These are brief notes for the lecture on Friday November 6, 2009: they are not complete, but they are a guide to what I want to say today. They are not guaranteed to be correct.

5.3. Eigenvectors and Eigenvalues

5.4. 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. There are lots of ways to prove this: for example, using co-factor expansions of determinants, using the fact that the determinant of a matrix can be expressed as a sum of products in which exactly one term from each column appears, or using clever algebraic methods which avoid determinants entirely.

Some matrices are such that all of their eigenvalues are distinct: however, sometimes a matrix can have repeated eigenvalues: for example, the $n \times n$ matrix I has n eigenvalues, all of which are 1, since $\det(I - \lambda I) = (1 - \lambda)^n$.

If A has characteristic equation $p(\lambda) = \det(A - \lambda I)$ and if λ_i is a multiple root of this polynomial, then we define the *multiplicity* of the eigenvalue λ_i to be the number of times $\lambda - \lambda_i$ is a root of the polynomial. (Note: sometimes there are other notions of multiplicity, and we'll need to call this the "algebraic" multiplicity of λ_i.)

Example: Find the eigenvalues and their multiplicities of

$$A = \begin{pmatrix} 6 & 5 & 0 & -5 \\ 0 & -3 & 1 & 2 \\ 0 & 0 & 6 & 3 \\ 0 & 0 & 0 & 7 \end{pmatrix}.$$

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.

THEOREM 5. *If $n \times n$ matrices A and B are similar, then they have the same characteristic polynomials.*

Proof: If $A = PBP^{-1}$ then

$$\begin{aligned} \det(A - \lambda I) &= \det(PBP^{-1} - \lambda I) = \det(PBP^{-1} - \lambda PIP^{-1}) \\ &= \det(P(B - \lambda I)(P^{-1}) = \det(P)\det(B - \lambda I)\det(P^{-1}) \\ &= \det(P)\det(P^{-1})\det(B - \lambda I) = \det(B - \lambda I) \end{aligned}$$

as claimed.

5.5. Diagonalization

Definition. *A square matrix D having ij^{th} element d_{ij} is* diagonal *if $d_{ij} = 0$ except for the elements d_{jj} on the diagonal: that is,*

$$\text{if } i \neq j \text{ then } d_{ij} = 0.$$

Often, if the eigenvalues and eigenvectors of a $n \times n$ matrix A are nice enough, we can use them to obtain a nice factorization of the form $A = PDP^{-1}$. Here we want D to be a nice matrix, for example, diagonal, and P to be invertible, so that its columns form a basis for $\mathbb{R}^n$.

A big part of the reason this is a useful thing is that we can compute powers of diagonal matrices very quickly: for example, if

$$D = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

then

$$D^2 = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 5^2 & 0 & 0 \\ 0 & 2^2 & 0 \\ 0 & 0 & 3^2 \end{pmatrix}$$

$$D^3 = DD^2 \begin{pmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 5^2 & 0 & 0 \\ 0 & 2^2 & 0 \\ 0 & 0 & 3^2 \end{pmatrix} = \begin{pmatrix} 5^3 & 0 & 0 \\ 0 & 2^3 & 0 \\ 0 & 0 & 3^3 \end{pmatrix}$$

and in general,

$$D^k = \begin{pmatrix} 5^k & 0 & 0 \\ 0 & 2^k & 0 \\ 0 & 0 & 3^k \end{pmatrix} \quad \text{for } k \geq 1.$$

If $A = PDP^{-1}$, where D is diagonal, then we can compute A^k easily in terms of D^k:

$$A^2 = (PDP^{-1})(PDP^{-1}) = (PD)(P^{-1}P)(DP^{-1}) = (PD)I(DP^{-1}) = PD^2P^{-1},$$
$$A^3 = (PDP^{-1})(PD^2P^{-1}) = (PD)(P^{-1}P)(D^2P^{-1}) = (PD)I(D^2P^{-1}) = PD^3P^{-1}$$

and in general

$$A^k = (PDP^{-1})(PD^{k-1}P^{-1}) = (PD)(P^{-1}P)(D^{k-1}P^{-1}) = (PD)I(D^{k-1}P^{-1}) = PD^kP^{-1}.$$

Without all the derivation crowding this out, it's clearer:

$$A = PDP^{-1},$$
$$A^2 = PD^2P^{-1},$$
$$A^3 = PD^3P^{-1}$$

and in general, if $k \geq 1$,

$$A^k = PD^kP^{-1}.$$

So, if we can compute D^k easily (which we can, for a diagonal matrix D) and if $A = PDP^{-1}$, then we can compute A^k easily too.