These are brief notes for the lecture on Friday October 29, 2010: they are not complete, but they are a guide to what I want to say today. They are not guaranteed to be correct.

0.1. Eigenvectors and Eigenvalues

Recall the observation from last class: eigenvectors with the same eigenvalue (together with the zero vector) form a subspace. In some sense, they behave similarly to each other. To continue the analogy, eigenvectors having different eigenvalues must be somehow very different from each other.

For example, if $\underline{v}_1, \underline{v}_2$ are eigenvectors of A with eigenvalues λ_1, λ_2 respectively, and $\lambda_1 \neq \lambda_2$, then $\underline{v}_1, \underline{v}_2$ cannot be scalar multiples of each other. Indeed if

$$\underline{v}_1 = c\underline{v}_2$$

then

$$A\underline{v}_1 = Ac\underline{v}_2 \implies \lambda_1\underline{v}_1 = c\lambda_2\underline{v}_2$$

so $c\lambda_1 = c\lambda_2$, so $c = 0$. But then $\underline{v}_1 = \underline{0}$, so it wasn't an eigenvector in the first place!

The following (often very useful) theorem is a typical rephrasing of that idea:

THEOREM 1. *If $\underline{v}_1, \ldots, \underline{v}_r$ are eigenvectors corresponding to distinct eigenvalues $\lambda_1, \ldots, \lambda_r$ of an $n \times n$ matrix A, then the set $\{\underline{v}_1, \ldots, \underline{v}_r\}$*

Proof: Suppose that the vectors $\underline{v}_1, \ldots, \underline{v}_r$ are eigenvectors with distinct eigenvalues $\lambda_1, \ldots, \lambda_r$, and that they are linearly dependent. Choose the largest p so that $\underline{v}_1, \ldots, \underline{v}_p$ are linearly independent: since $\underline{v}_1$ is an eigenvector, it is non-zero, and so p exists. Then $\underline{v}_{p+1}$ is a linear combination of $\underline{v}_1, \ldots, \underline{v}_p$,

$$\underline{v}_{p+1} = c_1\underline{v}_1 + \cdots + c_p\underline{v}_p.$$

Now multiply both sides of this equation by $A - \lambda_{p+1}I$ to obtain

$$(\lambda_{p+1} - \lambda_{p+1})\underline{v}_{p+1} = (\lambda_1 - \lambda_{p+1})c_1\underline{v}_1 + \cdots + (\lambda_p - \lambda_{p+1})c_p\underline{v}_p.$$

Since the vectors $\underline{v}_1, \ldots, \underline{v}_p$ are linearly independent, and none of $\lambda_1, \ldots, \lambda_p$ are equal to λ_{p+1}, this implies that each of the c_j's are zero. But this contradicts our assumption that $\underline{v}_{p+1}$ is both non-zero and a linear combination of $\underline{v}_1, \ldots, \underline{v}_p$.

Hence $\underline{v}_1, \ldots, \underline{v}_r$ are linearly independent, as claimed.

An application to linear recurrences: Suppose that we have a constant coefficient homogeneous linear recurrence: for example:

$$f_{n+1} = 3f_n - 2f_{n-1}.$$

If we create a vector

$$\underline{x}_n = \begin{pmatrix} f_n \\ f_{n-1} \end{pmatrix}$$

then the recurrence can be written as

$$\begin{pmatrix} f_{n+1} \\ f_n \end{pmatrix} = \begin{pmatrix} 3f_n - 2f_{n-1} \\ f_n \end{pmatrix}$$

or

$$\begin{pmatrix} f_{n+1} \\ f_n \end{pmatrix} = \begin{pmatrix} 3 & -2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} f_n \\ f_{n-1} \end{pmatrix}$$

that is, as

$$\underline{x}_{n+1} = A\underline{x}_n.$$

where $A = \begin{pmatrix} 3 & -2 \\ 1 & 0 \end{pmatrix}$. Iterating this equation, we obtain $x_n = A^{n-1}x_1$.

Now, it happens that this matrix has nice eigenvalues and eigenvectors: indeed, $\lambda_1 = 1$ is an eigenvalue, with eigenvector $\underline{u} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\lambda_2 = 2$ is an eigenvalue, with eigenvector $\underline{v} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$. Hence $A^{n-1}\underline{u} = \lambda_1^{n-1}\underline{u} = \underline{u}$, and $A^{n-1}\underline{v} = \lambda_2^{n-1}\underline{v} = 2^{n-1}\underline{v}$. Thus, if $\underline{x}_1 = c\underline{u} + d\underline{v}$, we have

$$\underline{x}_{n-1} = A^{n-1}(c\underline{u} + d\underline{v}) = c\underline{u} + d2^{n-1}\underline{v}.$$

So now, if we know f_0 and f_1, since $\underline{u}$ and $\underline{v}$ span $\mathbb{R}^2$, we can find c, d so that $\underline{x}_1 = c\underline{u} + d\underline{v}$, and hence determine a formula for f_n.

These sorts of techniques are very useful for solving much larger difference equations.

0.2. The Characteristic Equation

As was asked last lecture, "It's easy to check that a vector is an eigenvector, and to check that a real number is an eigenvalue, row reduce $A - \lambda I$ and check that it has rows without pivots, and this way we can find eigenvectors. But how do we find the eigenvalues to use?"

In this section we'll show that eigenvalues are the roots of a certain polynomial.

Suppose we want to find the eigenvalues of a 2×2 matrix A: we want to find λ so that there is a non-zero solution $\underline{x}$ to $(A - \lambda I)\underline{x} = \underline{0}$. Such solutions exist precisely when the determinant

$$\det(A - \lambda I) = 0$$

(see Theorem 4 in section 2.2).

So, for example, if we consider the matrix $A = \begin{pmatrix} 3 & -2 \\ 1 & 0 \end{pmatrix}$ from the example at the end of the last section,

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

Hence λ is an eigenvalue of A if and only if $\lambda^2 - 3\lambda + 2 = 0$: since $\lambda^2 - 3\lambda + 2 = (\lambda - 1)(\lambda - 2)$, the only eigenvalues are $1, 2$ as we claimed earlier.

Notice that we have taken an equation involving a real unknown λ and a vector unknown $\underline{x}$, and isolated just λ. Once we have found the values of λ which make the determinant zero, we can solve for the eigenvectors $\underline{x}$ which work.

For a general 2×2 matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, the formula for the determinant enables us to compute the eigenvalues easily:

$$\det \begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix} = (a-\lambda)(d-\lambda) - bc = \lambda^2 - (a+d)\lambda + (ad-bc)$$

so that the quadratic formula gives

$$\lambda = \frac{(a+d) \pm \sqrt{(a+d)^2 - 4(ad-bc)}}{2}.$$

This is *absolutely* **not** a formula to memorize! It's far easier in practice to work out for a given matrix what the eigenvalues are by doing it from scratch.

Note: if the equation $\det(A-\lambda I) = 0$ has no real roots, then there are no real eigenvalues (or eigenvectors). However, if we work over the complex numbers $\mathbb{C}$ we can still do something. This is beyond the scope of our discussion here, but is of great importance. For example, if

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

corresponding to a rotation through an angle θ, and if the angle is not a half or a full turn (i.e. π or 2π) then

$$\det(A - \lambda I) = \lambda^2 - 2\lambda\cos\theta + 1$$

so $\lambda = \cos\theta \pm \sqrt{cos^2\theta - 1} = \cos\theta + i\sin\theta$, where $i = \sqrt{-1}$, so unless $\sin\theta = 0$, there are no real eigenvalues.

For a 3×3 matrix A we can do the same analysis, and using the formula for the determinant of A, show that the eigenvalues of A must satisfy a cubic polynomial. Here we see a formula for the polynomial which is much worse than for the 2×2 case:

$$-\lambda^3 + (a+e+i)\lambda^2 + (bd - ai + cg - ei - ae + fh)lambda + (aei + bfh + cdh - afh - bdi - ceg)$$

Again, this is not a formula to memorize! Rather, it is of theoretic importance to know that it is a cubic polynomial, and hence there are at most three real eigenvalues.

The big theorem is an extension of the Invertible Matrix theorem:

Theorem (The Invertible Matrix Theorem continued). *Let A be an $n \times n$ matrix. Then A is invertible if and only if*

s *The number 0 is* not *an eigenvalue of A.*

t *The determinant of A is not zero.*

We'll also need to recall the following facts about determinants.

Theorem 4 (Properties of Determinants). *Let A and B be $n \times n$ matrices.*

a *A is invertible if and only if* $\det(A) \neq 0$

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

c $\det(A^T) = \det(A)$.

d *If A is triangular, then* $\det(A)$ *is the product of the entries on the main diagonal.*

e *A row replacement operation on A doesn't change the determinant. A row interchange switches the sign of the determinant. A row scaling also scales the determinant by the same scale factor.*

Then we see the fundamental fact that

THEOREM. *A scalar λ is an eigenvalue of an $n \times n$ matrix if and only if λ satisfies the characteristic equation*

$$\det(A - \lambda I) = 0.$$

It is true, but beyond the scope of this course, that if A is an $n \times n$ matrix then the characteristic polynomial $\det(A - \lambda I)$ has degree n. Hence an $n \times n$ matrix has at most n eigenvalues.

DEFINITION. *Two $n \times n$ matrices A and B are similar if there is an invertible matrix P so that $P^{-1}AP = B$, or equivalently $A = PBP^{-1}$.*

Note that if P is invertible, then so is $Q = P^{-1}$, so if $A = PBP^{-1}$ then $B = QAQ^{-1}$, so that A is similar to B if and only if B is similar to A.