MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Electrical Engineering and Computer Science

6.241: Dynamic Systems—Fall 2003

Recitation 5

Jordan Decomposition

In this recitation, the decomposition of the $n \times n$ matrix A into its Jordan form is reviewed. The matrix exponential, e^{tA} is then generalized to the case where A is not diagonalizable.

Motivation

The Matrix Exponential. As we have seen in lecture, the matrix exponential, e^{tA}, where A is $n \times n$, is a state transition matrix for LTI systems with a state space representation. The matrix exponential was defined as follows:

$$e^{tA} = I + \sum_{k=1}^{\infty} \frac{t^k A^k}{k!} = I + At + \frac{t^2 A^2}{2!} + \cdots + \frac{t^k A^k}{k!} + \cdots$$

When the $n \times n$ matrix A is diagonalizable[1], e^{tA} becomes particularly easy to evaluate. Recall that if A is diagonalizable, then there exists a (nonsingular) similarity transformation, V, such that $A = V \Lambda V^{-1}$ where Λ is a diagonal matrix, $\Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Thus, $A^k = (V \Lambda V^{-1})^k = V \Lambda^k V^{-1}$, where $\Lambda^k = diag(\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k)$. So, substituting into the series expansion of e^{tA},

$$e^{tA} = V(I + t\Lambda + \frac{t^2 \Lambda^2}{2!} + \cdots + \frac{t^k \Lambda^k}{k!} + \cdots)V^{-1}$$

Now using the definition of the function e^x as power series:

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!},$$

we have that

$$e^{At} = V \begin{pmatrix} e^{\lambda_1 t} & & & \\ & e^{\lambda_2 t} & & \\ & & \ddots & \\ & & & e^{\lambda_n t} \end{pmatrix} V^{-1}.$$

Diagonalizable Matrices. In recitation 2, we said A is diagonalizable, if and only if there exists a (nonsingular) similarity transformation, V, such that $A = V \Lambda V^{-1}$ where Λ is a diagonal matrix, $\Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_n)$. A necessary condition for A to be diagonalizable is that A has n linearly

[1] See to recitation 2.

independent eigenvectors. A sufficient condition for A to be diagonalizable is that A has n distinct eigenvalues, because when A has n distinct eigenvalues, then the corresponding eigenvectors are linearly independent. One can restate the conditions for the diagonalizability of A in terms of the so-called algebraic multiplicity and geometric multiplicity of A's eigenvalues.

Recall that the eigenvalues of the $n \times n$ matrix A are the roots of its characteristic polynomial, $\chi(\lambda) = \det(A - \lambda I) = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2} \ldots (\lambda - \lambda_k)^{m_k} = 0$, where $\sum m_i = n$. The number of times an eignevalue λ_i is repeated, m_i, is known as the *algebraic multiplicity* of λ_i, and is denoted by AM(λ_i)= m_i. The number of linearly independent eigenvectors that are associated with the eigenvalue λ_i is known as the *geometric multiplicity* of λ_i, and is denoted by GM(λ_i). Verify that GM(λ_i)=dim$\mathcal{N}(A - \lambda_i I)$, i.e., the dimension of the nullspace of $A - \lambda_i I$. Also, verify that in general GM(λ_i)$\leq$AM(λ_i).

Now, we can state the condition on the diagonalizability of A as follows: A is diagonalizable if and only if GM(λ_i)=AM(λ_i) for every distinct eigenvalue, λ_i, of A. In otherwords, for every eigenvalue that is repeated m_i times, we can find m_i linearly independent eignevectors associated with that eigenvalue.

Example 1: (From [1] page 129.) Consider the matrix

$$A = \begin{bmatrix} 1 & 3 & -2 \\ 0 & 4 & -2 \\ 0 & 3 & -1 \end{bmatrix}.$$

From the characteristic polynomial of this matrix, $(1 - \lambda)^2(2 - \lambda)$, the eigenvalues of A are $\lambda_1 = 1$ and $\lambda_2 = 2$, where AM$(\lambda_1) = 2$. The eigenvalues associated with λ_1 are the vectors in the nullspace of

$$A - \lambda_1 I = \begin{bmatrix} 0 & 3 & -2 \\ 0 & 3 & -2 \\ 0 & 3 & -2 \end{bmatrix}.$$

Note that there are two such linearly independent eigenvectors, for example, $v_1 = [\ 1 \quad 2 \quad 3\ ]'$ and $v_2 = [\ 1 \quad 0 \quad 0\ ]'$. Thus, GM$(\lambda_1) = 2$. Finally, $v_3 = [\ 1 \quad 1 \quad 1\ ]'$ is the eigenvector associated with λ_2. Let $V = [\ v_1 \quad v_2 \quad v_3\ ]$ and $\Lambda = diag(1, 1, 2)$, and verify that $A = V \Lambda V^{-1}$.

Jordan Form

When it is not possible to find n linearly independent eigenvectors of A, the matrix A cannot be diagonalized. However, there exists a similarity transformation, M that transforms A to a matrix J, $A = MJM^{-1}$, such that J is as close as possible to a diagonal form. J is the Jordan form matrix, and each A is similar to only one such J (except for a reordering of the blocks):

$$J = \begin{pmatrix} J_1 & & & \\ & J_2 & 0 & \\ & 0 & \ddots & \\ & & & J_r \end{pmatrix}.$$

J is thus block diagonal and J_i are square matrices with the following structure:

$$J_i = \begin{pmatrix} \lambda_j & 1 & 0 & \cdots & 0 \\ 0 & \lambda_j & 1 & 0 & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots 0 & \lambda_j & 1 \\ 0 & \cdots & \cdots & 0 & \lambda_j \end{pmatrix}.$$

J_i is called a *Jordan block*, in this case J_i is Jordan block associated with λ_j. The number of Jordan blocks associated with an eigenvalue λ_j is equal to GM(λ_j). The sum of the dimensions of all Jordan blocks associated with λ_j is equal to AM(λ_j).

Example 2: From the structure of the Jordan form of A one can obtain information about the algebraic and geometric multiplicity of the eigenvalues of A. Consider

$$J = \left(\begin{array}{ccc|cc|cc|c} \lambda_1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ & \lambda_1 & 1 & 0 & 0 & 0 & 0 & 0 \\ & & \lambda_1 & 0 & 0 & 0 & 0 & 0 \\ \hline & & & \lambda_1 & 1 & 0 & 0 & 0 \\ & & & & \lambda_1 & 0 & 0 & 0 \\ \hline & & 0 & & & \lambda_2 & 1 & 0 \\ & & & & & & \lambda_2 & 0 \\ \hline & & & & & & & \lambda_3 \end{array}\right).$$

We can immediately say that: AM(λ_1) = 5 and GM(λ_1) = 2 (two Jordan block for λ_1), AM(λ_2) = 2, and GM(λ_2) = 1, and AM(λ_3) = GM(λ_3) = 1. The fact that any matrix can be put in its Jordan form by an opportune similarity transformation, makes the Jordan form very useful in linear system theory. Unfortunately, the computation of the Jordan form is very sensitive to computational errors.

Constructing the Jordan Form of A. When A does not have n linearly independent eigenvectors to form the columns of the matrix M (and a basis for $\mathbb{R}^n$), one can add linearly independent vectors to the eigenvectors in order to complete the basis. In the following discussion, we explain which vectors to add so that the associated similarity transformation of A gives J.

Generalized Eigenvectors. First, consider an eigenvalue with GM(λ_i)<AM(λ_i). If GM(λ_i)= p_i, AM(λ_i)= m_i then we need to find $m_i - p_i$ linearly independent vectors to associate with this eignevalue. These vectors are generated from the eigenvectors and are called *generalized eigenvectors* of A.

Suppose λ is an eignevalue of A and x_1 is a corresponding eigenvector. $k-1$ generalized eigenvectors,

$\{x_2, \ldots x_k\}$ can be generated as follows:

$$\begin{aligned} &Ax_1 = \lambda x_1 \\ &Ax_2 = \lambda x_2 + x_1 \\ &\vdots \\ &Ax_k = \lambda x_k + x_{k-1}. \end{aligned}$$

$\{x_2, \ldots x_k\}$ form a string of vectors that is said to be headed by x_1. $\{x_1, x_2, \ldots x_k\}$ correspond to a single Jordan block; in particular, notice that the above equations give rise to the structure of a Jordan block of dimension k relative to an eigenvalue λ:

$$A[x_1, \ldots, x_k] = [x_1, \ldots, x_k] \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & 0 & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots 0 & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix}.$$

An equivalent definition of generalized eigenvectors that you may see in some references is the following: a vector $x \neq 0$ is said to be a generalized eigenvector of order k of A relative to λ if and only if $(A - \lambda I)^k x = 0$ and $(A - \lambda I)^{k-1} x \neq 0$. Note that if $k = 1$, the above definition reduces to $(A - \lambda I)x = 0$ and $x \neq 0$, which is the definition of an eigenvector. Starting from a generalized eigenvector of order k of A relative to λ, denoted by x_k, one can generate a the Jordan chain of generalized eigenvectors as follows:

$$\begin{aligned} &x_k \\ &x_{k-1} = (A - \lambda I)x_k \\ &\vdots \\ &x_i = (A - \lambda I)^{k-i} x_k \\ &\vdots \\ &x_1 = (A - \lambda I)^{k-1} x_k. \end{aligned}$$

Note that x_1 is an eigenvector of A, since $(A - \lambda I)x_1 = (A - \lambda I)(A - \lambda I)^{k-1} x_k = 0$. Note also that any x_i, $i = 1, \ldots, k$, is a generalized eigenvector of order i of A. This is true since: $(A - \lambda I)^i x_i = (A - \lambda I)^i (A - \lambda I)^{k-i} x_k = 0$ and $(A - \lambda I)^{i-1} x_i = (A - \lambda I)^{k-1} x_k \neq 0$. Another immediate consequence of the definition is that all the generalized eigenvectors in the Jordan chain of length k belong to $\mathcal{N}(A - \lambda I)^k$.

Finally, the the chain of vectors headed by x_1 are linearly independent, as we require. This is proved in the following theorem.

Theorem 0.1 *Let x_k be a generalized eigenvector of A of order k relative to λ, and let $x_1, \ldots x_k$ be the chain of generalized eigenvectors generated by x_k. Then $x_1, \ldots x_k$ are linearly independent vectors.*

Proof. We will prove that if $\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_k x_k = 0$ then $\alpha_1, \ldots, \alpha_k = 0$. First notice that $(A - \lambda I)^{k-1} x_i = 0$ for $1 \leq i \leq k-1$. This follows immediately from the fact that x_{k-1} is a generalized eigenvector of A of order $k-1$. Form the definition of x_k it also follows that $(A - \lambda I)^{k-1} x_k \neq 0$. Now, if $y = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_k x_k = 0$, then $(A - \lambda I)^{k-1} y = 0$ but this can only be true if

$\alpha_k = 0$. If we multiply y by $(A - \lambda I)^{k-2}$ then it follows that $\alpha_{k-1} = 0$. Continuing the procedure we obtain that $\alpha_1, \ldots, \alpha_k = 0$.

The following method for the construction of the Jordan form is due to Filippov, see [2].

Step 1. Suppose the dimension of the column space of A is $r < n$. There must be r independent vectors, x_i, in the column space which are either eigenvectors or generalized eigenvectors, i.e., $Ax_i = \lambda_i x_i$ or $Ax_i = \lambda_i x_i + x_{i-1}$.

Step 2. Suppose further that the nullspace and column space of A have an intersection of dimension p. Each vector, x_i, in the nullspace of A is an eigenvector corresponding to $\lambda = 0$, that is, $Ax_i = 0$. Now if x_i is also in the column space of A, then $x_i = Ay_i$ for some y_i.

Step 3. Finally, since the dimension of the nullspace of A is $n - r$, and p of the vectors are in both the null space and the column space, then there must be $n - r - p$ vectors, z_i, that are in the nullspace but not in the column space.

The vectors found in step 1, 2 and 3, namely x_i, y_i, and z_i are linearly independent (this is shown in [2] pare 455). They form the columns of M, and $J = M^{-1}AM$ is the Jordan form. The order of these vectors in M is important. Each string of generalized eigenvectors follows the eigenvector that heads that string, and each y_i follows the x_i from which it came, and completes a string which has $\lambda_i = 0$. The z_i come at the end, each in a string by itself.

Example 3:(From [2] page 456.) We demonstrate Jordan decomposition in the following example. Suppose

$$A = \begin{bmatrix} 8 & 0 & 0 & 8 & 8 \\ 0 & 0 & 0 & 8 & 8 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 8 \end{bmatrix}.$$

A has eigenvalues 8 (repeated twice) and 0 (repeated three times). The rank of A, $r = 3$. Note that the column space of A is spanned by the vectors $\{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \end{bmatrix}', \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \end{bmatrix}',$ $\begin{bmatrix} 0 & 0 & 0 & 0 & 1 \end{bmatrix}'\}$, and the nullspace of A is spanned by $\{\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \end{bmatrix}', \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \end{bmatrix}'\}$. Now, we look for the eigenvectors of A. For $\lambda = 8$, we have

$$A - 8I = \begin{bmatrix} 0 & 0 & 0 & 8 & 8 \\ 0 & -8 & 0 & 8 & 8 \\ 0 & 0 & -8 & 0 & 0 \\ 0 & 0 & 0 & -8 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Notice that $x_1 = \begin{bmatrix} 8 & 0 & 0 & 0 & 0 \end{bmatrix}'$ is an eigenvector associated with $\lambda = 8$. There are no more eigenvectors associated with $\lambda = 8$, so, since we need another vector to associate with $\lambda = 8$ (because $\lambda = 8$ has algebraic multiplicity 2), we seek a generalized eigenvector, i.e., x_2 such that $Ax_2 = 8x_2 + x_1$ i.e. $(A - 8I)x_2 = x_1$. Note that $x_2 = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \end{bmatrix}'$ does the job. So, $\{x_1, x_2\}$ are a jordan string for $\lambda = 8$. Now, we turn to $\lambda = 0$. Notice that $x_3 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \end{bmatrix}'$ is in the nullspace of A, hence is an eigenvector that is associated with $\lambda = 0$. Now, x_3 is also in the column space of A, so we must find y such that $Ay = x_3$. Note that $y = \begin{bmatrix} -1 & 0 & 0 & 1 & 0 \end{bmatrix}'$ does the job.So, $\{x_3, y\}$ are a jordan string for $\lambda = 0$. Finally, we complete the basis with z from the nullspace of A (but not in the column space of A). Take $z = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \end{bmatrix}'$. So, $M = \begin{bmatrix} x_1 & x_2 & x_3 & y & z \end{bmatrix}$, and

$$J = \begin{bmatrix} 8 & 1 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Example 4: Let

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

$\chi(\lambda) = \lambda^4$, therefore A has 4 repeated eigenvalues $\lambda = 0$, i.e., $AM(\lambda) = 4$. It is also easy to see that $GM(\lambda) = 2$. Therefore we expect a Jordan form for A with 2 Jordan blocks.

Two linearly independent eigenvectors corresponding to $\lambda = 0$, $\{w_1, w_2\}$ (which are also in $\mathcal{N}(A)$), are given by: $w_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}'$ and $w_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix}'$. Now notice that w_1 and w_2 are also in the column space of A, so we must find the vectors y_1 and y_2 such that $Ay_1 = w_1$ and $Ay_2 = w_2$, where y_1 and y_2 together with w_1 and w_2 form a linearly independent set. Verify that the following vectors work: $y_1 = \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix}'$ and $y_2 = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}'$.

Therefore the matrix $M = [w_1, y_1, w_2, y_2]$ of the similarity transformation and the relative Jordan form are

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad J = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The Matrix Exponential. The existence of the Jordan form allows us to generalize the results on power of a matrix and matrix exponential that we derived for diagonalizable matrices. Note that a diagonal matrix is a special case of a Jordan form.

$$A^k = (MJM^{-1})^k = MJ^kM^{-1}$$

with

$$J^k = \begin{pmatrix} J_1^k & & & \\ & J_2^k & & \\ & & \ddots & \\ & & & J_r^k \end{pmatrix},$$

where

$$J_i^k = \begin{pmatrix} \lambda_j^k & k\lambda_j^{k-1} & k^2\lambda_j^{k-2} & \cdots & k^{s-1}\lambda_j^{(k-s-1)} \\ 0 & \lambda_j^k & k\lambda_j^{k-1} & k^2\lambda_j^{k-2} & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \\ \vdots & \cdots & 0 & \lambda_j^k & \lambda_j^{k-1} \\ 0 & \cdots & \cdots & 0 & \lambda_j^k \end{pmatrix}.$$

J_i is $s \times s$. This result can be used to obtain a general expression for the matrix exponential.

$$e^{tA} = \sum_{k=0}^{\infty} \frac{t^k A^k}{k!} = Me^{tJ}M^{-1}$$

where

$$e^{tJ} = \begin{pmatrix} e^{tJ_1} & & & \\ & e^{tJ_2} & & \\ & & \ddots & \\ & & & e^{tJ_r} \end{pmatrix},$$

and

$$e^{tJ_i} = \begin{pmatrix} e^{\lambda_j t} & te^{\lambda_j t} & \frac{t^2}{2!}e^{\lambda_j t} & \cdots & \frac{t^{s-1}}{(s-1)!}e^{\lambda_j t} \\ 0 & e^{\lambda_j t} & te^{\lambda_j t} & \frac{t^2}{2!}e^{\lambda_j t} & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \\ \vdots & \cdots & 0 & e^{\lambda_j^t} & te^{\lambda_j t} \\ 0 & \cdots & \cdots & 0 & e^{\lambda_j t} \end{pmatrix}.$$

Example 5: Suppose

$$J = \begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}.$$

Then

$$e^{tJ} = \begin{pmatrix} e^{\lambda_1 t} & te^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_1 t} & 0 \\ 0 & 0 & e^{\lambda_2 t} \end{pmatrix}.$$

Let $J_1 = \begin{bmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{bmatrix}$ and $J_2 = [\lambda_2]$. Now,

$$e^{tJ} = \sum_{k=0}^{\infty} \frac{t^k J^k}{k!} = \sum_{k=0}^{\infty} \frac{t^k}{k!} \begin{bmatrix} J_1^k & 0 \\ 0 & J_2^k \end{bmatrix} = \begin{bmatrix} \sum_{k=0}^{\infty} \frac{t^k J_1^k}{k!} & 0 \\ 0 & \sum_{k=0}^{\infty} \frac{t^k J_2^k}{k!} \end{bmatrix}.$$

Now, $e^{\lambda_2 t} = \sum_{k=0}^{\infty} \frac{t^k J_2^k}{k!}$. Furthermore, $J_1^0 = I$, $J_1^1 = \begin{bmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{bmatrix}$, $J_1^2 = \begin{bmatrix} \lambda_1^2 & 2\lambda_1 \\ 0 & \lambda_1^2 \end{bmatrix}$, $J_1^3 = \begin{bmatrix} \lambda_1^2 & 3\lambda_1^2 \\ 0 & \lambda_1^2 \end{bmatrix}$, and so on. So, in $\sum_{k=0}^{\infty} \frac{t^k J_1^k}{k!}$, the diagonal terms sum to $e^{\lambda_1 t}$, and the off diagonals generate the sum $0 + \frac{t}{1!} + \frac{2\lambda_1 t^2}{2!} + \frac{3\lambda_1 t^3}{3!} + \ldots = te^{\lambda_1 t}$.

References

[1] Antsakalis, P.J., and Michel, A.N. “Linear Systems,” *McGraw Hill*: 1998.

[2] Strang, G. “Linear Algebra and its Applications,” *Saunders HBJ*: 1988.