Let $A$ be a constant $n \times n$ real matrix. We use the same symbol $A$ to denote the corresponding linear operator $A : \mathbb{R}^n \to \mathbb{R}^n$. Because a real matrix $A$ can have nonreal (complex with nonzero real part) eigenvalues, we also consider the complexification of $A$, the linear operator $A : \mathbb{C}^n \to \mathbb{C}^n$ that considers the same real matrix $A$ acting as a linear operator from $\mathbb{C}^n$ into $\mathbb{C}^n$.

Eigenvalues, eigenvectors, multiplicity and spectrum

An eigenvalue of $A$ is a complex number $\lambda_j$ that satisfies

$$Av_j = \lambda_j v_j$$

for some nonzero vector $v_j \in \mathbb{C}^n$. Such a vector $v_j$ is called an eigenvector for the eigenvalue $\lambda_j$. If $\lambda_j$ is a real eigenvalue then it is possible to choose an eigenvector $v_j$ for $\lambda_j$ so that it is real. Let us assume that this is always done. Then we have $Av_j = \lambda_j v_j$ in $\mathbb{R}^n$. If $\lambda_j \in \mathbb{C}$ is a nonreal eigenvalue, then any eigenvector for $\lambda_j$ is a vector $v_j = a_j + ib_j \in \mathbb{C}^n$ whose real and imaginary parts $a_j$ and $b_j$ are linearly independent (real) vectors in $\mathbb{R}^n$.

The eigenvalues of $A$ are the same as the roots of the characteristic polynomial $h(\lambda) = \det(A - \lambda I_n)$, where $I_n$ is the $n \times n$ identity matrix. Since $h(\lambda)$ is a polynomial of degree $n$, it has $n$ complex roots, not necessarily distinct. An eigenvalue $\lambda_j$ has multiplicity $m_j$ if $\lambda_j$ is a distinct root of the characteristic polynomial $h(\lambda)$ of multiplicity $m_j$, i.e. $h(\lambda)$ factors as $h(\lambda) = (\lambda - \lambda_j)^{m_j} p(\lambda)$, with $p(\lambda_j) \neq 0$.

The spectrum of a matrix or linear operator $A$ is the set, denoted $\sigma(A)$, of all the distinct eigenvalues $\{\lambda_1, \ldots, \lambda_d\}$ of $A$, where the number of distinct eigenvalues $d$ is no greater than the dimension $n$ of the vector space $\mathbb{C}^n$. If $m_j = 1$ for all eigenvalues $\lambda_j$, then $d = n$; if $m_j > 1$ for some eigenvalue $\lambda_j$, then $d < n$.

If $\lambda_j$ is an eigenvalue of $A$ with multiplicity $m_j$, then its complex conjugate $\overline{\lambda_j}$ is also an eigenvalue of $A$ (recall that $A$ is a real matrix), and it has the same multiplicity $m_j$ (of course, this is trivial if $\lambda_j$ is real).

Direct sums, generalized eigenvectors, and the Jordan normal form

If $V_1$ and $V_2$ are two subspaces of a vector space, then the sum of $V_1$ and $V_2$ is the subspace $V_1 + V_2$ of all vectors of the form $v = v_1 + v_2$ with $v_1 \in V_1$ and $v_2 \in V_2$. If the subspaces $V_1$ and $V_2$ satisfy $V_1 \cap V_2 = \{0\}$, then their sum is called a direct sum, and the representation $v = v_1 + v_2$ with $v_1 \in V_1$ and $v_2 \in V_2$ is unique. We write the direct sum of $V_1$ and $V_2$ as $V_1 \oplus V_2$.

A subspace $V$ is (algebraically) invariant under $A$ if $v \in V$ implies $Av \in V$. The statement that $\mathbb{C}^n = V_1 \oplus V_2$ where $V_1$ and $V_2$ are both invariant under $A$ is the same as the statement that there exists some basis of $\mathbb{C}^n$ with respect to which the matrix representation of $A$ is block diagonal. We will find the “best” basis (consisting of generalized eigenvectors) so that the matrix representation of $A$ is block diagonal and is “as simple as possible”. Details can get a bit complicated if there are nonreal eigenvalues, or multiple eigenvalues ($m_j > 1$) (or nonreal multiple eigenvalues).
If $\lambda_j$ is a distinct eigenvalue of $A$ with multiplicity $m_j$, we consider the subspace of $\mathbb{C}^n$ given by

$$E(\lambda_j) = \mathcal{N}((A - \lambda_j I_n)^{m_j}),$$

where $\mathcal{N}$ denotes the null space (or kernel) of a linear operator. A nonzero vector in the subspace $E(\lambda_j)$ is called a generalized eigenvector corresponding to $\lambda_j$. A theorem states that there always exist $m_j$ linearly independent generalized eigenvectors $v_j^{[1]}, \ldots, v_j^{[m_j]}$ in $\mathbb{C}^n$ that span this subspace:

$$E(\lambda_j) = \text{span}\{v_j^{[1]}, \ldots, v_j^{[m_j]}\},$$

so the dimension (in $\mathbb{C}^n$) of the subspace $E(\lambda_j)$ is $m_j$. If $\lambda_1, \ldots, \lambda_d$ are all the distinct eigenvalues of $A$, the direct sum of all the corresponding subspaces $E(\lambda_j)$ is all of $\mathbb{C}^n$:

$$E(\lambda_1) \oplus \cdots \oplus E(\lambda_d) = \mathbb{C}^n,$$

or, written more conveniently,

$$\bigoplus_{\lambda_j \in \sigma(A)} E(\lambda_j) = \mathbb{C}^n.$$

This direct sum decomposition of $\mathbb{C}^n$ into the $m_j$-dimensional complex subspaces $E(\lambda_j)$, corresponding to the distinct eigenvalues $\lambda_j$ with their multiplicities $m_j$, implies that the matrix $A$ can be block-diagonalized by a complex nonsingular similarity transformation (i.e. a complex linear change of coordinates).

We construct an $n \times n$ complex change of coordinates matrix $P$ by using as the columns of $P$ suitably chosen basis vectors $v_j^{[1]}, \ldots, v_j^{[m_j]}$ for each of the subspaces $E(\lambda_j) = \mathcal{N}((A - \lambda_j I_n)^{m_j})$, for all the distinct eigenvalues $\lambda_j$, $j = 1, \ldots, d$. If $A$ has a nonreal eigenvalue $\lambda_j$, then one can always choose the basis vectors for $E(\bar{\lambda}_j)$ to be the complex conjugates of the basis vectors for $E(\lambda_j)$. In this case, certain columns of the change of basis matrix $P$ will be the complex conjugates of some other columns.

**Theorem 1.2** (Jordan Normal Form). Let $A$ be a real $n \times n$ matrix. Then there exists a linear nonsingular change of variables $x = Pz$ (if all the eigenvalues of $A$ are real, then the matrix $P$ can be chosen to be real; if $A$ has any nonreal eigenvalue, then $P$ is nonreal) that transforms

$$\dot{x} = Ax \quad \text{into} \quad \dot{z} = Jz,$$

where $J = P^{-1}AP$ is block-diagonal, with square blocks of various sizes along the diagonal

$$J = \begin{pmatrix} J_1 & & \\ & J_2 & \\ & & \ddots \\ & & & J_L \end{pmatrix}$$

and zeros elsewhere. Each block $J_\ell$, $\ell = 1, \cdots, L$ has the form

$$J_\ell = (\lambda_j) \quad \text{or} \quad J_\ell = \begin{pmatrix} \lambda_j & 1 \\ \lambda_j & \ddots \\ \ddots & \ddots & \ddots \\ \lambda_j & \ddots & 1 \\ & \cdots & \cdots & \lambda_j \end{pmatrix},$$

i.e. a $1 \times 1$ block containing an eigenvalue, or a larger block with the eigenvalue $\lambda_j$ in each position on the main diagonal, the number $1$ in each position on the first superdiagonal, and $0$ in every other position.
The blocks $J_i$ are called \textit{elementary Jordan blocks} and the matrix $J$ (which is unique up to the ordering of the elementary Jordan blocks) is called the \textit{Jordan normal form} (or \textit{Jordan canonical form}) of $A$.

Note that the same eigenvalue can appear on the main diagonal of more than one elementary Jordan block (examples given in the lectures and homework), so the size of an elementary Jordan block is not necessarily the same as the multiplicity $m_j$ of the corresponding eigenvalue $\lambda_j$, and the number of elementary Jordan blocks in the Jordan normal form is not necessarily the same as the number of distinct eigenvalues.

If $A$ has any nonreal eigenvalues, then for the Jordan normal form we are forced to work in $\mathbb{C}^n$, even if $A$ is a real matrix (this problem is fixed below).

\section*{Generalized eigenspaces and the real normal form}

Once we have the Jordan normal form of a matrix, it is a simple matter (except possibly for some tedious bookkeeping) to get a corresponding normal form for $A$ that is real, even if there are nonreal eigenvalues.

Let us return to the discussion above, of the $m_j$-dimensional subspaces $E(\lambda_j) = \mathcal{N}((A - \lambda_j I_n)^{m_j})$ of (complex) generalized eigenvectors, corresponding to each distinct eigenvalue $\lambda_j$ with multiplicity $m_j$.

If the eigenvalue $\lambda_j$ is real, then it is possible to choose the $m_j$ linearly independent vectors $v_j^{[1]}, \ldots, v_j^{[m_j]}$ that span $E(\lambda_j)$ so that they are all in $\mathbb{R}^n$. We will always assume that this is done. In this case we define the \textit{generalized eigenspace} for the real eigenvalue $\lambda_j$ to be the $m_j$-dimensional subspace of $\mathbb{R}^n$ given by

$$X(\lambda_j) = \text{span} \{v_j^{[1]}, \ldots, v_j^{[m_j]}\}.$$  

In other words, if $\lambda_j$ is real, we take

$$X(\lambda_j) = E(\lambda_j) \cap \mathbb{R}^n.$$  

Any nonzero vector in $X(\lambda_j)$ is a real generalized eigenvector corresponding to $\lambda_j$. All real eigenvectors corresponding to the real eigenvalue $\lambda_j$ are automatically included in $X(\lambda_j)$.

If the eigenvalue $\lambda_j$ is nonreal, then the $m_j$ linearly independent generalized eigenvectors $v_j^{[1]} = a_j^{[1]} + ib_j^{[1]}, \ldots, v_j^{[m_j]} = a_j^{[m_j]} + ib_j^{[m_j]}$ that span $E(\lambda_j)$ are nonreal, and the $2m_j$ real and imaginary parts of the generalized eigenvectors $a_j^{[1]}, \ldots, a_j^{[m_j]}, b_j^{[1]}, \ldots, b_j^{[m_j]}$ are linearly independent vectors in $\mathbb{R}^n$. In this case we define the \textit{generalized eigenspace} for the nonreal eigenvalue $\lambda_j$ (and also of $\bar{\lambda}_j$) to be the $2m_j$-dimensional subspace of $\mathbb{R}^n$ given by

$$X(\lambda_j, \bar{\lambda}_j) = \text{span} \{a_j^{[1]}, \ldots, a_j^{[m_j]}, b_j^{[1]}, \ldots, b_j^{[m_j]}\}.$$  

It can be shown that

$$X(\lambda_j, \bar{\lambda}_j) = (E(\lambda_j) \oplus E(\bar{\lambda}_j)) \cap \mathbb{R}^n.$$  

The real and imaginary parts of all eigenvectors and generalized eigenvectors for nonreal $\lambda_j$ are included in the generalized eigenspace $X(\lambda_j, \bar{\lambda}_j)$, which is a real subspace. Taking all the (real or nonreal) eigenvalues of $A$ together, the direct sum of their generalized eigenspaces is all of $\mathbb{R}^n$.

By carefully choosing basis vectors from the generalized eigenspaces to form a basis for $\mathbb{R}^n$, we can use these basis vectors as the columns of a real nonsingular $n \times n$ matrix $Q$ that makes the desired transformation of $A$.

\textbf{Theorem 1.3 (Real Normal Form).} Let $A$ be a real $n \times n$ matrix. Then there is always a real linear nonsingular change of variables $x = Qy$ that transforms

$$\dot{x} = Ax \quad \text{into} \quad \dot{y} = Ry,$$  

into
where $R = Q^{-1}AQ$ is block diagonal with square blocks of various sizes along the diagonal

$$R = \begin{pmatrix} R_1 & & \\ & R_2 & \\ & & \ddots \\ & & & R_M \end{pmatrix}$$

and zeros elsewhere.

i) If $\lambda_j$ is a real eigenvalue, then a corresponding block $R_m$ ($m = 1, \cdots, M$) is an elementary Jordan block

$$R_m = \begin{pmatrix} \lambda_j \\ 1 \\ & \lambda_j \\ & & \ddots \\ & & & 1 \\ & & & & \lambda_j \end{pmatrix};$$

ii) If $\lambda_j = \mu_j + i\omega_j$ ($\mu_j$ and $\omega_j$ both real, $\omega_j > 0$) is a nonreal eigenvalue, then a corresponding block $R_m$ ($m = 1, \cdots, M$) has the form

$$R_m = D_j \quad \text{or} \quad R_m = \begin{pmatrix} D_j & I_2 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots \\ I_2 & \cdots & D_j \end{pmatrix},$$

where $D_j$ and $I_2$ are the $2 \times 2$ matrices

$$D_j = \begin{pmatrix} \mu_j & -\omega_j \\ \omega_j & \mu_j \end{pmatrix}, \quad I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The blocks $R_m$ are called elementary real blocks and the block-diagonal matrix $R$ (which is unique up to the ordering of the elementary real blocks) is called the real normal form (or real canonical form) of the matrix $A$.

**NOTE:** If $\lambda_j = \mu_j + i\omega_j$ ($\mu_j$, $\omega_j$ real, $\omega_j > 0$) is a nonreal eigenvalue and the complex vector $v_j^{[p]} = a_j^{[p]} + ib_j^{[p]}$ ($a_j^{[p]}$, $b_j^{[p]}$ both real) is a basis vector of $E(\lambda_j)$ used as a column of the matrix $P$ to obtain the Jordan normal form $J$, then by convention we use the two real vectors $a_j^{[p]}$ and $-b_j^{[p]}$ (NOTE THE MINUS SIGN!!) as columns in the matrix $Q$ to obtain the real normal form $R$. (The minus sign in $-b_j^{[p]}$ is to obtain $-\omega_j < 0$ in the upper right corner of $D_j$ and $\omega_j > 0$ in the lower left corner; it will be important to remember this convention to avoid confusion when we do calculations at a Hopf bifurcation.)
Fredholm Alternative Theorem

For two real vectors $u, v$ of the same dimension $n$, their inner product (or scalar product) is

$$\langle u, v \rangle = u^\top v = \begin{pmatrix} u_1 & \cdots & u_n \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \sum_{j=1}^{n} u_j v_j.$$

Let $A$ be a real $n \times m$ matrix and let $y \in \mathbb{R}^n$ be a real vector.

**Theorem 3.9** (Fredholm Alternative). The equation $Ax = y$ has a solution if and only if $\langle p, y \rangle = 0$ for every $p \in \mathbb{R}^n$ satisfying $A^\top p = 0$.

By definition, the nonhomogeneous linear equation

$$Ax = y$$

has a solution $x \in \mathbb{R}^m$ if and only if $y$ belongs to the range of $A$. Thus, the Fredholm Alternative Theorem states that the range of $A$, denoted $A(\mathbb{R}^m)$, is the orthogonal complement in $\mathbb{R}^n$ of the null space (or kernel) of the transposed matrix $A^\top$:

$$A(\mathbb{R}^m) = (\mathcal{N}(A^\top))^\perp.$$

If $A$ is a complex matrix and $y \in \mathbb{C}^n$, the Fredholm Alternative Theorem remains valid if we define the inner product as

$$\langle u, v \rangle = \bar{u}^\top v = \begin{pmatrix} \bar{u}_1 & \cdots & \bar{u}_n \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \sum_{j=1}^{n} \bar{u}_j v_j,$$

and replace the transposed matrix $A^\top$ by the adjoint matrix $A^* = \bar{A}^\top$ (transposed and conjugated).

**Additional References:** Hale; Hirsch, Smale & Devaney; Meiss.