Math 1180, Hastings

Notes, part 9

First I want to recall the following very important theorem, which only concerns square matrices. We will need to use parts of this frequently.

Theorem 1 *Suppose that A is an $n \times n$ matrix. Then either all of the following statements are true for A, or they are all false.*
(1) An echelon form of A has n pivots
(2) The rank of A is equal to n.
(3) The set of equations $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b}$ in R^n.
(5) $A\mathbf{x} = \mathbf{0}$ has only the zero solution.
(6) A is nonsingular
(7) $\det A \neq 0$
(8) The reduced row echelon form of A is the $n \times n$ identity matrix I.

Next I want to give more detail about another theorem stated earlier. The suggestion for proof given before (at the end of notes 8), does not work so easily. First we need a result about the matrix transpose:

Proposition 2 *If A is $m \times n$ and B is $n \times p$, so that AB is defined, then $(AB)^T = B^T A^T$.*

Proof. *We first check that this makes sense. AB is $m \times p$ so $(AB)^T$ is $p \times m$. Also, B^T is $p \times n$ and A^T is $n \times m$, so indeed, $B^T A^T$ is $p \times m$, which is compatible with the proposition, but does not prove it. To prove it we need to use summations and the definition of matrix multiplication.*

$$(AB)^T_{ij} = (AB)_{ji} = \sum_{k=1}^{n} a_{jk} b_{ki}.$$

On the other hand,

$$\left(B^T A^T\right)_{ij} = \sum_{k=1}^{n} \left(B^T\right)_{ik} \left(A^T\right)_{kj} = \sum_{k=1}^{n} b_{ki} a_{jk}.$$

Comparing these, we see that they are equal. ■

Theorem 3 *For any $n \times n$ matrix A,* $\det A = \det A^T$.

Proof. (This proof is from the text, but I've added details which they leave to an exercise.) Recall that for any square matrix A there is a permutation matrix P such that

$$PA = LU,$$

where L is lower triangular, with 1's on the diagonal, and U is upper triangular. From the proposition,

$$A^T P^T = U^T L^T,$$

and from the product rule for determinants,

$$\det P \det A = \det L \det U$$
$$\det A^T \det P^T = \det U^T \det L^T.$$

Since L and U are triangular, their determinants are the products of their diagonal elements, and L^T and U^T are also triangular, with the same diagonal elements. So to prove $\det A^T = \det A$, we need to show that $\det P^T = \det P$.

P is a permutation matrix, meaning that it is obtained from I by a sequence of row exchanges. First let's look at an example:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Notice that each time we exchange two rows, we can get the same result by exchanging two columns, though not necessarily the same two columns. For example in the second step we exchanged row 2 and row 3. The same result is obtaining by exchanging column 1 and column 2. Let's see if we can prove this in general.

A permutation matrix is of the form

$$P = (p_{ij})$$

where each row and each column has exactly one 1 and the rest 0's. Now suppose that before the exchange, $p_{ir} = 1$ and $p_{jq} = 1$. This means that row i has a 1 in column r, and row j has a 1 in column q.

Now we exchange row i and row j. If the new matrix is $\hat{P}$, then

$$p_{ir} = 1, p_{jq} = 1$$
$$\hat{p}_{jr} = 1, \hat{p}_{iq} = 1.$$

But suppose instead that we had exchanged column r and column q, to get $\tilde{P}$. Then

$$\begin{aligned} p_{ir} &= 1, p_{jq} = 1 \\ \tilde{p}_{iq} &= 1, \tilde{p}_{jr} = 1. \end{aligned}$$

Hence, $\hat{P} = \tilde{P}$. So exchanging two rows has exactly the same effect on the determinant (a sign change) as exchanging two columns. It follows that $\det P^T = \det P$, and so $\det A^T = \det A$. ∎

0.0.1 Eigenvalues and Eigenvectors

We now move on to Chapter 6 which has a peculiar title: "Eigenvalues, eigenvectors and canonical forms". However, some parts of that chapter require the preceding chapters, which we will come to later. So for now, concentrate on these notes.

It must be emphasized that this entire chapter is only about square matrices. Start with the same definition as in the text:

Definition 8.2. *Let A be an $n \times n$ matrix. A* ***non-zero*** *vector $\mathbf{v}$ in R^n is called an "eigenvector" if $A\mathbf{v}$ is a scalar multiple of $\mathbf{v}$. The corresponding scalar multiple, often denoted by λ, is called an "eigenvalue". Hence $\mathbf{v}$ and λ satisfy the equation*

$$A\mathbf{v} = \lambda\mathbf{v} \tag{1}$$

Note: An eigenvector must be a nonzero vector.

Example:

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \ \mathbf{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Then

$$A\mathbf{v} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} = 3\mathbf{v} \tag{2}$$

so $\lambda = 3$.

The same matrix has another eigenvector, which is $\mathbf{w} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. To see this we compute $A\mathbf{w}$:

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

so $\lambda = -1$.

Generally, in applications, the eigenvalue is more important than the eigenvector, though you can't have one without the other. It is important to emphasize once again that an eigenvector cannot be the zero vector, but the eigenvalue could be zero.

Example:

$$A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \mathbf{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \ \lambda = 0.$$

This is easy to check and I leave it to you.

Note that any singular matrix has an eigenvalue 0. Can you see why? Look at the theorem at the beginning of these notes.

Generally, though not always, a 2×2 matrix has two eigenvalues and two linearly independent eigenvectors. For the example just above, the second eigenvalue is $\lambda = 2$ and the second eigenvector is $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

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

But some matrices may have only one eigenvalue. The most obvious example is the zero matrix, for which every nonzero vector is an eigenvector with eigenvalue $\lambda = 0$.

The practical thing one has to learn is: How do we find eigenvalues and eigenvectors? To answer this we start with the key equation (1) and write it differently:

$$A\mathbf{v} - \lambda \mathbf{v} = \mathbf{0}$$
$$(A - \lambda I)\,\mathbf{v} = \mathbf{0}.$$

Since $\mathbf{v} \neq \mathbf{0}$, this is possible if and only if the matrix $A - \lambda I$ is singular. This is a restriction on λ.

Now recall from the Theorem at the beginning of this set of notes saying (among other things) that a matrix is singular if and only if its determinant is zero. We apply this to the matrix $A - \lambda I$ to see that λ is an eigenvalue if and only if $\det(A - \lambda I) = 0$.

Look at the example from above,

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

We find $A - \lambda I$:

$$A - \lambda I = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{pmatrix}.$$

We then take the determinant of this 2×2 matrix to get

$$\begin{aligned} \det(A - \lambda I) &= (1-\lambda)^2 - 4 \\ &= 1 - 2\lambda + \lambda^2 - 4 \\ &= \lambda^2 - 2\lambda - 3 = (\lambda - 3)(\lambda + 1). \end{aligned}$$

Setting this equal to zero gives the two values of λ used earlier: $\lambda_1 = 3$, $\lambda_2 = -1$.

Next we find eigenvectors. We choose one of the eigenvalues, say $\lambda = 3$. We form the matrix

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

We wish to find a nonzero vector $\mathbf{v}$ such that

$$(A - 3I)\,\mathbf{v} = \mathbf{0}.$$

We can do this using Gaussian elimination:

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

which is now in echelon form. There is one pivot, in the first column, and the second variable is a free variable.

We wish to solve

$$\begin{pmatrix} -2 & 2 \\ 0 & 0 \end{pmatrix} \mathbf{v} = \mathbf{0}$$

or

$$\begin{pmatrix} -2 & 2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{3}$$

or

$$
\begin{aligned}
-2v_1 + 2v_2 &= 0 \\
v_1 &= v_2.
\end{aligned}
$$

Choosing some nonzero value for the free variable $v_{2,}$ we then solve for v_1. The simplest choice is $v_2 = 1$, whereupon we derive that

$$\mathbf{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

is an eigenvector. So is any nonzero multiple of this vector.

Next we choose the second eigenvalue, $\lambda = -1$. This time

$$A - \lambda I = A + I = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 2 \\ 0 & 0 \end{pmatrix}.$$

There are one pivot and one free variable. We find that one eigenvector is

$$\begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

and any other one is a nonzero constant times this vector.

Now let's do a 3×3 example, namely

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

Therefore

$$A - \lambda I = \begin{pmatrix} 2-\lambda & 0 & 1 \\ 0 & 2-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{pmatrix}.$$

Since $A - \lambda I$ is upper triangular, its determinant is the product of its diagonal elements. Hence the characteristic polynomial of A is

$$\det(A - \lambda I) = (2-\lambda)^3$$

and the eigenvalues of A are the roots of $(2-\lambda)^3$, or $\lambda_1 = \lambda_2 = \lambda_3 = 2$.

Definition 4 *The "characteristic polynomial" of* A *is the polynomial* $p(\lambda) = \det(A - \lambda I)$.

Definition 5 *If* $p(\lambda)$ *has a term* $(\lambda - c)^m$, *then* c *is called an eigenvalue with "algebraic multiplicity"* m.

Hence, in the example above, 2 is an eigenvalue with algebraic multiplicity 3.

We now look solutions of $(A - 2I)\mathbf{v} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \mathbf{v} = \mathbf{0}$. This is in echelon form, with one pivot, in the third column, so v_1 and v_2 are free variables. Setting $v_1 = 1, v_2 = 0$ and vice-versa gives the two eigenvectors

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Any nonzero multiple of one of these is also an eigenvector. Their sum is also, as well as any nonzero linear combination of these two vectors.

Notice that in this example A is upper triangular and the eigenvalues are the numbers on the diagonal. This is always the case: *The eigenvalues of a triangular matrix are the numbers on the diagonal of the matrix.*

Here is another example:

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

From what we just said, the eigenvalues are $\lambda_1 = 2, \lambda_2 = 2, \lambda_3 = 2$.

Then $A - 2I = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. There is only one free variable, and one solution of $(A - 2I)\mathbf{v} = \mathbf{0}$ is $\mathbf{v} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. This is then the only linearly independent eigenvector. (i.e. every other eigenvector is a multiple of this one.) We say that $\lambda = 2$ has "geometric multiplicity " one.

In the previous example, $\begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$, $\lambda = 2$ had geometric multiplicity two, because there were two eigenvectors, neither one a multiple of the other. (We will generalize this idea later. We can't give a good general definition of geometric multiplicity now. However, it turns out that the geometric multiplicity of an eigenvalue λ is the number of free variables (or parameters) in the solution of $(A - \lambda I)\,\mathbf{v} = \mathbf{0}$. This is $n - r$ where r is the rank of $A - \lambda I$.)

Warning: Do not use Gaussian elimination on A in order to find eigenvalues. Example showing why not:

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 2 \\ 0 & -3 \end{pmatrix} \to \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I. \tag{4}$$

We saw that A has eigenvalues 3 and -1. But B has eigenvalues 1 and 1. Thus Gaussian elimination changed the eigenvalues. (You can use Gaussian elimination on the matrix $A - \lambda I$, though this rarely is needed in homework problems.)

Example with complex eigenvalues:
Let $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.. Then

$$\det (A - \lambda I) = \det \begin{pmatrix} -\lambda & 1 \\ -1 & -\lambda \end{pmatrix} = \lambda^2 + 1.$$

There are no real numbers which are eigenvalues. The only eigenvalues are complex, indeed, purely imaginary. They are $\lambda_1 = i$ and $\lambda_2 = -i$.

The case of complex eigenvalues is very important for math 1270.

Another example is

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}. \tag{5}$$

Then

$$\det (A - \lambda I) = \det \begin{pmatrix} -\lambda & -1 \\ 1 & 1 - \lambda \end{pmatrix} = \lambda^2 - \lambda + 1.$$

The roots are

$$\lambda_1 = \frac{-1 + \sqrt{-3}}{2} = -\frac{1}{2} + \frac{\sqrt{3}}{2} i, \;\; \lambda_2 = -\frac{1}{2} - \frac{\sqrt{3}}{2} i$$

Hence the eigenvalues can have both real and imaginary parts. ($\mathrm{Re}\lambda_1 = \mathrm{Re}\,\lambda_2 = -\frac{1}{2}$, $\mathrm{Im}\,\lambda_1 = \frac{\sqrt{3}}{2}$, $\mathrm{Im}\,\lambda_2 = -\frac{\sqrt{3}}{2}$ in this example.)

If the eigenvalues are complex, then the eigenvectors will be complex also. The text book gives the eigenvectors in some cases with complex eigenvalues.

Homework (due Wednesday, October 7): pg. 547, # 2,6,10,12,14 . Hint: Each of these is a matrix, and you are supposed to find the eigenvalues and eigenvectors. You are also supposed to find the geometric and algebraic multiplicities of the eigenvalues.

One of these matrices has complex eigenvalues. The real and imaginary parts of these complex eigenvalues are small integers. **You do not have to find the eigenvectors for the one with complex eigenvalues, as the eigenvectors are usually not very useful.**

The other four matrices have only real eigenvalues, all of which are also small integers. If you get any fractions or square roots in your eigenvalues, you have made a mistake.

In finding the roots of cubics, since you know they are small integers, you can try substituting some guesses until you find one. Then use long division to get it down to a quadratic. (Divide the cubic by $\lambda - \lambda_1$, where λ_1 is the root you found by guessing.)

Your eigenvectors might contain fractions, but you can multiply them by constants to remove all fractions from your answers.