These are brief notes for the lecture on Wednesday November 11, 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. Diagonalization

Definition. *An $n \times n$ matrix A is said to be diagonalizable if there exists a an invertible matrix P and a diagonal matrix D (both of which must be $n \times n$) so that $A = PDP^{-1}$ (or equivalently, since P is invertible, $AP = PD$): that is, A is similar to a diagonal matrix.*

Some matrices turn out not to be diagonalizable: we'll see shortly that, for example,

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

is not diagonalizable.

Eigenvalues and eigenvectors turn out to be exactly what we need to study diagonalization of matrices. In fact, we can characterize exactly when a square matrix can be diagonalized.

Theorem 6 (The Diagonalization Theorem). *An $n \times n$ matrix A is diagonalizable if and only if it has n linearly independent eigenvectors. In fact, $A = PDP^{-1}$, with D a diagonal matrix, if and only if the columns of P are n linearly independent eigenvectors of A. In this case the diagonal entries of D are the eigenvalues of A, with the jj^{th} entry of D being the eigenvalue corresponding to the eigenvector which is the j^{th} column of A.*

So, this means that we can diagonalize an $n \times n$ matrix A if and only if we can find a basis for $\mathbb{R}^n$ consisting of eigenvectors of A. Before we see why this is, let's revisit the matrix $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ above.

The characteristic equation for A is $\lambda^2 = 0$, so the only eigenvalue is 0. Now, the null space of A is 1 dimensional, so we can only find one linearly independent eigenvector. Hence by the theorem, A cannot be diagonalized!

Proof of the theorem: If P is invertible, and $A = PDP^{-1}$, then $AP = PD$. If

$$P = [\underline{v}_1 \ \underline{v}_2 \ \dots \ \underline{v}_n]$$

and D has diagonal entries $\lambda_1, \dots \lambda_n$, then PD has columns

$$PD = [\lambda_1 \underline{v}_1 \ \lambda_2 \underline{v}_2 \ \dots \ \lambda_n \underline{v}_n].$$

Hence $A\underline{v}_j = \lambda_j \underline{v}_j$, and $\underline{v}_j$ is an eigenvector of A. Since P is invertible, its columns are linearly independent, and hence we have n linearly independent eigenvectors of A.

Conversely, if we have n linearly independent eigenvectors, $\underline{v}_1, \dots \underline{v}_n$ (having corresponding eigenvalues $\lambda_1, \dots \lambda_n$ respectively), we construct a matrix P having them as columns. Then

$$AP = [\lambda_1 \underline{v}_1 \ \lambda_2 \underline{v}_2 \ \dots \ \lambda_n \underline{v}_n] = PD$$

Since the n vectors in $\mathbb{R}^n$ are linearly independent, P is invertible and hence

$$A = PDP^{-1}$$

and A is diagonalizable. □

COROLLARY. *An $n \times n$ matrix A having n distinct eigenvalues is diagonalizable.*

Proof: We saw earlier that eigenvectors corresponding to distinct eigenvalues are linearly independent. Each eigenvalue has at least one corresponding eigenvector: hence we have n linearly independent eigenvectors, and A is diagonalizable.

How to find eigenvalues numerically

The characteristic equation is nice theoretical tool, but it is not the best way to find eigenvalues numerically, especially if the matrix has floating point entries.

Here are two methods which give more success.

To simplify things, we will assume that A is $n \times n$ and is diagonalizable, so that it has n real eigenvalues, and that they all have distinct sizes,

$$|\lambda_1| > |\lambda_2 > \cdots > |\lambda_n|$$

with eigenvectors $\underline{v}_j$ corresponding to λ_j.

The power method will enable us to find λ_1 and a corresponding eigenvector. Suppose we have a vector $\underline{x} \in \mathbb{R}^n$, so that

$$\underline{x} = c_1\underline{v}_1 + \cdots + c_n\underline{v}_n$$

Then if we multiply $\underline{x}$ by A k times, we obtain

$$\begin{aligned} A^k\underline{x} &= c_1 A^k \underline{v}_1 + \cdots + c_n A^k \underline{v}_n \\ &= c_1\lambda_1^k\underline{v}_1 + \cdots + c_n\lambda_n^k\underline{v}_n \end{aligned}$$

If we divide both sides by λ_1^k, we get

$$\frac{1}{(\lambda_1)^k}A^k\underline{x} = c_1\underline{v}_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\underline{v}_2 + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\underline{v}_n.$$

The coefficients $c_j\left(\frac{\lambda_j}{\lambda_1}\right)^k$ for $j \geq 2$ tend to zero as $k \to \infty$. This means, so long as $c_1 \neq 0$, that as k gets large, $A^k\underline{x}$ points in nearly the same direction as $\underline{v}_1$.

This gives rise to
The Power Method

(1) Select an initial vector $\underline{x}_0$ whose largest entry is 1.

(2) For $k = 0, 1, \ldots$

  (a) Compute $A\underline{x}_k$

  (b) Let μ_k be an entry in $A\underline{x}_k$ whose absolute value is as large as possible.

  (c) Compute $\underline{x}_{k+1} = (1/\mu_k)A\underline{x}_k$.

(3) For almost all choices of $\underline{x}_0$, the sequence $\{\mu_k\}$ approaches the dominant eigenvalue λ_1, and the sequence $\{\underline{x}_k\}$ approaches a corresponding eigenvector.

This method works fastest when λ_1 is much bigger than the other eigenvalues. We can exploit this, if we have a good guess at an eigenvalue using the inverse power method.

Suppose that A has an eigenvalue λ_j, which we don't know, but we have an estimate $\alpha \neq \lambda_j$ which is closer to λ than to any other eigenvalue. We will consider the matrix $B = (A - \alpha I)^{-1}$: it is not hard to show that B has the same eigenvectors as A, and that its eigenvalues are

$$\frac{1}{\lambda_1 - \alpha}, \frac{1}{\lambda_2 - \alpha}, \ldots, \frac{1}{\lambda_n - \alpha}.$$

So, if α is closer to λ_j than to any other eigenvalue of A, the eigenvalue $\frac{1}{\lambda_j - \alpha}$ is dominant.

We exploit this as follows:

The Inverse Power Method

(1) Select an initial estimate α close enough to λ.

(2) Select an initial vector $\underline{x}_0$ whose largest entry is 1.

(3) For $k = 0, 1, \ldots$

  (a) Solve $(A - \alpha I)\underline{y}_k = \underline{x}_k$ for $\underline{y}_k$.

  (b) Let μ_k be an entry in $\underline{y}_k$ whose absolute value is as large as possible.

  (c) Compute $\nu_k = \alpha + (1/\mu_k)$.

  (d) Compute $\underline{x}_{k+1} = (1/\mu_k)\underline{y}_k$.

(4) For almost all choices of $\underline{x}_0$, the sequence $\{\nu_k\}$ approaches the eigenvalue λ of A, and the sequence $\{\underline{x}_k\}$ approaches a corresponding eigenvector.