A\mathbf{x} = \mathbf{b}$ which we can solve by Cramer's Rule. In the end we get the following method for finding A^{-1}. Just as with Cramer's rule, it's usually not practical for hand computations on large matrices, but it is useful for theoretical calculations and hand computations on 2×2 and 3×3 matrices, especially when row operations involve lots of fractions.

Method:

(1) Compute $\det A$. If you get 0, then stop as there is no inverse.

(2) Compute all the cofactors and put them into a matrix C, so that C_{ij} is in row i, column j.

(3) A^{-1} is given by $C^T / \det A$.

One way to do step 2 is to compute all the determinants you get by crossing out a row and a column of A, and put them into a matrix. The determinant calculated by crossing out row i and column j goes into row i, column j of this new matrix. Then get the signs right by using the sign matrix . Don't touch entries which are in the same location as the + signs of the sign matrix . However, do change the sign of entries which are in the same location as the − signs of the sign matrix.

Example: Find the inverse of $\begin{pmatrix} 0 & 1 & 3 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix}$.

Solution:

(1) Compute $\det A$ to get 5.

(2) Compute all the determinants you get by crossing out a row and a column of A, and put them into a matrix.

$$\begin{matrix} \begin{smallmatrix} 0&1&3\\1&0&1\\2&1&0 \end{smallmatrix} & \begin{smallmatrix} 0&1&3\\1&0&1\\2&1&0 \end{smallmatrix} & \begin{smallmatrix} 0&1&3\\1&0&1\\2&1&0 \end{smallmatrix} \\ \begin{smallmatrix} 0&1&3\\1&0&1\\2&1&0 \end{smallmatrix} & \begin{smallmatrix} 0&1&3\\1&0&1\\2&1&0 \end{smallmatrix} & \begin{smallmatrix} 0&1&3\\1&0&1\\2&1&0 \end{smallmatrix} \\ \begin{smallmatrix} 0&1&3\\1&0&1\\2&1&0 \end{smallmatrix} & \begin{smallmatrix} 0&1&3\\1&0&1\\2&1&0 \end{smallmatrix} & \begin{smallmatrix} 0&1&3\\1&0&1\\2&1&0 \end{smallmatrix} \end{matrix} \longrightarrow \begin{pmatrix} \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 2 & 0 \end{vmatrix} & \begin{vmatrix} 1 & 0 \\ 2 & 1 \end{vmatrix} \\ \begin{vmatrix} 1 & 3 \\ 1 & 0 \end{vmatrix} & \begin{vmatrix} 0 & 3 \\ 2 & 0 \end{vmatrix} & \begin{vmatrix} 0 & 1 \\ 2 & 1 \end{vmatrix} \\ \begin{vmatrix} 1 & 3 \\ 0 & 1 \end{vmatrix} & \begin{vmatrix} 0 & 3 \\ 1 & 1 \end{vmatrix} & \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} \end{pmatrix} = \begin{pmatrix} -1 & -2 & 1 \\ -3 & -6 & -2 \\ 1 & -3 & -1 \end{pmatrix}$$

Now apply the sign matrix.

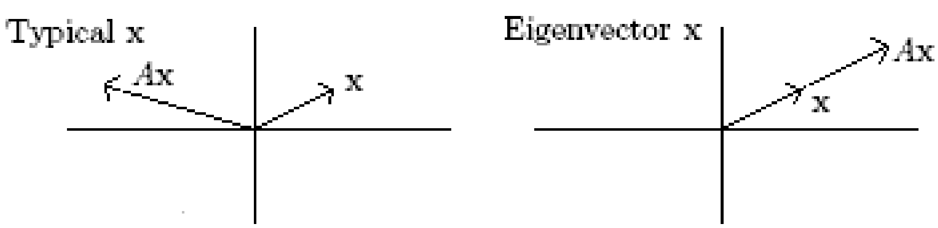
$$\begin{pmatrix} -1 & -2 & 1 \\ -3 & -6 & -2 \\ 1 & -3 & -1 \end{pmatrix} \longrightarrow \begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix} \longrightarrow \begin{pmatrix} -1 & 2 & 1 \\ 3 & -6 & 2 \\ 1 & 3 & -1 \end{pmatrix} = C$$

(3) Finally, transpose C and divide by $\det A$.

$$A^{-1} = \frac{1}{5} \begin{pmatrix} -1 & 3 & 1 \\ 2 & -6 & 3 \\ 1 & 2 & -1 \end{pmatrix}$$

Eigenvalues and Eigenvectors

Consider multiplying a vector $\mathbf{x}$ by a matrix A. The resulting vector $A\mathbf{x}$ most likely is of a different length and points in a different direction than $\mathbf{x}$.



Those vectors $\mathbf{x}$ for which $A\mathbf{x}$ points in the same direction as $\mathbf{x}$ are called **eigenvectors**. Essentially this means that multiplication by A on $\mathbf{x}$ acts just like multiplying $\mathbf{x}$ by a constant. If we call the constant λ, then we get the equation

$$A\mathbf{x} = \lambda\mathbf{x} \quad \text{if } \mathbf{x} \text{ is an eigenvector.}$$

The constant λ is called an **eigenvalue**.

How to Find Eigenvalues and Eigenvectors

The equation $A\mathbf{x} = \lambda\mathbf{x}$ tells us how to find the eigenvalues and eigenvectors. Subtract $\lambda\mathbf{x}$ from both sides and factor it out to get $(A - \lambda I)\mathbf{x} = \mathbf{0}$. The only way for there to be nonzero eigenvectors (i.e. interesting eigenvectors) is if $A - \lambda I$ is not invertible. Recall that this is true when $\det(A - \lambda I) = 0$. Thus to find the eigenvalues we solve the polynomial equation given by $\det(A - \lambda I) = 0$. For each λ found above there are different eigenvectors. The eigenvectors $\mathbf{x}$ satisfy $(A - \lambda I)\mathbf{x} = \mathbf{0}$, i.e. they are elements of the nullspace of $A - \lambda I$.

Note that $A - \lambda I$ has the same entries as A except that the diagonal entries have λ subtracted from them. For example:

$$A = \begin{pmatrix} 2 & 4 \\ 3 & 5 \end{pmatrix} \qquad A - \lambda I = \begin{pmatrix} 2-\lambda & 4 \\ 3 & 5-\lambda \end{pmatrix}$$

In summary,

• To find the eigenvalues, compute $\det(A-\lambda I)$. This is a polynomial equation. Set it equal to 0 and solve for λ.

• To find the eigenvectors for an eigenvalue λ, find the nullspace of $A-\lambda I$. The vectors in the nullspace are the eigenvectors corresponding to λ. Do this for all the eigenvalues.

Example: Find the eigenvalues and eigenvectors of $\begin{pmatrix} 4 & 3 \\ 1 & 2 \end{pmatrix}$.

Solution:

• First find the eigenvalues:

$$\begin{vmatrix} 4-\lambda & 3 \\ 1 & 2-\lambda \end{vmatrix} = (4-\lambda)(2-\lambda)-3 = \lambda^2-6\lambda+5 = (\lambda-1)(\lambda-5)$$

So $\lambda = 1, 5$ are the eigenvalues.

• Eigenvectors for $\lambda = 1$: Find the nullspace of $A-3I$.

$$A-3I = \begin{pmatrix} 3 & 3 \\ 1 & 1 \end{pmatrix} \xrightarrow{(1/3)r1} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \xrightarrow{r2-r1} \begin{pmatrix} \mathbf{1} & 1 \\ 0 & 0 \end{pmatrix} \to x_1+x_2=0 \quad \to x_1 = -x_2$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Therefore the eigenvectors for $\lambda = 1$ are all multiples of $(-1, 1)$.

• Eigenvectors for $\lambda = 5$: Find the nullspace of $A-5I$.

$$A-5I = \begin{pmatrix} -1 & 3 \\ 1 & -3 \end{pmatrix} \xrightarrow{r3+r1} \begin{pmatrix} -1 & 3 \\ 0 & 0 \end{pmatrix} \xrightarrow{-r1} \begin{pmatrix} \mathbf{1} & -3 \\ 0 & 0 \end{pmatrix} \to x_1-3x_2=0 \quad \to x_1 = 3x_2$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3x_2 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

Therefore the eigenvectors for $\lambda = 5$ are all multiples of $(3, 1)$.

Example: Find the eigenvalues and eigenvectors of $\begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix}$.

Solution:

• First find the eigenvalues: $\det(A-\lambda I)$ might look imposing, but use row operations to simplify it, and then expand across row 1.

$$\begin{vmatrix} 2-\lambda & 2 & 2 \\ 2 & 2-\lambda & 2 \\ 2 & 2 & 2-\lambda \end{vmatrix} \overset{\substack{r1-r3 \\ r2-r3}}{=} \begin{vmatrix} -\lambda & 0 & \lambda \\ 0 & -\lambda & \lambda \\ 2 & 2 & 2-\lambda \end{vmatrix} = -\lambda \begin{vmatrix} -\lambda & \lambda \\ 2 & 2-\lambda \end{vmatrix} + \lambda \begin{vmatrix} 0 & -\lambda \\ 2 & 2 \end{vmatrix}$$

$$= -\lambda(-\lambda(2-\lambda)-2\lambda) + \lambda(0--2\lambda) = 4\lambda^2-\lambda^3+2\lambda^2 = \lambda^2(6-\lambda)$$

Thus the eigenvalues are $\lambda = 0, 6$.

• Eigenvectors for $\lambda = 0$: Find the nullspace of $A-0I$.

$$A-0I = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix} \xrightarrow{rref} \begin{pmatrix} \mathbf{1} & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \to x_1+x_2+x_3=0 \quad \to x_1 = -x_2-x_3$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -x_2-x_3 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Therefore the eigenvectors for $\lambda = 0$ are all linear combinations of $(-1, 1, 0)$ and $(-1, 0, 1)$.

• Eigenvectors for $\lambda = 6$: Find the nullspace of $A - 6I$.

$$A - 6I = \begin{pmatrix} -4 & 2 & 2 \\ 2 & -4 & 2 \\ 2 & 2 & -4 \end{pmatrix} \xrightarrow{rref} \begin{pmatrix} \mathbf{1} & 0 & -1 \\ 0 & \mathbf{1} & -1 \\ 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{array}{l} x_1 - x_3 = 0 \\ x_2 - x_3 = 0 \end{array} \quad \rightarrow \quad \begin{array}{l} x_1 = x_3 \\ x_2 = x_3 \end{array}$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_3 \\ x_3 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Therefore the eigenvectors for $\lambda = 6$ are all multiples of $(1, 1, 1)$.

Remarks about Eigenvalues and Eigenvectors:

- There are a few ways to check your work:
  1. The product of the eigenvalues equals $\det A$.
  2. The sum of the eigenvalues equals the sum of the diagonal entries of A.
  3. Eigenvalues are numbers which make $A - \lambda I$ not invertible. Therefore, if you find that the only eigenvector you get is $\mathbf{0}$ (or equivalently, if you have a pivot in every column), then there is certainly a mistake somewhere.
- The eigenvalues of a triangular matrix are the entries on the diagonal.
- It's possible that the all the eigenvalues are imaginary numbers. This means that the only (real) eigenvector is $\mathbf{0}$.
- Finding the eigenvalues of a $n \times n$ involves solving a polynomial of degree n which is often difficult for $n > 2$.

Diagonalizing a Matrix

A diagonal matrix is a matrix whose entries above and below the diagonal are all zero. Diagonal matrices look like the identity matrix, except that the entries on the diagonal don't have to be all ones. Diagonal matrices are nice to work with because they have so many zero entries. Using eigenvalues and eigenvectors, we can rewrite a matrix in terms of a diagonal matrix. To do this, i.e. to *diagonalize* an $n \times n$ matrix, we need the matrix to have n linearly independent eigenvectors, otherwise we can't do it.

Example:

(1) If all the eigenvalues of a matrix are different, then there will be n linearly independent eigenvectors, so the matrix is diagonalizable.

(2) Suppose a 3×3 matrix has eigenvalues 3 and 4, and in finding the eigenvectors you find for $\lambda = 3$, the nullspace of $A - \lambda I$ is given by $\mathbf{x} = x_2(2, 1, 0) + x_3(-1, 0, 1)$ and for $\lambda = 4$, the nullspace of $A - \lambda I$ is given by $\mathbf{x} = x_3(3, 0, 1)$. The vectors $(2, 1, 0)$, $(-1, 0, 1)$, and $(-3, 0, 1)$ are three linearly independent eigenvectors, so the matrix is diagonalizable.

(3) Suppose in the above example, the nullspace of $A - \lambda I$ for $\lambda = 3$ didn't have the second term. Then we could only find two linearly independent eigenvectors, and so we couldn't diagonalize the matrix.

When you find the nullspace of $A - \lambda I$, the vectors at the last step are a basis for the nullspace. (For example, in (2), the vectors we're referring to are $(2, 1, 0)$ and $(-1, 0, 1)$.) It turns out that the set of all the basis vectors for all the eigenvalues is linearly independent. So you just have to count up the total number of basis vectors you find, and if the total is n, then the matrix is diagonalizable, otherwise it is not.

How to Diagonalize A Matrix

Let Λ be the diagonal matrix whose entries along the diagonal are the eigenvalues of A. Let S be the matrix whose columns are the eigenvectors of A. It is important that the order of the eigenvectors in S corresponds to the order of the eigenvalues in Λ. For example, if the eigenvalue $\lambda = -4$ is in column 3 of Λ, then the eigenvector in column 3 of S must be an eigenvector you found using $\lambda = -4$. We can *diagonalize* A by writing it as

$$A = S\Lambda S^{-1}$$

Example: Diagonalize $A = \begin{pmatrix} 2 & 1 \\ 0 & 5 \end{pmatrix}$.

Solution: The matrix is upper triangular so the eigenvalues are $\lambda = 2, 5$, the entries on the diagonal.

$$\lambda = 2: \ A - \lambda I = \begin{pmatrix} 0 & 1 \\ 0 & 3 \end{pmatrix} \xrightarrow{rref} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \rightarrow x_2 = 0 \rightarrow \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\lambda = 5: \ A - \lambda I = \begin{pmatrix} -3 & 1 \\ 0 & 0 \end{pmatrix} \xrightarrow{rref} \begin{pmatrix} 1 & -1/3 \\ 0 & 0 \end{pmatrix} \rightarrow x_1 = x_2/3 \rightarrow \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2/3 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} 1/3 \\ 1 \end{bmatrix}$$

$$\Lambda = \begin{pmatrix} 2 & 0 \\ 0 & 5 \end{pmatrix} \qquad S = \begin{pmatrix} 1 & 1 \\ 0 & 3 \end{pmatrix} \qquad S^{-1} = \frac{1}{3}\begin{pmatrix} 3 & -1 \\ 0 & 1 \end{pmatrix}$$

Notice that for the second column instead of using $(1/3, 1)$ we used $(1, 3)$. This is ok since the eigenvectors for $\lambda = 5$ are all the multiples of $(1/3, 1)$. $(1, 3)$ is such a multiple, and we chose it since it has no fractions. This wasn't necessary, but it gives a nicer S. So we diagonalize A as

$$A = S\Lambda S^{-1} = \frac{1}{3}\begin{pmatrix} 1 & 1 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & 5 \end{pmatrix}\begin{pmatrix} 3 & -1 \\ 0 & 1 \end{pmatrix}.$$

Using Diagonalization to Find Powers of A

Suppose A is diagonalizable, so that $A = S\Lambda S^{-1}$. Observe the following:

$$A^2 = (S\Lambda S^{-1})(S\Lambda S^{-1}) = S\Lambda^2 S^{-1}$$
$$A^3 = A^2 A = (S\Lambda^2 S^{-1})(S\Lambda S^{-1}) = S\Lambda^3 S^{-1}$$

We used the fact that $SS^{-1} = I$ to simplify both expressions. In general we get $A^k = S\Lambda^k S^{-1}$. This is useful because raising diagonal matrices to powers is particularly simple – just raise each diagonal entry to the power. This doesn't usually work with non-diagonal matrices. Thus to compute A^k we only need three multiplications instead of $k + 3$ multiplications.

Example: Use the diagonalization of the matrix in the example above to compute A^k.

Solution:

$$A^k = S\Lambda^k S^{-1} = \frac{1}{3}\begin{pmatrix} 1 & 1 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} 2^k & 0 \\ 0 & 5^k \end{pmatrix}\begin{pmatrix} 3 & -1 \\ 0 & 1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 2^k & 5^k \\ 0 & 3 \cdot 5^k \end{pmatrix}\begin{pmatrix} 3 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2^k & (5^k - 2^k)/3 \\ 0 & 5^k \end{pmatrix}$$

We just multiplied all the matrices together. Notice that we brought the 1/3 inside at the last step.

Math 43 Review Notes - Chapters 6 & 8.3

Symmetric Matrices and Orthogonal Diagonalization

A symmetric matrix has only real eigenvalues. Its eigenvectors can always be chosen to be orthonormal. A symmetric matrix can *always* be diagonalized, unlike other matrices. When we diagonalize a symmetric matrix we get a special diagonalization called an *orthogonal diagonalization*.

We diagonalize A as $A = Q\Lambda Q^T$, where Λ is the diagonal matrix with the eigenvalues of A on the diagonal, and Q is the matrix whose columns are *unit* eigenvectors.* Make sure that the unit eigenvectors in Q line up with the corresponding eigenvalues in Λ.

This diagonalization is very similar to the usual $S\Lambda S^{-1}$ diagonalization. The difference here is that we use *unit* eigenvectors, and since the columns of Q are orthonormal, $Q^{-1} = Q^T$, so we don't have to compute an inverse.

Example: Orthogonally diagonalize $\begin{pmatrix} 8 & 6 \\ 6 & -8 \end{pmatrix}$.

Solution: First find the eigenvalues.

$$\begin{vmatrix} 8-\lambda & 6 \\ 6 & -8-\lambda \end{vmatrix} = \lambda^2 - 100 = (\lambda - 10)(\lambda + 10).$$

Thus the eigenvalues are 10 and -10. Now find the eigenvectors and unit eigenvectors.

$$\lambda = 10: \begin{pmatrix} -2 & 6 \\ 6 & -18 \end{pmatrix} \overset{rref}{\to} \begin{pmatrix} 1 & -3 \\ 0 & 0 \end{pmatrix} \Rightarrow \begin{bmatrix} 3 \\ 1 \end{bmatrix} \text{ is an eigenvector, and } \begin{bmatrix} 3/\sqrt{10} \\ 1/\sqrt{10} \end{bmatrix} \text{ is the unit eigenvector.}$$

$$\lambda = -10: \begin{pmatrix} 18 & 6 \\ 6 & 2 \end{pmatrix} \overset{rref}{\to} \begin{pmatrix} 1 & 1/3 \\ 0 & 0 \end{pmatrix} \Rightarrow \begin{bmatrix} -1 \\ 3 \end{bmatrix} \text{ is an eigenvector, and } \begin{bmatrix} -1/\sqrt{10} \\ 3/\sqrt{10} \end{bmatrix} \text{ is the unit eigenvector.}$$

Finally, we write

$$A = Q\Lambda Q^T = \begin{pmatrix} 3/\sqrt{10} & -1/\sqrt{10} \\ 1/\sqrt{10} & 3/\sqrt{10} \end{pmatrix} \begin{pmatrix} 10 & 0 \\ 0 & -10 \end{pmatrix} \begin{pmatrix} 3/\sqrt{10} & 1/\sqrt{10} \\ -1/\sqrt{10} & 3/\sqrt{10} \end{pmatrix}.$$

Markov Matrices

A *Markov matrix* is a matrix whose entries satisfy:

1. All the entries are ≥ 0.
2. The entries in each column add up to 1.

For example, $\begin{pmatrix} .2 & .5 \\ .8 & .5 \end{pmatrix}$ and $\begin{pmatrix} 1 & 1/3 \\ 0 & 2/3 \end{pmatrix}$ are 2×2 Markov matrices.

A Markov matrix always has 1 as an eigenvalue. The eigenvectors corresponding to $\lambda = 1$ are called *steady-state eigenvectors*. To see why they are called this, recall the equation $A\mathbf{x} = \lambda\mathbf{x}$ that defines eigenvalues and eigenvectors. With $\lambda = 1$ this says $A\mathbf{x} = \mathbf{x}$; in other words multiplication by A doesn't change $\mathbf{x}$. Many real-life phenomena are modelled by the equation $\mathbf{x}_{n+1} = A\mathbf{x}_n$, with A a Markov matrix. The long range behavior of such a system is determined by the steady state eigenvectors.

Example: Find the eigenvalues and a steady-state eigenvector for $A = \begin{pmatrix} .3 & .4 \\ .7 & .6 \end{pmatrix}$.

Solution: We know right away that 1 is an eigenvalue for A, since A is a Markov matrix. To find the other, remember that the sum of the eigenvalues is equal to the sum of the diagonal entries of A. So we solve $1 + \lambda = .3 + .6$ to get the other eigenvalue $-.1$.

The steady-state eigenvectors are eigenvectors for $\lambda = 1$:

$$A - \lambda I = \begin{pmatrix} -.8 & .5 \\ .8 & -.5 \end{pmatrix} \overset{rref}{\to} \begin{pmatrix} 1 & -5/8 \\ 0 & 0 \end{pmatrix}.$$

Thus the steady state eigenvectors are multiples of $\begin{bmatrix} 5/8 \\ 1 \end{bmatrix}$. For instance, $\begin{bmatrix} 5 \\ 8 \end{bmatrix}$ or $\begin{bmatrix} 5/13 \\ 8/13 \end{bmatrix}$ are two examples.

*If one eigenvalue gives more than one linearly independent eigenvector, then you will have to orthogonalize the vectors using Gram-Schmidt or something else. We didn't consider anything like this in class, however, so don't worry about it.

More about Steady-State Eigenvectors

Markov matrices are useful for modelling many things. Below is a simple example.

Example: In any year, 92% of deer in the forest remain there, while 8% find their way into the suburbs (and people's backyards, where they eat their shrubbery). In addition, 88% of the deer in the suburbs remain there, while 12% are caught and returned into the forest. (Note that this is a simplified example; it doesn't take into account a lot of factors. Can you think of any assumptions in this model?)

Let f_n and s_n denote the number of deer in the forest and suburbs, respectively, in year n. Then we can write the above paragraph mathematically as

$$\begin{aligned} f_{n+1} &= .92f_n + .12s_n \\ s_{n+1} &= .08f_n + .88s_n \end{aligned}.$$

We can write this system of equations in matrix form as

$$\begin{bmatrix} f_{n+1} \\ s_{n+1} \end{bmatrix} = \begin{pmatrix} .92 & .12 \\ .08 & .88 \end{pmatrix} \begin{bmatrix} f_n \\ s_n \end{bmatrix}.$$

Let's call the matrix A. It is a Markov matrix, and thus it has $\lambda = 1$ as an eigenvalue. The steady-state eigenvectors (eigenvectors for $\lambda = 1$) are found by computing the nullspace of $A - \lambda I$.

$$A - \lambda I = \begin{pmatrix} -.08 & .12 \\ .08 & -.12 \end{pmatrix} \stackrel{rref}{\rightarrow} \begin{pmatrix} 1 & -3/2 \\ 0 & 0 \end{pmatrix}.$$

Thus the steady state eigenvectors are multiples of $(3/2, 1)$. One multiple without fractions is $(3, 2)$. To express this in terms of percents, divide each entry by the sum of the two entries to get $(3/5, 2/5) = (.6, .4)$.

Thus we expect that after many years (i.e., when n is large) 60% of the total deer population will be in the forest and 40% will be in the suburbs. To see precisely why this is true, notice from the matrix equation above that

$$\begin{bmatrix} f_2 \\ s_2 \end{bmatrix} = A \begin{bmatrix} f_1 \\ s_1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} f_3 \\ s_3 \end{bmatrix} = A \begin{bmatrix} f_2 \\ s_2 \end{bmatrix} = A \left(A \begin{bmatrix} f_1 \\ s_1 \end{bmatrix} \right) = A^2 \begin{bmatrix} f_1 \\ s_1 \end{bmatrix}.$$

In general, we see that $\begin{bmatrix} f_{n+1} \\ s_{n+1} \end{bmatrix} = A^n \begin{bmatrix} f_1 \\ s_1 \end{bmatrix}$.

So the population is closely related to A^n. Recall that we can use diagonalization to find A^n. We diagonalize A as $A = S\Lambda S^{-1}$, and from there we get $A^n = S\Lambda^n S^{-1}$.

To diagonalize A we need the other eigenvalue and its eigenvector. As in the previous example, to find the other eigenvalue we solve $\lambda + 1 = .92 + .88$ to get $\lambda = .8$. An eigenvector for $\lambda = .8$ turns out to be $(-1, 1)$. Thus we can write

$$A^n = S\Lambda^n S^{-1} = \begin{pmatrix} 3 & -1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & .8^n \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -2 & 3 \end{pmatrix} \frac{1}{5} = \frac{1}{5} \begin{pmatrix} 3 + 2(.8^n) & 3 - 3(.8^n) \\ 2 - 2(.8^n) & 2 + 3(.8^n) \end{pmatrix}.$$

The last equality comes from multiplying the three matrices together. Notice that for fairly large values of n, $.8^n$ is very small. For example, $.8^{25} \approx .004$ and $.8^{50} \approx .000014$. So after a number of years we can ignore the $.8^n$ term and say

$$A \approx \begin{pmatrix} 3/5 & 3/5 \\ 2/5 & 2/5 \end{pmatrix}.$$

Notice that the columns are both the steady-state eigenvector of A that we found above. Moreover,

$$\begin{bmatrix} f_{n+1} \\ s_{n+1} \end{bmatrix} = A^n \begin{bmatrix} f_1 \\ s_1 \end{bmatrix} \approx \begin{pmatrix} 3/5 & 3/5 \\ 2/5 & 2/5 \end{pmatrix} \begin{bmatrix} f_1 \\ s_1 \end{bmatrix} = \begin{bmatrix} \frac{3}{5}f_1 + \frac{3}{5}s_1 \\ \frac{2}{5}f_1 + \frac{2}{5}s_1 \end{bmatrix} = (f_1 + s_1) \begin{bmatrix} 3/5 \\ 2/5 \end{bmatrix}.$$

So we see that in the long run (in this case after roughly 20-50 years) 3/5 (60%) of the deer population will be in the forest and 2/5 (40%) in the suburbs. In addition, the equation above says that the percentages will remain the same for all later years. This is why we call the vector $(3/5, 2/5)$ a *steady*-state.

Finally, notice that the percentage of deer in the forest versus the suburbs during year 1 had *absolutely no effect* on the outcome in the long run. Whether all the deer start out in the forest during year 1, or if it was 50/50, or whatever, makes no difference in the long run.

Positive Definite Matrices

A symmetric matrix with positive eigenvalues is called *positive definite*. The following gives a test for positive definiteness.

Let A be a symmetric matrix. If any one of the following is true, then the others are also true.

1. Every eigenvalue is positive.
2. Every upper left determinant is positive.
3. Every pivot is positive.
4. $\mathbf{x}^T A\mathbf{x}$ is positive except at $\mathbf{x} = \mathbf{0}$.

Remember, by positive we mean strictly greater than zero. Zero is *not* positive.

By "upper left determinant" we mean the determinant of the upper left part of the matrix. The four upper left determinants of this matrix are indicated on the right.

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 6 & 8 \\ 2 & 5 & 8 & 1 \\ 1 & 3 & 5 & 7 \end{pmatrix} \qquad 1, \quad \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = -2, \quad \begin{vmatrix} 1 & 2 & 3 \\ 3 & 4 & 6 \\ 2 & 5 & 8 \end{vmatrix} = -1, \quad \begin{vmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 6 & 8 \\ 2 & 5 & 8 & 1 \\ 1 & 3 & 5 & 7 \end{vmatrix} = -10$$

Example: Are all the eigenvalues of $\begin{pmatrix} 3 & 4 \\ 4 & 2 \end{pmatrix}$ positive?

Solution: The upper left determinants are 3 and $\begin{vmatrix} 3 & 4 \\ 4 & 2 \end{vmatrix} = -10$. They are not both positive, so neither are the eigenvalues.

Quadratic Forms

The term $\mathbf{x}^T A\mathbf{x}$ in (4) is called a *quadratic form*. Let's compute it for the 2×2 symmetric matrix $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$.

$$\mathbf{x}^T A\mathbf{x} = \begin{bmatrix} x & y \end{bmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax+by & bx+cy \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = ax^2 + 2bxy + cy^2.$$

Example: Using this formula, we can easily convert between a 2-dimensional quadratic form and its corresponding matrix. For example, the symmetric matrix $\begin{pmatrix} 3 & 2 \\ 2 & 7 \end{pmatrix}$ has quadratic form $3x^2 + 4xy + 7y^2$, and conversely, the quadratic form $5x^2 + 3xy + 2y^2$ corresponds to the matrix $\begin{pmatrix} 5 & 3/2 \\ 3/2 & 2 \end{pmatrix}$.

Completing the Square

You can check that we can write $ax^2 + 2bxy + cy^2 = a\left(x + \frac{b}{a}y\right)^2 + \left(\frac{ac-b^2}{a}\right)y^2$.

This is a lot to remember; however, notice that the coefficients of the square terms are actually *the pivots* of the matrix corresponding to this quadratic form. Breaking the quadratic form into these two square terms can be helpful sometimes.

Example: Are there any values of x and y that make $3x^2 + 12xy + 5y^2$ negative? If so, what are they?

Solution: The matrix corresponding to $3x^2 + 12xy + 5y^2$ is $\begin{pmatrix} 3 & 6 \\ 6 & 5 \end{pmatrix}$. Subtract 2 *row* 1 from *row* 2 to see what the pivots are. We get $\begin{pmatrix} \mathbf{3} & 6 \\ 0 & \mathbf{-7} \end{pmatrix}$. Now complete the square:

$$3x^2 + 12xy + 5y^2 = \mathbf{3}(x+2y)^2 + \mathbf{-7}y^2.$$

The negative pivot indicates there are values of x and y that make this negative. To find them, choose x and y to make the first term 0, say $x = -2$, $y = 1$. With this x and y the quadratic form $3x^2 + 12xy + 5y^2$ evaluates to -7.

Finding the Axes of a Tilted Ellipse

This is a nice example of how quadratic forms and orthogonal diagonalization can help with an algebra problem. The ellipse on the left is an ordinary ellipse. Its equation is given by $ax^2 + by^2 = 1$. Its major and semi-major axes are the x and y axis, respectively.

However, we could have a "tilted" ellipse, one whose axes lie on diagonal lines, instead of the x and y axes. Its equation will have an additional term, an xy term. We want to be able to find its major and semi-major axes.

Example: Find the axes of the ellipse $11x^2 - 6xy + 19y^2 = 1$.

Solution: Notice that $11x^2 - 6xy + 19y^2$ is the quadratic form associated to $A = \begin{pmatrix} 11 & -3 \\ -3 & 19 \end{pmatrix}$.

Now orthogonally diagonalize A. Calculating the eigenvalues and eigenvectors, we find the eigenvalues are 20 and 10 with corresponding eigenvectors $(-1, 3)$ and $(3, 1)$ and corresponding unit eigenvectors $(-1, 3)/\sqrt{10}$ and $(3, 1)/\sqrt{10}$. Thus we can orthogonally diagonalize A as

$$A = Q\Lambda Q^T = \frac{1}{\sqrt{10}} \begin{pmatrix} -1 & 3 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 20 & 0 \\ 0 & 10 \end{pmatrix} \begin{pmatrix} -1 & 3 \\ 3 & 1 \end{pmatrix} \frac{1}{\sqrt{10}}.$$

Now simplifying and substituting $Q\Lambda Q^T$ for A, we can write

$$11x^2 - 6xy + 19y^2 = \begin{bmatrix} x & y \end{bmatrix} A \begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{10} \begin{bmatrix} x & y \end{bmatrix} \begin{pmatrix} -1 & 3 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 20 & 0 \\ 0 & 10 \end{pmatrix} \begin{pmatrix} -1 & 3 \\ 3 & 1 \end{pmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Simplify this a bit by multiplying the first two parts together and multiplying the last two together:

$$\begin{bmatrix} x & y \end{bmatrix} \begin{pmatrix} -1 & 3 \\ 3 & 1 \end{pmatrix} = \begin{bmatrix} -x+3y & 3x+y \end{bmatrix}, \qquad \begin{pmatrix} -1 & 3 \\ 3 & 1 \end{pmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x+3y \\ 3x+y \end{bmatrix}.$$

By doing this, the big messy term ends up looking just like a quadratic form:

$$\frac{1}{10} \begin{bmatrix} -x+3y & 3x+y \end{bmatrix} \begin{pmatrix} 20 & 0 \\ 0 & 10 \end{pmatrix} \begin{bmatrix} -x+3y \\ 3x+y \end{bmatrix} = \frac{1}{10} \begin{bmatrix} X & Y \end{bmatrix} \begin{pmatrix} 20 & 0 \\ 0 & 10 \end{pmatrix} \begin{bmatrix} X \\ Y \end{bmatrix},$$

where $X = -x + 3y$ and $Y = 3x + y$. Use the formula $ax^2 + 2bxy + cy^2$ on the previous page to write this as

$$\frac{1}{10}(20X^2 + (2)(0)XY + 10Y^2)$$

Plugging back in for X and Y and simplifying, we see that we have rewritten the original ellipse equation as

$$2(-x + 3y)^2 + (3x + y)^2 = 1.$$

The axes of the ellipse are given by the setting the terms being squared equal to 0. So the equations of the axes are

$$\begin{matrix} -x + 3y = 0 \\ 3x + y = 0 \end{matrix} \quad \text{or} \quad \begin{matrix} y = x/3 \\ y = -3x \end{matrix}.$$

Similar Matrices

We say B is *similar* to A if there is an M so that $B = M^{-1}AM$. For example, $\begin{pmatrix} 4 & 2 \\ -1 & 1 \end{pmatrix}$ is similar to $\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ since $\begin{pmatrix} 4 & 2 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix}$. This example comes from diagonalization.

Fact: Similar matrices have the same eigenvalues.

Example: $\begin{pmatrix} 2 & 5 \\ 0 & 3 \end{pmatrix}$ is not similar to $\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$ because they have different eigenvalues.

Example: Find an M to show that $B = \begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix}$ is similar to $A = \begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix}$.

Solution: We want to find an M so that $B = M^{-1}AM$. Multiply both sides of this equation by M to get $MB = AM$. Now let M be a generic 2×2 matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and compute MB and AM:

$$MB = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} a & 3a+2b \\ c & 3c+2d \end{pmatrix}, \qquad AM = \begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} -a+6c & -b+6d \\ -a+4c & -b+4d \end{pmatrix}.$$

Since $MB = AM$, the corresponding entries in each must be equal, so we get the following system of equations:

$$\begin{array}{lcl} a & = & -a+6c \\ 3a+2b & = & -b+6d \\ c & = & -a+4c \\ 3c+2d & = & -b+4d \end{array} \quad \overset{\text{Simplify}}{\longrightarrow} \quad \begin{array}{l} a = 3c \\ a+b = 2d \\ a = 3c \\ 3c-b = 2d \end{array}.$$

This has an infinite number of solutions. Pick any value of c and d, say $c = d = 1$. Plugging into the equations, we get $a = 3$ and $b = -1$. Thus $M = \begin{pmatrix} 3 & -1 \\ 1 & 1 \end{pmatrix}$.

The only thing to be careful about when choosing c and d is to make sure that M is invertible. For instance $c = d = 0$ would be a bad choice. In general, it might not be this easy to solve the equations, but you can always move all the terms to the left hand side, write the system in matrix form and row reduce to find the solution. If there is no solution, or the only solution is all zeroes, then there is no M to be found, so the matrices are not similar. Note also that this same procedure (with a bigger M) works for larger matrices.

Singular Value Decomposition (SVD)

The singular value decomposition is a way to kind of diagonalize matrices of any shape. It uses the fact that A^TA is a square symmetric matrix which can be orthogonally diagonalized.

Let A be an $m \times n$ matrix. The SVD is $A = U\Sigma V^T$, where

Σ — This is an $m \times n$ matrix whose diagonal entries are the square roots of the eigenvalues of A^TA. (These are called the *singular values* of A. All the other entries are 0.

V — This is an $n \times n$ matrix whose columns are the unit eigenvectors of A^TA.* Its columns are orthonormal.

U — This is an $m \times m$ matrix. For each nonzero eigenvector $\mathbf{v}$ of A^TA, compute $A\mathbf{v}$ and find the corresponding unit vector. These are the columns of U.† Its columns are also orthonormal.

Order matters. Arrange the singular values in Σ from largest to smallest so that the largest is the first entry, the next largest is the second entry, etc. The vectors in V and U must line up with their corresponding eigenvalues in Σ just like in diagonalization.

Helpful Hints:

(1) If A has more columns than rows, then AA^T will be a smaller matrix than A^TA and thus its eigenvalues may be easier to find. In this case, any eigenvalue of AA^T is also an eigenvalue of A^TA, and the rest of the eigenvalues of A^TA are 0.

(2) When finding a unit vector in the direction of a given vector, factor out anything you can from the vector and then ignore the number you factored out. For example, to find a unit vector in the same direction as $(4, 8)$, factor out a 4 to get $4(1, 2)$, ignore the 4 and just find the unit vector in the direction of $(1, 2)$, which is $(1, 2)/\sqrt{5}$. Since $(1, 2)$ and $(4, 8)$ both point in the same direction, the unit vector in the direction of either will be the same.

Example: Find the SVD of $\begin{pmatrix} 2 & 0 \\ 3 & 2 \end{pmatrix}$.

Solution: First compute $A^TA = \begin{pmatrix} 13 & 6 \\ 6 & 4 \end{pmatrix}$. Its eigenvalues are 16 and 1 with corresponding eigenvectors $(2, 1)$ and $(-1, 2)$. The corresponding unit eigenvectors are $(2, 1)/\sqrt{5}$ and $(-1, 2)/\sqrt{5}$. Thus

$$\Sigma = \begin{pmatrix} \sqrt{16} & 0 \\ 0 & \sqrt{1} \end{pmatrix} \quad \text{and} \quad V = \begin{pmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix}.$$

Now multiply each eigenvector by A and find the corresponding unit vectors to find the columns of U:

$$\begin{pmatrix} 2 & 0 \\ 3 & 2 \end{pmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \end{bmatrix} \longrightarrow \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \qquad \begin{pmatrix} 2 & 0 \\ 3 & 2 \end{pmatrix} \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \end{bmatrix} \longrightarrow \frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 1 \end{bmatrix}.$$

Thus $A = U\Sigma V^T = \begin{pmatrix} 1/\sqrt{5} & -2/\sqrt{5} \\ 2/\sqrt{5} & 1/\sqrt{5} \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix}$.

---

*Footnote on first page applies here, too.

†If $n < m$, then there will not be enough vectors to fill up U. More work is needed as the columns U must be extended to an orthonormal basis of $\mathbb{R}^m$. However, we didn't consider this in class, so don't worry about it.

Example: Find the SVD of $\begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$.

Solution: First compute $A^TA = \begin{pmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 8 \end{pmatrix}$. Use helpful hint (1) to compute its eigenvalues.

$AA^T = \begin{pmatrix} 17 & 8 \\ 8 & 17 \end{pmatrix}$ has eigenvalues 25 and 9, thus A^TA has eigenvalues 25, 9, and 0.

The corresponding eigenvectors are $(1,1,0)$, $(1,-1,4)$, and $(-2,2,1)$ with corresponding unit vectors $(1,1,0)/\sqrt{2}$, $(1,-1,4)/\sqrt{18}$, and $(-2,2,1)/3$. Thus

$$\Sigma = \begin{pmatrix} \sqrt{25} & 0 & 0 \\ 0 & \sqrt{9} & 0 \end{pmatrix} \quad \text{and} \quad V = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{18} & -2/3 \\ 1/\sqrt{2} & -1/\sqrt{18} & 2/3 \\ 0 & 4/\sqrt{18} & 1/3 \end{pmatrix}.$$

Remember that Σ is the same shape as A, 3×2, so we can only fit the first two singular values in. Now multiply each eigenvector by A and find the corresponding unit vectors to find the columns of U:

$$\begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \end{bmatrix} \longrightarrow \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix} \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix} = \begin{bmatrix} 9 \\ -9 \end{bmatrix} \longrightarrow \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Notice that we did not use the third eigenvector, the one corresponding to the eigenvalue 0. This is because since A is 3×2, U is going to be 2×2 with vectors corresponding to the largest two eigenvalues.

So we have $A = U\Sigma V^T = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{18} & -1/\sqrt{18} & 4/\sqrt{18} \\ -2/3 & 2/3 & 1/3 \end{pmatrix}$.

Orthonormal Bases for the Four Fundamental Spaces

The SVD gives orthonormal bases for the fundamental spaces. Let A be a matrix with rank r (i.e., A has r nonzero pivots).

(1) The first r columns of V are an orthonormal basis for the row space of A.

(2) The remaining columns of V are an orthonormal basis for the nullspace of A.

(3) The first r columns of U are an orthonormal basis for the column space of A.

(4) The remaining columns of U are an orthonormal basis for the nullspace of A^T (called the left nullspace).

Example: In the second example above A has rank 2. Thus the orthonormal bases are the following:

Row space — $\left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{18}} \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix} \right\}$ Nullspace — $\left\{ \frac{1}{3} \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix} \right\}$

Column space — $\left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$ Left nullspace — None: only $\mathbf{0}$ is in the left nullspace.

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

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

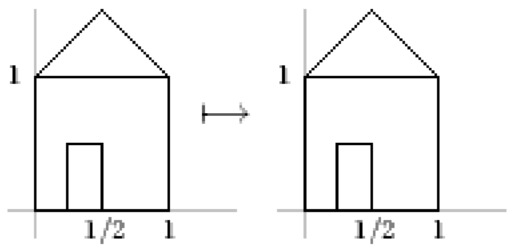
We can understand ordinary functions better by graphing them. We can't graph most linear transformations, but we can instead see how they transform geometric shapes.

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

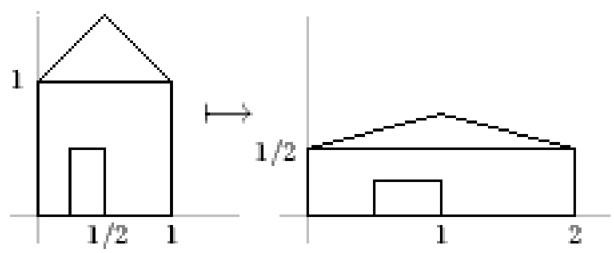
$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} 0 \\ 1 \end{bmatrix}$



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

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

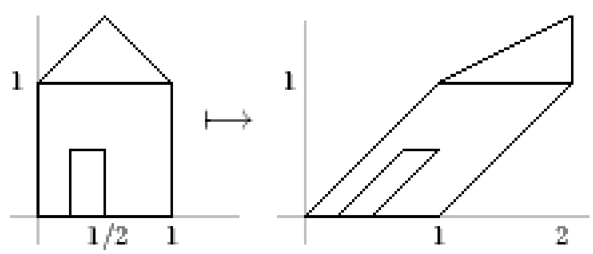
$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} 0 \\ .5 \end{bmatrix}$



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

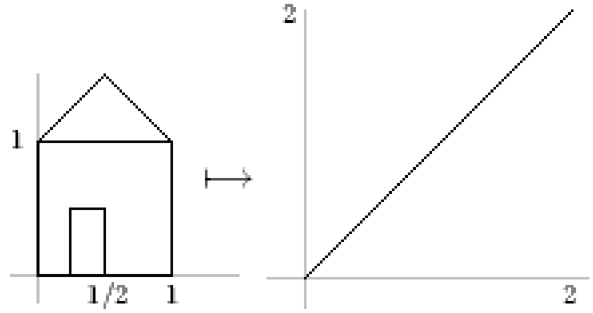
$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} 1 \\ 1 \end{bmatrix}$



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

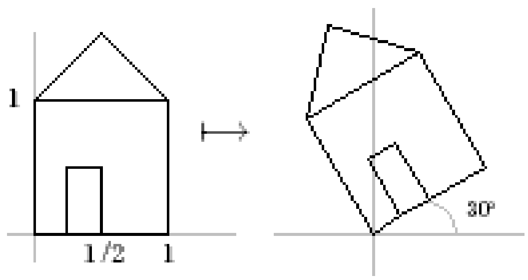
$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} 1 \\ 1 \end{bmatrix}$



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

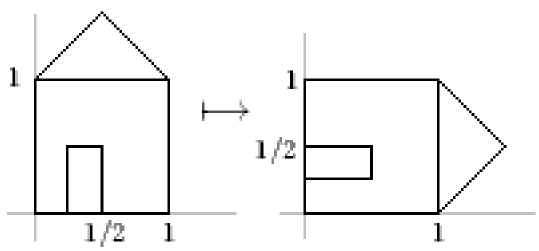
$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} \cos 30^\circ \\ \sin 30^\circ \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} -\sin 30^\circ \\ \cos 30^\circ \end{bmatrix}$



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

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

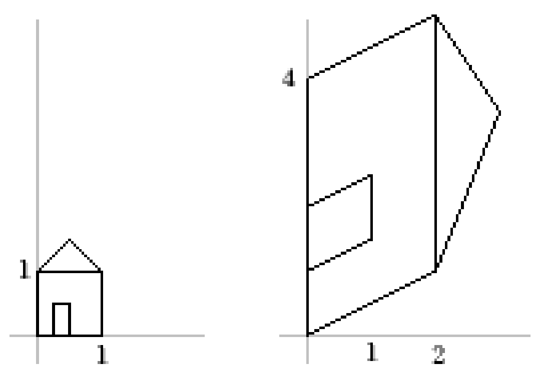
$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} 1 \\ 0 \end{bmatrix}$



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.

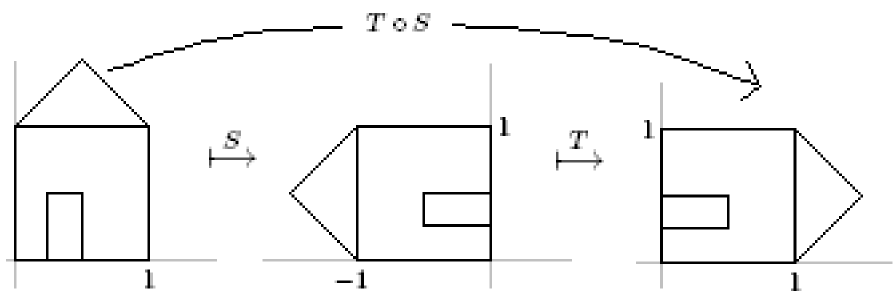
From these pictures we can see why these transformations are called linear: straight lines are transformed into straight lines or points. A nonlinear transformation might transform them into curves. Notice also that in all cases the transformed house touches the origin $(0, 0)$. This is because $T(\mathbf{0})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

Compositions

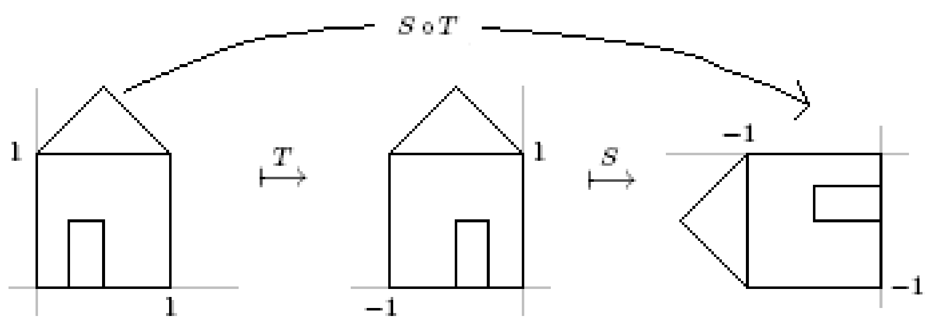
The *composition* of two linear transformations S and T, denoted $T \circ S$, is the linear transformation obtained by first doing S and then doing T. If the matrix of S is A and the matrix of T is B, then the linear transformation $T \circ S$ has matrix BA.

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

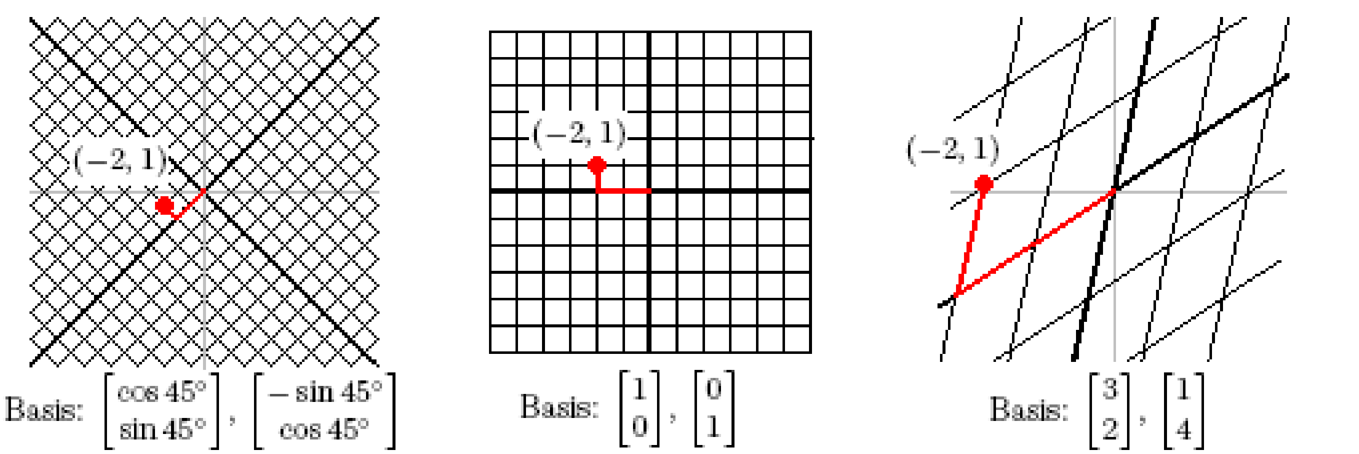
Write the output vector as $x \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + y \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. Thus the range consists of all linear combinations of the vectors $(1, 2, 1)$ and $(0, 0, 1)$.

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.



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.