1. Jordan Canonical Form

An \( n \times n \) matrix \( A \) can be decomposed as

\[
A = QJQ^*
\]

where the matrix \( J \) is a block diagonal matrix

\[
J = \begin{bmatrix}
J_1 & & \\
& \ddots & \\
& & J_k
\end{bmatrix}
\]

and each block \( J_r \), for \( r = 1, \ldots, k \), has the form

\[
J_r = \begin{bmatrix}
\lambda_r & 1 & & \\
& \ddots & \ddots & \\
& & \ddots & 1 \\
& & & \lambda_r
\end{bmatrix}
\]

where \( J_r \) is \( n_r \times n_r \). This decomposition of \( A \) is known as the Jordan canonical form.

The Jordan canonical form provides valuable information about the eigenvalues of \( A \). The values \( \lambda_j \), for \( j = 1, \ldots, k \), are the eigenvalues of \( A \). For each distinct eigenvalue \( \lambda \), the number of Jordan blocks having \( \lambda \) as a diagonal element is equal to the number of linearly independent eigenvectors associated with \( \lambda \). This number is called the geometric multiplicity of the eigenvalue \( \lambda \). The sum of the sizes of all of these blocks is called the algebraic multiplicity of \( \lambda \).

We now consider \( J_r \)'s eigenvalues. We have

\[
\lambda(J_r) = \lambda_r, \ldots, \lambda_r
\]

where \( \lambda_r \) is repeated \( n_r \) times. But, because

\[
J_r - \lambda_r I = \begin{bmatrix}
0 & 1 & & \\
& \ddots & \ddots & \\
& & \ddots & 1 \\
& & & 0
\end{bmatrix}
\]

is a matrix of rank \( n_r - 1 \), it follows that the homogeneous system \((J_r - \lambda_r I)x = 0\) has only one vector (up to a scalar multiple) for a solution, and therefore there is only one eigenvector associated with this Jordan block.

Date: October 17, 2005, version 1.0.
Notes originally due to James Lambers. Minor editing by Lek-Heng Lim.
The unique unit vector that solves \((J_r - \lambda_r I)x = 0\) is the vector \(e_1 = [1, 0, \ldots, 0]^	op\). Now, consider the matrix

\[
(J_r - \lambda_r I)^2 = \begin{bmatrix}
0 & 1 & & \\
& & \ddots & \\
& & & 1 \\
\end{bmatrix}
\begin{bmatrix}
0 & 1 & & \\
& & \ddots & \\
& & & 1 \\
\end{bmatrix} = \begin{bmatrix}
0 & 0 & 1 \\
& & \\
& & \\
\end{bmatrix}.
\]

It is easy to see that \((J_r - \lambda_r I)^2e_2 = 0\). Continuing in this fashion, we can conclude that \((J_r - \lambda_r I)^k e_k = 0\), \(k = 1, \ldots, n_r - 1\).

The Jordan form can be used to easily compute powers of a matrix. For example, one can easily show that

\[
A^2 = QJQ^{-1}QJQ^{-1} = QJ^2Q
\]

and, in general,

\[
A^k = QJ^kQ.
\]

Due to its structure, it is easy to compute powers of a Jordan block \(J_r\). We have

\[
J_r^k = \begin{bmatrix}
\lambda_r & k \\
& & \ddots & k \\
& & & \lambda_r \\
\end{bmatrix}
= (\lambda_r I + K)^k, \quad K = \begin{bmatrix}
0 & 1 & & \\
& & \ddots & \\
& & & 1 \\
\end{bmatrix}
= \sum_{j=0}^k \binom{k}{j} \lambda_r^{k-j} K^j
\]

which yields, for \(k > n_r\),

\[
J_r^k = \begin{bmatrix}
\lambda_r^k & (k) \lambda_r^{k-1} & (k^2) \lambda_r^{k-2} & \cdots & (k^{n_r-1}) \lambda_r^{k-(n_r-1)} \\
& & \ddots & \ddots & \\
& & & \ddots & \lambda_r^k \\
\end{bmatrix}.
\]

For example,

\[
\begin{bmatrix}
\lambda & 1 & 0 \\
0 & \lambda & 1 \\
0 & 0 & \lambda \\
\end{bmatrix}^3 = \begin{bmatrix}
\lambda^3 & 3\lambda^2 & 3\lambda \\
0 & \lambda^3 & 3\lambda^2 \\
0 & 0 & \lambda^3 \\
\end{bmatrix}.
\]

We now consider an application of the Jordan canonical form. Consider the system of differential equations

\[
y'(t) = Ay(t), \quad y(t_0) = y_0.
\]

Using the Jordan form, we can rewrite this system as

\[
y'(t) = QJQ^{-1}y(t).
\]
Multiplying through by \( Q^{-1} \) yields
\[
Q^{-1} y'(t) = JQ^{-1} y(t),
\]
which can be rewritten as
\[
z'(t) = Jz(t),
\]
where \( z = Q^{-1} y(t) \). This new system has the initial condition
\[
z(t_0) = z_0 = Q^{-1} y_0.
\]
If we assume that \( J \) is a diagonal matrix (which is true in the case where \( A \) has a full set of linearly independent eigenvectors), then the system decouples into scalar equations of the form
\[
z'_i(t) = \lambda_i z_i(t),
\]
where \( \lambda_i \) is an eigenvalue of \( A \). This equation has the solution
\[
z_i(t) = e^{\lambda_i (t-t_0)} z_i(0),
\]
and therefore the solution to the original system is
\[
y(t) = Q \begin{bmatrix} e^{\lambda_1 (t-t_0)} & \cdots & e^{\lambda_n (t-t_0)} \end{bmatrix} Q^{-1} y_0.
\]

2. Systems of Linear Equations

The main content of the course will be concerned with solving problems of the form
\[
Ax = b
\]
where \( A \) is an \( m \times n \) matrix of rank \( r \). Of course, we cannot solve such problems generally, because the nature of the solution is influenced greatly by the relationship between \( m, n, \) and \( r \). For example, if \( m = n = r \), then there is a unique solution. On the other hand, if \( m < n \) and \( m = r \), then there are infinitely many solutions, and often we need to find the unique solution that satisfies certain constraints. This is the basis for linear programming. Finally, if \( m > n \), there may not be a solution, in which case we need to solve a least squares problem to find the vector \( x \) that comes as close as possible, in some sense, to solving the problem.

3. Some Results Involving Norms

If \( \|A\| < 1 \), then \( \|A^m\| \to 0 \) as \( m \to \infty \). Since \( \|A\| \) is a continuous function of the elements of \( A \), it follows that \( A^m \to 0 \). However, if \( \|A\| > 1 \), it does not follow that \( \|A^m\| \to \infty \). For example, suppose
\[
A = \begin{bmatrix} 0.99 & 10^6 \\ 0 & 0.99 \end{bmatrix}.
\]
In this case, \( \|A\|_\infty > 1 \), but because \( \rho(A) < 1 \), there must exist some norm \( \|A\|_\alpha \) such that \( \|A\|_\alpha < 1 \).

For matrices \( A \) such that \( \|A\| < 1 \) for some natural norm, we also have the following result.

**Theorem 1.** If, for some natural norm, \( \|A\| < 1 \), then
(a) \( I - A \) is nonsingular
(b) \[
\frac{1}{1 + \|A\|} \leq \|(I - A)^{-1}\| \leq \frac{1}{1 - \|A\|}.
\]

**Proof.** (a) Assume \( I - A \) is singular. Then, there exists a vector \( z \neq 0 \) such that \( (I - A)z = 0 \). Therefore \( z = Az \) and \( \|z\| = \|Az\| \leq \|A\| \|z\| \).

Therefore \( \|A\| \geq 1 \), which is a contradiction.
(b) Since \( I = (I - A)(I - A)^{-1} \), we have

\[
\|I\| \leq \|(I - A)\||I - A||^{-1}\|
\]

but since we are using a natural norm, \( \|I\| = 1 \), so, by the triangle inequality, we have

\[
1 \leq (1 + \|A\|)\|I - A||^{-1}\|
\]

thus proving the left inequality. For the right inequality, \((I - A)^{-1}(I - A) = I \) yields

\[
(I - A)^{-1} - (I - A)^{-1}A = I
\]

or

\[
(I - A)^{-1} = I + (I - A)^{-1}A.
\]

Taking norms, we obtain

\[
\|(I - A)^{-1}\| \leq 1 + \|A\||I - A||^{-1}\|
\]

which, by the fact that \( \|A\| < 1 \), proves the right inequality true. \( \square \)

Department of Computer Science, Gates Building 2B, Room 280, Stanford, CA 94305-9025

E-mail address: golub@stanford.edu