Math Jams Worksheet 4:
Jordan Canonical Form and Cayley-Hamilton

April 14th, 2008

Jordan Canonical Form

Definition 0.1. If A is an $n \times n$-matrix of the form

$$A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & A_k \end{pmatrix},$$

where A_1,... A_k are square matrices on the diagonal of A,we say A is *block-diagonal* with blocks $A_1, \dots, A_k$ on the diagonal. We also say A is the *direct* sum of the matrices $A_1, \dots A_k$, and we write $A = A_1 \oplus \cdots \oplus A_k$.

Note in particular that if $A = A_1 \oplus \cdots \oplus A_k$ and $B = B_1 \oplus \cdots \oplus B_k$ and A_i and B_i are square matrices of the same size for each i, then $AB = A_1B_1 \oplus \cdots \oplus A_kB_k$ (this is easier to see if you write this out in matrix form).

Let $\lambda \in \mathbb{C}$ be an arbitrary complex number. The $k \times k$ matrix

$$J = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & \ddots & & \\ & & \ddots & 1 & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}$$

Is called the *Jordan block of size k with eigenvalue λ.*

A matrix A is said to be in *Jordan canonical form* if it is a block diagonal matrix with Jordan blocks along the diagonal. If $T : V \to V$ is a linear transformation (where V is some finite-dimensional vector space over any field), then a *Jordan canonical form* for T is a matrix representing T which is in Jordan canonical form.

Theorem 0.2 (Jordan Canonical Form)**.** *If V is an n-dimensional complex vector space and $T : V \to V$ is a linear transformation, then there exists a basis β of V so that $[T]_\beta$, the matrix of T with respect to the basis β, is in Jordan canonical form. Moreover, up to permutation of the Jordan blocks, the Jordan canonical form of T is* unique*.*

We can actually phrase this theorem slightly more strongly: the reason we use the field $\mathbb{C}$ here is that $\mathbb{C}$ is algebraically closed, so the characteristic polynomial of T factors into linear factors over $\mathbb{C}$; we say here that $\mathbb{C}$ *contains all the eigenvalues of T*. The statement of JCF actually applies to a finite-dimensional vector space V over any field $\mathbb{F}$ and a linear transformation $T : V \to V$ such that $\mathbb{F}$ contains all the eigenvalues of T.

This theorem is extremely powerful; let's do some exercises to show off just how cool it is.

JCF Exercises

For all of these exercises, we assume our base field is $\mathbb{C}$.

1. Suppose A and B are $n \times n$-matrices with entries in $\mathbb{C}$. Show that A is similar to B (that is, there exists an invertible $n \times n$-matrix P such that $P^{-1}AP = B$) if and only if A and B have the same Jordan canonical form.
2. As a special case of the previous exercise, show that two diagonal matrices are similar if and only if their diagonal entries are the same up to a permutation.
3. Prove that if $\lambda_1, \ldots, \lambda_n$ are the (not necessarily distinct) eigenvalues of the $n \times n$-matrix A, then $\lambda_1^k, \ldots, \lambda_n^k$ are the eigenvalues of A^k for any $k \geq 0$.
4. Show that any matrix A is similar to its transpose, A^T (Hint: reduce to showing this in the case that A is a Jordan block).

The Minimal Polynomial and Cayley-Hamilton

The theory of the JCF is intricately tied to something called the minimal polynomial. To really get into the nitty-gritty of this theory, we need a bit (not much, mind you!) of abstract algebra, so I'll leave off the proofs and just give the facts.

Lemma 0.3 (The Minimal Polynomial)**.** *Let A be an $n \times n$-matrix over the field $\mathbb{F}$. Then there exists a unique monic polynomial $m_A(x)$ of least degree with coefficients in $\mathbb{F}$ such that $m_A(A) = 0$. The polynomial $m_A(x)$ is called the* minimal polynomial *for A. If $f(x)$ is another polynomial with coefficients in $\mathbb{F}$ such that $f(A) = 0$, then $m_A(x)$ divides $f(x)$.*

Theorem 0.4 (The Cayley-Hamilton Theorem)**.** *Let A be an $n \times n$-matrix over the field $\mathbb{F}$, let $p_A(x)$ be the characteristic polynomial of A, and let $m_A(x)$ be the minimal polynomial of A. Then $m_A(x)$ divides $p_A(x)$. Put another way, $p_A(A) = 0$.*

First, a few exercises to show how this is linked to JCF:

Exercises: Minimal Polynomials and JCF

1. Let J be a $k \times k$-Jordan block with eigenvalue λ. Show that the minimal polynomial of J is $(x - \lambda)^k$.
2. Show that if A is similar to B, then A and B have the same minimal polynomial.
3. Show that if A is an $n \times n$ matrix and $A = B \oplus C$, then the minimal polynomial of A is the least common multiple of the minimal polynomials for B and C.
4. Let A be an $n \times n$-matrix such that $A = J_1 \oplus J_2 \oplus \cdots \oplus J_k$ where each J_i is an $k_i \times k_i$-Jordan block with eigenvalue λ. Show that the minimal polynomial of A is $(x - \lambda)^{\max_i k_i}$.
5. Now, let A be a matrix in Jordan canonical form; say $A = J_1 \oplus J_2 \oplus \cdots \oplus J_k$, where each J_i is a Jordan block. Describe the minimal polynomial of A as completely as you can.
6. Find the minimal polynomials of the matrices

$$A_1 = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{and} \qquad A_2 = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

What we discussed in these exercises shows a number of great facts linking the minimal polynomial and JCF. In particular, we see that if A is an $n \times n$-matrix over $\mathbb{C}$ and λ is an eigenvalue of A, then $(x - \lambda)$ divides $p_A(x)$, and the largest power of $(x - \lambda)^k$ which divides $m_A(x)$ is the size of the largest Jordan block with eigenvalue λ appearing in the JCF of A. We can use this to prove the following great fact:

Theorem 0.5. *If A is an $n \times n$-matrix over $\mathbb{C}$, then A is diagonalizable if and only if $m_A(x)$ has no repeated roots.*

Proof. Suppose that A is diagonalizable, so A is similar to a diagonal matrix D. Since similar matrices have the same minimal polynomial, $m_A(x) = m_D(x)$. Let $a_1, \ldots, a_k$ be the distinct elements appearing on the diagonal of D. Then the minimal polynomial of D is $m_D(x) = (x - a_1) \cdots (x - a_k)$, which has no repeated roots.

On the other hand, if $m_A(x)$ has no repeated roots, then by the above discussion, every Jordan block of the JCF of A has size 1, i.e. the JCF of A is a diagonal matrix, so A is diagonalizable. $\square$

Again, the above theorem could be better stated: suppose A is an $n \times n$ matrix over the field $\mathbb{F}$ containing all the eigenvalues of A. Then A is diagonalizable if and only if $m_A(x)$ has no repeated roots. This immediately gives us the following corollary, which we had proved another way:

Corollary 0.6. *If A is an $n \times n$-matrix with n distinct eigenvalues, then A is diagonalizable.*

Proof. The characteristic polynomial of A has n distinct roots, and $m_A(x)$ divides $p_A(x)$, so also $m_A(x)$ has no repeated roots. $\square$

Now, we have a serious powerhouse behind us now. Let's use it to prove some cool results which were previously pretty tricky or impossible:

More Exercises

1. Determine all possible Jordan canonical forms for a linear transformation with characteristic polynomial $(x - 2)^3(x - 3)^2$. Find the corresponding minimal polynomial for each JCF.
2. Let A be an $n \times n$ matrix such that $A^k = I$ for some positive integer k. Prove that A is diagonalizable.
3. Prove that if $A^2 = A$ then A is similar to a diagonal matrix which has only 0's and 1's along the diagonal.
4. Prove that an $n \times n$-matrix over $\mathbb{C}$ with $A^3 = A$ can be diagonalized. Harder question: is this true over *any* field $\mathbb{F}$?
5. Prove that if A is an invertible $n \times n$-matrix such that some power of A is diagonalizable, then A is diagonalizable.