Jordan normal form notes

(version date: 11/21/07)

If A has an eigenbasis $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$, i.e. a basis made up of eigenvectors, so that $A\mathbf{u}_j = \lambda_j \mathbf{u}_j$, then A is diagonal with respect to that basis. To see this, let $U = (\mathbf{u}_1 \cdots \mathbf{u}_n)$, i.e. let U denote the matrix with j-th column $\mathbf{u}_j$, $j = 1, \ldots, n$. Then

$$AU = (A\mathbf{u}_1 \ \cdots \ A\mathbf{u}_n) = (\lambda_1 \mathbf{u}_1 \ \cdots \ \lambda_n \mathbf{u}_n) = U \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$$

implies that $U^{-1}AU = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$.

What happens if A *doesn't* have an eigenbasis?

Important example: Let $\mathbf{e}_j$ denote the j-th Euclidean basis vector, with 1 in the j-th postion and 0 elsewhere. For example,

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad \text{and} \qquad \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Define

$$N_n := (0 \ \mathbf{e}_1 \ \cdots \ \mathbf{e}_{n-1}) = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ & & & \ddots & & \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

The characteristic polynomial of N_n is $\chi_{N_n}(\lambda) = \lambda^n$; hence zero is the only eigenvalue of N_n.

$$N_n \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 0 + x_2 \mathbf{e}_1 + \cdots x_n \mathbf{e}_{n-1} = \begin{pmatrix} x_2 \\ \vdots \\ x_n \\ 0 \end{pmatrix}$$

implies that $\ker N_n = \operatorname{span}\{\mathbf{e}_1\}$.

For any $n \times n$ matrix $B = (b_1 \cdots b_n)$,

$$B\,N_n = (B\ 0\ B\ \mathbf{e}_1 \ \cdots \ B\ \mathbf{e}_{n-1}) = (0\ b_1 \ \cdots \ b_{n-1});$$

the columns of B are shifted right, with a zero column being introduced on the left. In particular,

$$\begin{aligned} N_n^2 &= (0\ 0\ \mathbf{e}_1 \ \cdots \ \mathbf{e}_{n-2}) \\ N_n^3 &= (0\ 0\ 0\ \mathbf{e}_1 \ \cdots \ \mathbf{e}_{n-3}) \\ &\vdots \\ N_n^n &= 0. \end{aligned}$$

Hence $\ker N_n^j = \operatorname{span}\{\mathbf{e}_1, \ldots, \mathbf{e}_j\}$. If we let $S_{n,k}$ denote the $n \times n$ matrix with 1 on the k super-diagonal and 0 elsewhere, i.e. with ij-th entry equal to 1 if $j = i + k$, 0 otherwise, then $N_n^k = S_{n,k}$.

The exponential of N_n can be computed using the power series for exp, since $S_{n,j} = \mathbb{O}$ if $j \geq n$ implies that

$$\exp(t\,N_n) = \sum_{j=0}^{\infty} \frac{t^j}{j!} N_n^j = \sum_{j=0}^{n-1} \frac{t^j}{j!} S_{n,j}. \tag{1}$$

□

We can use this example as a guide in handling arbitrary repeated eigenvalues with "insufficient eigenspace", i.e. eigenvalues λ whose algebraic multiplicity (the power to which $(x - \lambda)$ appears in the characteristic polynomial is greater than the geometric multiplicity (the dimension of the eigenspace). Define $s : \mathbb{R}^{n\times n} \times \mathbb{R} \to \mathbb{R}^{n\times n}$ by

$$s(A, \lambda) := A - \lambda\,\mathbb{I}.$$

If λ is an eigenvalue of A, then the eigenspace of λ is the kernel of $s(A, \lambda)$:

$$\ker s(A, \lambda) = \{x \in \mathbb{R}^n : s(A, \lambda)x = 0\} = \{x \in \mathbb{R}^n : A\,x = \lambda\,x\}.$$

If λ is an eigenvalue of A with (algebraic) multiplicity ℓ, then the generalized eigenspace of λ is the kernel of $s(A, \lambda)^\ell$.

Example: $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$.
The characteristic polynomial of A is $\chi_A(\lambda) = (\lambda - 2)^2(\lambda - 3)$. $s(A, 2) = (0\ \mathbf{e}_1\ \mathbf{e}_3)$, with $\ker s(A, \lambda) = \operatorname{span}\{\mathbf{e}_1\}$. $s(A, 2)^2 = (0\ 0\ \mathbf{e}_3)$, with $\ker s(A, \lambda)^2 = \operatorname{span}\{\mathbf{e}_1, \mathbf{e}_2\}$. □

Consider the case in which λ has algebraic multiplicity k, but geometric multiplicity 1 (the eigenspace of λ is one dimensional). Let $\mathbf{u}_1$ be an eigenvector of A with eigenvalue λ. If there are $\mathbf{u}_2, \ldots, \mathbf{u}_k$ such that

$$s(A, \lambda)\mathbf{u}_{j+1} = \mathbf{u}_j \qquad \text{(equivalently } A\mathbf{u}_{j+1} = \mathbf{u}_j + \lambda\,\mathbf{u}_{j+1}\text{)} \qquad j = 1, \ldots, k-1,$$

then

$$\begin{aligned} A(\mathbf{u}_1\ \mathbf{u}_2\ \cdots\ \mathbf{u}_k) &= (A\,\mathbf{u}_1\ A\,\mathbf{u}_2\ \cdots\ A\,\mathbf{u}_k) \\ &= (\lambda\,\mathbf{u}_1\ \mathbf{u}_1 + \lambda\,\mathbf{u}_2\ \cdots\ \mathbf{u}_{k-1} + \lambda\,\mathbf{u}_k) \\ &= \lambda(\mathbf{u}_1\ \mathbf{u}_2\ \cdots\ \mathbf{u}_k) + (0\ \mathbf{u}_2\ \cdots\ \mathbf{u}_{k-1}) \\ &= (\mathbf{u}_1\ \mathbf{u}_2\ \cdots\ \mathbf{u}_k)(\lambda\,\mathbb{I} + N_k). \end{aligned}$$

(Here $\mathbb{I}$ is the $k \times k$ identity matrix.)

Example: $A = \begin{pmatrix} -1 & 1 & 1 \\ -2 & 2 & 1 \\ -1 & 1 & 1 \end{pmatrix}$.

The characteristic polynomial of A is $\chi_A(\lambda) = \lambda(\lambda-1)^2$, so 0 is an eigenvalue with multiplicity one and 1 is an eigenvalue with algebraic multiplicity 2. Summing all three columns of $s(A,1) = \begin{pmatrix} -2 & 1 & 1 \\ -2 & 1 & 1 \\ -1 & 1 & 0 \end{pmatrix}$ gives 0, so $\mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ is an eigenvector of A with eigenvalue 1. One can verify that the eigenspace of 1 is one dimensional, so we need to determine the generalized eigenspace. The second column of $s(A,\lambda)$ equals $\mathbf{u}_1$, so $s(A,\lambda)\mathbf{e}_2 = \mathbf{u}_1$; hence we can take $\mathbf{u}_2 = \mathbf{e}_2$. The generalized eigenspace equals span $\{\mathbf{u}_1, \mathbf{u}_2\}$. Summing the first two columns of A gives 0, so $\mathbf{u}_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ is an eigenvector of A with eigenvalue 0.

Note: This basis is a permutation of the one I used in lecture—here I've put the double eigenvalue first and the single one last. The intermediate matrices are different, but the final answer is, as it has to be, the same. The choice of basis, and hence of U, is *not* unique!

If we set

$$U = (\mathbf{u}_1\ \mathbf{u}_2\ \mathbf{u}_3) = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \qquad \text{then} \qquad U^{-1}A\,U = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

$U^{-1}A\,U$ is almost diagonal, and almost as easy to exponentiate as a diagonal matrix. One way of computing $\exp(t\,A)$ is the following:

$$U^{-1}AU = C + D, \qquad \text{where} \qquad C = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \text{and} \qquad D = \operatorname{diag}(1,1,0).$$

$C\,D = D\,C$ and $C^2 = \mathbb{O}$ imply that

$$\exp(t\,U^{-1}A\,U) = \exp(t\,(C+D)) = \exp(t\,C)\exp(t\,D) = (\mathbb{I}+t\,C)\operatorname{diag}\left(e^t, e^t, 1\right) = \begin{pmatrix} e^t & t\,e^t & 0 \\ 0 & e^t & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Hence

$$\exp(t\,A) = U\exp(t\,U^{-1}A\,U)\,U^{-1} = \begin{pmatrix} 1-t\,e^t & t\,e^t & e^t-1 \\ 1-(1+t)\,e^t & (1+t)\,e^t & e^t-1 \\ -t\,e^t & t\,e^t & e^t \end{pmatrix}.$$

□

A *block diagonal matrix* B consists of a collection of smaller square matrices straddling the diagonal, with zeroes elsewhere:

$$B = \mathrm{block}\,(B_1, \ldots, B_n) = \begin{pmatrix} B_1 & \mathbb{O} & \cdots & \mathbb{O} \\ \mathbb{O} & B_2 & \cdots & \mathbb{O} \\ \vdots & & \ddots & \vdots \\ \mathbb{O} & \cdots & \mathbb{O} & B_k \end{pmatrix},$$

where B_j is a $d_j \times d_j$ matrix, $j = 1, \ldots, k$. $\mathrm{block}\,(B_1, \ldots, B_n)^j = \mathrm{block}\left(B_1^j, \ldots, B_n^j\right)$; hence

$$\exp(t\,\mathrm{block}\,(B_1, \ldots, B_n)) = \mathrm{block}\,(\exp(t\,B_1), \ldots, \exp(t\,B_n))\,.$$

Exponentiating several little matrices is generally easier than exponentiating one big one; for example, $\exp(t\,\mathrm{diag}\,(\lambda_1, \ldots, \lambda_n)) = \mathrm{diag}\left(e^{\lambda_1 t}, \ldots, e^{\lambda_n t}\right)$ is a very easy calculation. Hence the following construction, called the *Jordan normal form* is very convenient. It guarantees that every matrix can be block diagonalized by a change of basis, with the blocks being matrices whose exponentials are explicitly known.

A *Jordan block* is a complex $k \times k$ matrix of the form

$$B = \lambda\,\mathbb{I} + N_k = \begin{pmatrix} \lambda & 1 & & \mathbb{O} \\ & \ddots & \ddots & \\ & & \lambda & 1 \\ \mathbb{O} & & & \lambda \end{pmatrix},$$

i.e. $b_{jj} = \lambda$, $j = 1, \ldots, k$, $b_{j\,j+1} = 1$, $j = 1, \ldots, k-1$, and all other entries are 0. A Jordan block is easily exponentiated: Since $\lambda\,\mathbb{I}$ commutes with everything,

$$\exp(t\,(\lambda\,\mathbb{I} + N_k)) = \exp(t\,\lambda\,\mathbb{I})\exp(t\,N_k) = e^{\lambda\,t}\sum_{j=0}^{k-1}\frac{t^j}{j!}S_{k,j}. \tag{2}$$

(Recall that $S_{k,j}$ has $i\ell$-th entry equal to 1 if $\ell = i + j$, 0 otherwise; see (1).)

Jordan normal form theorem: If A is an $n \times n$ complex matrix with eigenvalues $\lambda_1, \ldots \lambda_\ell$ (possibly repeated), then there is an invertible matrix U such that $U^{-1}\,A\,U = \mathrm{block}\,(B_1, \ldots, B_r)$ for some Jordan blocks $B_1, \ldots, B_r$. The blocks are uniquely determined, but the ordering depends on the choice of U.

Using the Jordan normal form, we can solve any linear homogeneous initial value problem $\dot{x} = A\,x$, $x(0) = x_0$, as follows: We know that the solution has the form $x(t) = \exp(t\,A)x_0$, so it suffices to compute $\exp(t\,A)$. Let U be a matrix that takes A to Jordan normal form, with blocks $B_1, \ldots B_r$. (The columns of U are (generalized) eigenvectors of A.) Then

$$\begin{aligned} \exp(t\,A) &= U\exp(t\,U^{-1}\,AU)\,U^{-1} \\ &= U\exp(t\,\mathrm{block}\,(B_1, \ldots, B_r))\,U^{-1} \\ &= U\,\mathrm{block}\,(\exp(t\,B_1), \ldots, \exp(t\,B_r))\;U^{-1}. \end{aligned}$$

The exponential of each of the Jordan blocks is given by (2).

Note: If we set $y(t) = U^{-1}\, x(t)$, then

$$\begin{aligned} y(t) &= U^{-1} \exp(t\, A)\, x_0 \\ &= U^{-1} \left(U \, \text{block}\left(\exp(t\, B_1), \ldots, \exp(t\, B_r)\right)\, U^{-1}\right) x_0 \\ &= \text{block}\left(\exp(t\, B_1), \ldots, \exp(t\, B_r)\right) y(0), \end{aligned}$$

so $y(t)$ satisfies $\dot{y} = \text{block}\left(B_1, \ldots, B_r\right)\, y$.