Contents

No diagonals

Chapter 14

Nondiagonalizable Matrices, the Jordan Normal Form

We study sparse matrix representations for square matrices A that cannot be diagonalized under similarity.

14.1 Lecture Fourteen (The Jordan Normal Form)

We develop the Jordan normal form for matrices A_{nn} that do not have a complete eigenvector basis.

This chapter extends Chapter 9. Both versions of Lecture 9 in Sections 9.1 and 9.1.D terminated with a subsection on diagonalizable matrices. We recall that a square matrix $A \in \mathbb{R}^{n,n}$ (or $\mathbb{C}^{n,n}$) is diagonalizable by matrix similarity $D = X^{-1}AX = \text{diag}(\lambda_i)$ if and only if there is a basis of eigenvectors in $\mathbb{R}^n$ (or $\mathbb{C}^n$) for A that appears in the columns of X. See specifically equation (7.4) and Theorems 9.4 and 9.2.D in Sections 7.1, 9.1, and 9.1.D, respectively.
If A is diagonalizable and $D = X^{-1}AX$ is diagonal, then A's eigenvalues λ_i occur on the diagonal of D and its eigenvectors x_i are the columns of X.
If A is not diagonalizable, then A cannot have a complete eigenvector basis in $\mathbb{R}^n$ (or $\mathbb{C}^n$).
Recalling the notion of eigenspace $E(\lambda_i) = \ker(A - \lambda_i I)$ of A for each distinct eigenvalue $\lambda_i,\ i = 1, ..., k \leq n$, we observe that for a **non diagonalizable matrix** A the join of its eigenspaces cannot encompass the whole space, i.e., with the subspace join notation from Section 4.2, we must have

$$E(\lambda_1) \ + \ ... \ + \ E(\lambda_k) \ \subsetneqq \ \mathbb{R}^n \ \ (\text{or } \mathbb{C}^n).$$

[0]© Copyright 2002 by Frank Uhlig; “Work in progress”; latest update of 16th November 2005.

According to Section 9.1, a non diagonalizable matrix A has a minimal polynomial of the form $p_A(x) = \prod_{i=1}^{k} (x-\lambda_i)^{j_i}$ where for at least one eigenvalue λ_i the ***index***, or exponent j is greater than 1.
Likewise according to Theorem 9.8 of section 9.2, a non diagonalizable matrix A must have at least one eigenvalue λ_i with differing geometric and algebraic multiplicity, namely

$$\dim(E(\lambda_i)) \;=\; \text{geom. mult.}(\lambda_i) \;<\; \text{alg. mult.}(\lambda_i) \;=\; n_i \;,$$

where $f_A(\lambda) = \prod_{i=1}^{k} (\lambda-\lambda_i)^{n_i}$ is the characteristic polynomial $\det(\lambda I - A)$ of A and the $\lambda_i \in \mathbb{R}$ or $\mathbb{C}$ are distinct. According to the Corollary that is common to both, the end of Sections 9.1 and 9.1.D, to study non diagonalizable matrices A_{nn} thus means to study matrices with at least one repeated eigenvalue.

If a matrix A is not diagonalizable, and if λ is one eigenvalue of A with insufficiently many eigenvectors, i.e., if $\text{index}(\lambda) > 1$ in A's minimal polynomial, or if the geometric multiplicity of λ is less than its algebraic multiplicity in A's characteristic polynomial, then $\ker(A-\lambda I)^2 \supsetneqq \ker(A-\lambda I) = E(\lambda)$ as we shall shortly see. For simplicity, we start this section with matrices A that have only one eigenvalue λ.

(a) ***A matrix A with only one eigenvalue λ***

If A is diagonalizable and has only one eigenvalue λ, then $X^{-1}AX = \lambda I$, or $A = \lambda I$.
For general, not necessarily diagonalizable matrices A we first study a generalization of matrix eigenspaces, namely the **principal subspaces**.

Definition 1: For an eigenvalue $\lambda \in \mathbb{R}$ or $\mathbb{C}$ of $A_{n,n}$, the subspace $P_k(\lambda) := \ker(A-\lambda I)^k$ of $\mathbb{R}^n$ or $\mathbb{C}^n$ is called the **principal subspace** of ***order*** k for A and λ where $k = 0, 1, 2, \ldots.$ ◀

Clearly $P_0(\lambda) = \ker(A-\lambda I)^0 = \ker(I) = \{0\}$ while $P_1(\lambda) = \ker(A-\lambda I) = E(\lambda)$ is the eigenspace of A for λ. Note that in general, i.e., for diagonalizable and non diagonalizable matrices alike, each principal subspace $P_k(\lambda)$ is a matrix kernel, and therefore according to Chapter 4 it is a subspace of $\mathbb{R}^n$ (or of $\mathbb{C}^n$ if $\lambda \notin \mathbb{R}$).
Moreover $P_1(\lambda) \subset P_2(\lambda) \subset \ldots$ is an ascending chain of subspaces of $\mathbb{R}^n$ (or $\mathbb{C}^n$) for each fixed eigenvalue λ of A. This follows since $y \in P_k(\lambda)$ implies $(A-\lambda I)^k y = 0$ and thus

$$(A-\lambda I)^{k+\ell} y = (A-\lambda I)^{\ell}\left((A-\lambda I)^k y\right) = (A-\lambda I)^{\ell} 0 = 0 \;,$$

or $y \in P_{k+\ell}(\lambda)$ for each $\ell \geq 0$. More holds for the vectors in a principal subspace chain $P_1(\lambda) \subset P_2(\lambda) \subset \ldots$ of A and an eigenvalue λ.

Lemma 1: Let A be an arbitrary square matrix with the eigenvalue λ. If for some order k the principal subspaces $P_k(\lambda) = P_{k+1}(\lambda)$ are equal, then $P_{k+\ell}(\lambda) = P_k(\lambda)$ for all $\ell \geq 0$.
In other words, if the sequence of principal subspaces $P_i(\lambda)$ becomes stationary for one order $i = k$, then it remains so for all larger orders $k + \ell$. ◀

Proof: We have already noticed that $P_k(\lambda) \subset P_{k+\ell}(\lambda)$ for all k and $\ell \geq 0$. Hence we only need to show that $P_{k+\ell}(\lambda) \subset P_k(\lambda)$ in order to ascertain equality, given that $P_k(\lambda) = P_{k+1}(\lambda)$.
If $x \in P_{k+\ell}(\lambda)$ for $\ell > 1$, then $(A-\lambda I)^{k+\ell}x = 0$, or $(A-\lambda I)^{k+1}\left((A-\lambda I)^{\ell-1}x\right) = 0$. Thus $(A-\lambda I)^{\ell-1}x \in P_{k+1}(\lambda) = P_k(\lambda)$ as assumed and $(A-\lambda I)^k\left((A-\lambda I)^{\ell-1}x\right) = (A-\lambda I)^{k+\ell-1}x = 0$ as well, or $x \in P_{k+\ell-1}(\lambda)$. By repeating this order reduction, we eventually obtain that $x \in P_{k+1}(\lambda) = P_k(\lambda)$. I.e., $P_{k+\ell}(\lambda) \subset P_k(\lambda)$ for all $\ell \geq 0$, provided that $P_k(\lambda) = P_{k+1}(\lambda)$. ■

We specialize this result now to a non diagonalizable matrix A with only one eigenvalue λ. Such a matrix has a minimal polynomial of the form $p_A(x) = (x-\lambda)^j$ for an exponent $2 \leq j \leq n$ according to Section 9.1. But the minimality of p_A together with the fact that $p_A(A) = (A-\lambda I)^j = O_n$ makes $P_j(\lambda) = \mathbb{R}^n$ (or $\mathbb{C}^n$) while $P_{j-1}(\lambda) \subsetneqq P_j(\lambda)$. Thus we have the following relations among the principal subspaces for the eigenvalue λ of precise index $j > 1$ when A has exactly one eigenvalue λ:

$$\{0\} = P_0(\lambda) \subsetneqq E(\lambda) = P_1(\lambda) = \ker(A-\lambda I) \subsetneqq P_2(\lambda) = \ker(A-\lambda I)^2 \subsetneqq \ldots.$$
$$\ldots. \subsetneqq P_{j-1}(\lambda) = \ker(A-\lambda I)^{j-1} \subsetneqq P_j(\lambda) = \ker(A-\lambda I)^j = \mathbb{R}^n \quad (\text{or } \mathbb{C}^n)\ .$$

Of particular interest for us are those vectors that lie in one principal subspace $P_\ell(\lambda)$, but do not lie in $P_{\ell-1}(\lambda)$. Since $P_{\ell-1}(\lambda) \subsetneqq P_\ell(\lambda)$, such vectors always exist as long as $1 \leq \ell \leq j$, where j is the index of λ.

Definition 2: The vectors $x \in P_\ell(\lambda)$ with $x \notin P_{\ell-1}(\lambda)$ are called **principal vectors** of ***grade*** ℓ for A and λ. ◀

Note that not all vectors in the principal vector subspace $P_k(\lambda)$ of order k for A and λ are principal vectors of grade k. Clearly $0 \in P_k(\lambda)$ has grade zero and any principal vector for λ of grade less than k will be annihilated by $(A-\lambda I)^k$, i.e., any lower grade principal vector belongs to $P_k(\lambda)$ by default.
Principal vectors of grade one are customarily called eigenvectors for λ. Principal vectors of arbitrary grade will help us find sparse and revealing matrix normal forms for non diagonalizable matrices A. Note that principal vectors exist for each eigenvalue λ of a square matrix A up to and including the grade equal to the index of λ.

In Chapter 9 we have seen that eigenvectors corresponding to distinct eigenvectors of one matrix are linearly independent. In this chapter we show that the same is true for carefully chosen **chains of principal vectors**. Moreover we learn how these principal vector chains can complete the incomplete eigenvector set for a non diagonalizable matrix to a full basis of $\mathbb{R}^n$ or $\mathbb{C}^n$. And thereby we obtain a non diagonal sparse normal form for such matrices under similarity.

Definition 3: If $x^{(k)}$ is a principal vector of grade $k \geq 1$ for an arbitrary square matrix A and its eigenvalue λ, then the k vectors

$$x^{(k)},\ x^{(k-1)} := (A-\lambda I)x^{(k)} \in P_{k-1}(\lambda),\ \ldots\ ,$$
$$x^{(l)} := (A-\lambda I)^{k-l}x^{(k)} = (A-\lambda I)x^{(l+1)} \in P_{k-l}(\lambda),\ \ldots\ ,$$
$$x^{(1)} := (A-\lambda I)^{k-1}x^{(k)} = (A-\lambda I)x^{(2)} \in P_1(\lambda) = E(\lambda)$$

form a **principal vector chain** for A and λ of length k. ◀

Lemma 2: The vectors $x^{(k)}$, $x^{(k-1)}$, ... $x^{(1)}$ of a principal vector chain are linearly independent. ◀

Proof: If $x^{(k)}$ is a principal vector of grade k, then

$$(A-\lambda I)^{k-\ell}(A-\lambda I)^{\ell}x^{(k)} = (A-\lambda I)^{k}x^{(k)} = 0$$

and therefore $(A-\lambda I)^{\ell}x^{(k)} \in P_{k-\ell}(\lambda)$ for each $k-1 \geq \ell \geq 0$.
To show linear independence we use the linear independence Definition 2 of Section 5.1: Assume that $\alpha_k x^{(k)} + \alpha_{k-1}x^{(k-1)} + \ \ldots\ + \alpha_1 x^{(1)} = 0$. Then

$$-\alpha_k x^{(k)} \ = \ \alpha_{k-1}x^{(k-1)} \ + \ \ldots \ + \ \alpha_1 x^{(1)}\ . \tag{14.1}$$

Since $x^{(k)}$ is a principal vector of grade k for A and its eigenvalue λ, the left hand vector $-\alpha_k x^{(k)}$ in (14.1) does not lie in $P_{k-1}(\lambda) = \ker(A-\lambda I)^{k-1}$ unless $\alpha_k = 0$. On the other hand, the right hand side linear combination in (14.1) does lie in $P_{k-1}(\lambda)$.
Thus $\alpha_k = 0$ and we are left to analyze the shortened equation $\alpha_{k-1}x^{(k-1)} + \ \ldots\ + \alpha_1 x^{(1)} = 0$, or $-\alpha_{k-1}x^{(k-1)} = \alpha_{k-2}x^{(k-2)} + \ \ldots\ + \alpha_1 x^{(1)}$. Again, the left hand side vector does not belong to $P_{k-2}(\lambda)$ unless $\alpha_{k-1} = 0$, while the right hand side does, implying that $\alpha_{k-1} = 0$. By repeating this argument we conclude that $\alpha_k = \alpha_{k-1} = \ \ldots\ = \alpha_2 = \alpha_1 = 0$, i.e., linear independence of the chain of specially chosen principal vectors $x^{(k)},\ \ldots\ , x^{(1)}$ for A and λ. ■

If we define the integers $d_k := \dim P_k(\lambda) - \dim P_{k-1}(\lambda) \geq 0$ and $p_k := \dim P_k(\lambda)$, then for each k there is a basis $\{x_1, ..., x_{p_{k-1}}\}$ for $P_{k-1}(\lambda)$ and a basis $\{x_1, ..., x_{p_k}\}$ for $P_k(\lambda)$ whose first p_{k-1} members form a basis for the principal subspace $P_{k-1}(\lambda)$. Similar to Lemma 2, the collection of all $d_k \cdot k$ principal vectors that originate from the $d_k = p_k - p_{k-1}$ linearly independent principal vectors $x_{p_{k-1}+1}, x_{p_{k-1}+2}, .., x_{p_k}$ of grade k for A and λ can be shown to be linearly independent. The key argument in proving this is the obvious assertion that any linear combination of the linearly independent principal vectors $x_{p_{k-1}+1}, x_{p_{k-1}+2}, .., x_{p_k} \in P_k(\lambda) \setminus P_{k-1}(\lambda)$ of grade k which lies in $P_{k-1}(\lambda)$, or has the grade $k-1$ or less as a principal vector, must be the zero vector.

Lemma 3: If $y_1, ..., y_{p_{k-1}}$ is a basis for the principal subspace $P_{k-1}(\lambda)$ of order $k-1$ for A and an eigenvalue λ of index at least k, and if $y_1, ..., y_{p_{k-1}}, ..., y_{p_k}$ is a basis for $P_k(\lambda)$, then the $d_k \cdot k$ vectors $(A-\lambda I)^{\ell}y_i,\ i = p_{k-1}+1, ..., p_k$, with $d_k = \dim P_k(\lambda) - \dim P_{k-1}(\lambda) = p_k - p_{k-1}$ and $\ell = 0, ..., k-1$ are linearly independent. ◀

Proof: Assume that for certain coefficients $\alpha_i,\ \beta_i,\ ...,\ \delta_i$ the linear combination

$$\sum_{i=p_{k-1}+1}^{p_k} \alpha_i y_i + (A-\lambda I)\left(\sum_{i=p_{k-1}+1}^{p_k} \beta_i y_i\right) + \ ...\ + (A-\lambda I)^{k-1}\left(\sum_{i=p_{k-1}+1}^{p_k} \delta_i y_i\right) \ = \ 0\ . \tag{14.2}$$

This vector equation contains k sum terms on its left. The first sum term unites the eigenvectors, the second term the principal vectors of precise grade 2, and so forth until the last summed term which contains the principal vectors of maximal grade k for A and λ. If we multiply (14.2) on the left by $(A-\lambda I)^{k-1}$, then the last $k-1$ sum terms vanish since by construction each y_i is a principal vector of grade k for each $p_{k-1} < i \leq p_k$. Thus we are left with the equation $(A-\lambda I)^{k-1}\sum \alpha_i y_i = 0$, which in turn implies $\sum_i \alpha_i y_i = 0$ since $\sum_i \alpha_i y_i \notin P_{k-1}(\lambda)$. Since the eigenvectors y_i are linearly independent by choice, all $\alpha_i = 0$.
Thus equation (14.2) simplifies to

$$(A-\lambda I)\sum_{i=p_{k-1}+1}^{p_k} \beta_i y_i + \ ...\ + (A-\lambda I)^{k-1}\sum_{i=p_{k-1}+1}^{p_k} \delta_i y_i \ = \ 0$$

with $k-1$ sum terms. If this equation is multiplied from the left by $(A-\lambda I)^{k-2}$ it yields the simpler equation $(A-\lambda I)^{k-1}\sum_i \beta_i y_i = 0$ since each of the remaining $k-2$ sum terms vanishes because $(A-\lambda I)^k y_i = 0$. As before, we conclude that all $\beta_i = 0$. Continuing in this fashion proves the linear independence of all $d_k \cdot k$ principal vectors $(A-\lambda I)^\ell y_i$ in the d_k distinct principal vector chains for A and λ of index k where $i = p_{k-1}+1, ..., p_k$ and $\ell = 0, 1, ..., k-1$. ■

Let us look at the effect of the linear independence of all principal vector chains for a matrix A with one eigenvalue λ. In particular we try to represent A with respect to each principal vector chain of order k. If $x^{(k)} \in P_k(\lambda)$ and $x^{(k)} \notin P_{k-1}(\lambda)$ is a principal vector of grade k for A and λ, then its associated principal vector chain consists of the vectors $x^{(k)},\ x^{(k-1)} := (A-\lambda I)x^{(k)},\ x^{(k-2)} := (A-\lambda I)x^{(k-1)},\ ...,\ x^{(1)} := (A-\lambda I)x^{(2)}$, where the upper index (ℓ) indicates the grade of the principal vector $x^{(\ell)} \in P_\ell(\lambda)$ with $x^{(\ell)} \notin P_{\ell-1}(\lambda)$ throughout. Now look at the matrix product $A_{nn} \cdot \begin{pmatrix} | & | & & | \\ x^{(k)} & x^{(k-1)} & \dots & x^{(1)} \\ | & | & & | \end{pmatrix}_{nk}$. Since $x^{(\ell-1)} = (A-\lambda I)x^{(\ell)}$ by construction for each $\ell = 2, ..., n$ we have $Ax^{(\ell)} = \lambda x^{(\ell)} + x^{(\ell-1)}$ for each $\ell = k, ..., 2$, while $Ax^{(1)} = \lambda \cdot x^{(1)}$. This is due to the construction of the principal vector chain. Thus

$$A_{nn}\begin{pmatrix} | & & | \\ x^{(k)} & \dots & x^{(1)} \\ | & & | \end{pmatrix}_{nk} = \begin{pmatrix} | & & | & | \\ \lambda x^{(k)} + x^{(k-1)} & \dots & \lambda x^{(2)} + x^{(1)} & \lambda x^{(1)} \\ | & & | & | \end{pmatrix} =$$

$$= \begin{pmatrix} | & & | \\ x^{(k)} & \dots & x^{(1)} \\ | & & | \end{pmatrix}_{nk} \begin{pmatrix} \lambda & 0 & & & 0 \\ 1 & \ddots & & & \\ & \ddots & \ddots & & \\ & & \ddots & \lambda & 0 \\ 0 & & & 1 & \lambda \end{pmatrix}_{kk} .$$

The k by k matrix $\begin{pmatrix} \lambda & 0 & & & 0 \\ 1 & \ddots & & & \\ & \ddots & \ddots & & \\ & & \ddots & \lambda & 0 \\ 0 & & & 1 & \lambda \end{pmatrix}$ is called a (lower) **Jordan block** $J(\lambda, k)$ of size k for the eigenvalue λ.

Note that a Jordan block $J(\lambda, k)$ with $k > 1$ cannot be diagonalized: its single eigenvalue λ appears k–fold on the diagonal, yet $\text{rank}(J(\lambda, k) - \lambda I_k) = k - 1$, since the lower co-diagonal contains precisely $k-1$ pivot candidates 1. Thus $J(\lambda, k)$ has precisely one eigenvector for the repeated eigenvalue λ. If the single repeated eigenvalue λ of A has the index j and if there are $d_j = \dim P_j(\lambda) - \dim P_{j-1}(\lambda)$ linearly independent principal vectors $x_m^{(j)}$ of grade j for $m = 1, ..., d_j$, then we set

$$X := \begin{pmatrix} | & & | & & | & & | \\ x_1^{(j)} & \dots & x_1^{(1)} & \dots & x_{d_j}^{(j)} & \dots & x_{d_j}^{(1)} \\ | & & | & & | & & | \end{pmatrix}_{n, d_j \cdot j}$$

and note that

$$A_{nn}X = X \begin{pmatrix} J(\lambda, j) & & \\ & \ddots & \\ & & J(\lambda, j) \end{pmatrix}_{d_j \cdot j, d_j \cdot j} = X \ \text{diag}(J(\lambda, j)) . \qquad (14.3)$$

Here $\text{diag}(J(\lambda, j))$ is a block diagonal matrix with d_j diagonal Jordan blocks and $d_j \cdot j \le n$. If $d_j \cdot j < n$, then for one first lower level r, less than the maximal index level j for λ, the principal subspace $P_r(\lambda)$ contains more linearly independent principal vectors of grade r than those generated from the maximal grade principal vectors $x_1^{(j)}, ..., x_{d_j}^{(j)}$ through repeated vector iteration with $(A - \lambda I)$. The extra $s := d_r - d_{r+1}$ linearly independent principal vectors of precise grade r and their associated principal vector chains of length r can be appended to X. And for the augmented principal vector matrix $X = X_{n, d_j \cdot j + s \cdot r}$ we have

$$AX = X \text{diag} \left(\text{diag}(J(\lambda, j)), \ \text{diag}(J(\lambda, r)) \right) .$$

Here the first diagonal block diag$(J(\lambda, j))$ contains d_j Jordan blocks for λ of size j according to (14.3) , and the second diagonal block, denoted by diag$(J(\lambda, r))$ above, contains $s = d_r - d_{r+1}$ Jordan blocks of size r for λ. If we collect all such lower grade maximal principal vectors and their principal vector chains in X, then eventually X will be n by n and $AX = XJ$, where J_{nn} is the **Jordan normal form** of the matrix A_{nn} with just one eigenvalue λ. This Jordan normal form consists of various Jordan blocks for λ on its block diagonal. The sizes of the Jordan blocks derive from the dimension differences between $P_k(\lambda)$ and $P_{k-1}(\lambda)$ for each $k = j,\ j-1,\ ....,\ 2$ if the single repeated eigenvalue λ of A has index j. The matrix X that transforms A to its Jordan normal form by similarity contains the vectors of a **Jordan basis** for A in its columns.

Theorem 14.1: (**Jordan Normal Form** (for a matrix with one eigenvalue))

Each matrix $A \in \mathbb{R}^{n,n}$ (or $\mathbb{C}^{n,n}$) with a single eigenvalue λ is similar over $\mathbb{R}$ (or $\mathbb{C}$, if $\lambda \notin \mathbb{R}$) to a block diagonal matrix $J =$ diag$\,(J(\lambda, n_{s_i}))$ of Jordan blocks $J(\lambda, n_{s_i})$. Here the dimensions n_{s_i} are equal to the maximal grades of principal vectors of A for λ that form a basis of $\mathbb{R}^n$ (or $\mathbb{C}^n$), called the **Jordan basis** of A.
In other words, the various dimensions n_{s_i} of Jordan blocks in the Jordan normal form of A are equal to the lengths of all complete length principal vector chains for A and λ.
The matrix J is called the **Jordan normal form** of A. ◀

The Jordan form of any matrix is generally not unique since a rearrangement of the principal vector chains in X rearranges the order of the individual Jordan blocks in the Jordan normal form J of A. However, the list of Jordan block dimensions $\{n_{s_i}\}$ is uniquely determined by A.

Example 1: Assume that a matrix $A_{14,14}$ has the single repeated eigenvalue λ. Assume further that the dimensions of the principal subspaces for A and λ are given as $p_1 =$ dim$E(\lambda) =$ dim$P_1(\lambda) = 5$, $p_2 =$ dim$P_2(\lambda) = 9$, $p_3 = 12$, and $p_4 = p_5 = 14$. Thus λ has index 4. By Theorem 14.1, our matrix A is similar to a block diagonal matrix with diagonal Jordan blocks for λ of sizes 4 by 4 or smaller for λ.

Which Jordan block sizes do occur here, given the above data?

Recall that $p_k = dim(P_k(\lambda))$ counts the total number of linearly independent vectors in $P_k(\lambda) = \bigcup_{i=1}^{k} P_i(\lambda)$. $P_k(\lambda)$ contains the principal vectors for A and λ of all grades $i \leq k$. Therefore this measure is too broad to be of immediate help with deciding the Jordan block structure of A.

However, $d_i = p_i - p_{i-1}$ counts the number of linearly independent principal vectors of precise grade i for A for each $i = 1, ..., k$, where k denotes the index of the eigenvalue λ of A. Thus if $d_i > d_{i+1}$, then there are $d_i - d_{i+1} > 0$ linearly independemt principal vectors of precise grade i for A and λ, each of which gives rise to a separate principal vector chain of order i that is linearly independent from all vectors of the higher order principal vector chains due to Lemma 3.

Note finally that for the eigenvalue λ of A of index k we have $d_{k+1} = 0$ since $p_k = p_{k-1}$ in this case.

Back to our example: Since $d_5 = 0$ and $d_4 := p_4 - p_3 = 14 - 12 = 2$, there are two linearly independent principal vector chains of maximal length 4 for A and λ. These result in 2 copies of $J(\lambda, 4)$ occuring in the Jordan normal form J of A. Next, $d_3 := p_3 - p_2 = 12 - 9 = 3$ gives A three linearly independent principal vectors of grade 3. Two of these are accounted for inside the two principal vector chains of order 4. Thus the third linearly independent principal vector of grade 3 gives rise to one separate Jordan block $J(\lambda, 3)$ of size 3 in J. Since $d_2 := p_2 - p_1 = 9 - 5 = 4$, there is one additional Jordan block $J(\lambda, 2)$ of size 2 by 2 in J, because only three grade 2 principal vectors are associated with the previously found larger dimensional Jordan blocks $J(\lambda, 4)$ (twice) and $J(\lambda, 3)$. Finally $d_1 := p_1 = 5$ counts the number of linearly independent eigenvectors for λ. Four of these occur in the two principal vector chains of length 4, in the one of length 3, and in the one of length 2. Thus there is one additional 1 by 1 Jordan block in J. And A is similar to its Jordan normal form

$$J_{14,14} = \operatorname{diag}(J(\lambda,4),\ J(\lambda,4),\ J(\lambda,3),\ J(\lambda,2),\ J(\lambda,1)) .$$

◀

For one eigenvalue λ the set of dimensions of its associated Jordan blocks can be illustrated in **diagram** form.

Example 2: We display the linearly independent principal vector chains from the previous example by the following dot diagram:

							no. of dots:
$P_5(\lambda)$:	no principal vectors of grade 5 :						$d_5 = 0$ dots
In $P_4(\lambda)$:	2 l. i. principal vectors of grade 4 :	•	•				$d_4 = 2$ dots
In $P_3(\lambda)$:	3 l. i. principal vectors of grade 3 :	•	•	•			$d_3 = 3$ dots
In $P_2(\lambda)$:	4 l. i. principal vectors of grade 2 :	•	•	•	•		$d_2 = 4$ dots
In $P_1(\lambda)$:	5 l. i. eigenvectors :	•	•	•	•	•	$d_1 = 5$ dots
	principal vector chain lengths :	4	4	3	2	1	$\sum_i d_i = 14$

Here the number of dots in each row is equal to the number of respective $d_{\dots} = p_{\dots} - p_{\dots-1}$, while the number of dots in each column describes the size of each Jordan block in the Jordan normal form of A. And the total number of dots (14) in the diagram equals the size of $A_{14,14}$

In this example we have $5 = p_1$ principal vector chains altogether, of lengths 4, 4, 3, 2, and 1 that give rise to Jordan blocks of the same dimensions as their principal vector chain lengths. Note that the sequence $\{d_i\} = \{5,\ 4,\ 3,\ 2\}$ of the numbers $d_i = p_i - p_{i-1}$ for $i = 1, \dots, k$ of linearly independent principal vectors of grade i is strictly decreasing in our example, namely $d_1 = 5 > d_2 = 4 > d_3 = 3 > d_4 = 2 >$

$d_5 = 0$.
For general matrices A and one eigenvalue λ of index j, we always have $d_i \geq d_{i+1}$ for $i \leq j$, the index of the eigenvalue λ of A. And equality $d_i = d_{i+1}$ holds for $i \leq j$ if there is no Jordan block of dimension i in the Jordan normal form of A. ◀

Example 3: (a) The sequence $p_1 = 4,\ p_2 = 8,\ p_3 = 9$, and $p_4 = p_5 = 10$ of principal subspace dimensions $p_i = \dim P_i(\lambda)$ does occur for a matrix A with a single repeated eigenvalue λ. The d_k dot diagram that we have introduced in Example 2 now looks as follows for $d_1 = p_1 = 4,\ d_2 = p_2 - p_1 = 4,\ d_3 = p_3 - p_2 = 1,\ d_4 = p_4 - p_3 = 1$:

					the d_i:
grade 4 :	•				$d_4 = 1$
grade 3 :	•				$d_3 = 1$
grade 2 :	•	•	•	•	$d_2 = 4$
grade 1 :	•	•	•	•	$d_1 = 4$
Jordan block sizes :	4	2	2	2	$\sum_i d_i = 10$

This tableau with 10 dots gives rise to the Jordan normal form

$$J = \text{diag}(J(\lambda, 4), J(\lambda, 2), J(\lambda, 2), J(\lambda, 2))$$

of size 10 by 10 for A. Note that the Jordan normal form for A contains no Jordan blocks of size 3 since $d_3 = d_4 = 1$, nor does it contain any of size 1 since $d_1 = d_2 = 4$.

(b) The sequence $p_1 = 3,\ p_2 = 7,\ p_3 = 9$, and $p_4 = p_5 = 10$ is not a legitimate principal subspace dimension sequence for any matrix A with one eigenvalue λ. To see this, we look at the dimension difference sequence $d_i = p_i - p_{i-1}$. For our data $d_1 = p_1 = 3 < d_2 = p_2 - p_1 = 4 > d_3 = p_3 - p_2 = 2 > d_4 = p_4 - p_3 = 1 > d_5 = 0$. Clearly the necessary inequality $d_i \geq d_{i+1}$ is not satisfied for $i = 1$. This contradicts the fact that the number $d_2 = 4$ of linearly independent principal vectors of level 2 must not be less than the number of linearly independent eigenvectors in $P_1(\lambda) = E(\lambda)$. And the d_i Jordan dot diagram does not look appropriate at all:

					the d_i:
grade 4 :	•				$d_4 = 1$
grade 3 :	•	•			$d_3 = 2$
grade 2 :	•	•	•	•	$d_2 = 4$
grade 1 :	•	•	•	?	$d_1 = 3$
Jordan block sizes :	4	2	2	?	

◀

Corollary 1: A sequence $\{p_i\}$ of principal subspace dimensions is admissible for some matrix A with one eigenvalue λ if and only if the numbers $d_i := p_i - p_{i-1} \geq 1$ satisfy $d_i \geq d_{i+1}$ for all $i = 1, \ldots, \text{index}(\lambda)$.

Specifically if $d_i > d_{i+1}$ for one $i \leq \text{index}(\lambda) - 1$, then the Jordan normal form of A contains $d_i - d_{i+1}$ Jordan blocks of size i by i for λ. ◀

Remark 1: According to Chapter 11, every real symmetric matrix $A = A^T$ has a diagonal Jordan normal form for real eigenvalues. In particular, it follows for $A = A^T \in \mathbb{R}^{n,n}$ that $\ker(A - \lambda I) = \ker(A - \lambda I)^k$ for all $k > 1$ and all $\lambda \in \mathbb{R}$. ◀

(b) *A non-diagonalizable matrix A with several distinct eigenvalues*

The ideas and results of subsection (a) can be applied to matrices with several eigenvalues. If a sparse normal form N can be obtained for A via matrix similarity $X^{-1}AX = N$, then this form can be achieved for every matrix B that is similar to A: In particular, if A and B are two similar matrices, i.e., if $Y^{-1}AY = B$ for a nonsingular matrix Y and if $X^{-1}AX = N$ is a normal form of A, then with $Z := Y^{-1}X$ and $Z^{-1} = X^{-1}Y$ according to Theorem 6.4, we have

$$Z^{-1}BZ \;=\; X^{-1}YBY^{-1}X \;=\; X^{-1}AX \;=\; N \quad .$$

Therefore $B = Y^{-1}AY$ is also similar to the normal form N of A.

Thus, when looking for sparse and structure revealing matrix representations of a given matrix A_{nn} under basis changes $X^{-1}..X$, we may start with any matrix B that is similar to A. Its normal form will be identical to the one for A. For the multi–eigenvalue and non diagonalizable case, we shall start with the Schur normal form S of A; see Theorem 11.1. S is upper triangular and has some repeated entries on its diagonal since a non diagonalizable matrix A must have repeated eigenvalues. We assume further without loss of generality that the distinct eigenvalues λ_1, ..., λ_k of S are arranged in consecutive diagonal entry clusters as depicted below

$$S \;=\; \begin{pmatrix}
\lambda_1 & * & & & & & & & & & * \\
0 & \ddots & & & & & & & & & \\
 & & \lambda_1 & & & & * & & & & \\
 & & & \lambda_2 & & & & & & & \\
 & & & & \ddots & & & & & & \\
 & & & & & \lambda_2 & & & & & \\
 & & & & & & \ddots & & & * & \\
 & & & & & & & \ddots & & & \\
 & & & & & & & & \lambda_k & & \\
 & & & & & & & & & \ddots & * \\
0 & & & & & & & & & 0 & \lambda_k
\end{pmatrix}$$

with $\lambda_i \neq \lambda_j$ for $i \neq j$. Note that this Schur form can be achieved by choosing $\lambda_2 = \lambda_1$ in the second step of the proof of Theorem 11.1 in case λ_1 is repeated and by further such repetitions for each repeated eigenvalue.

Our main task now is to use matrix similarity to reduce the Schur normal form S of A with its repeated eigenvalues appearing in clusters on its diagonal to a block diagonal matrix of the form

$$\widehat{S} \;=\; \begin{pmatrix} \boxed{\begin{matrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_1 \end{matrix}} & 0 & 0 \\ & \boxed{\begin{matrix} \lambda_2 & & * \\ & \ddots & \\ 0 & & \lambda_2 \end{matrix}} & 0 \\ & & \ddots \\ 0 & & & \ddots \end{pmatrix}$$

with upper triangular blocks for distinct eigenvalues on the diagonal and zero blocks in all of its off-diagonal positions.

If A_{nn} has only two distinct eigenvalues $\lambda_1 \neq \lambda_2$ of algebraic multiplicities $r > 1$ and $s \geq 1$, respectively, then $r + s = n$ and its Schur normal form can be written as

$$S \;=\; \begin{pmatrix} \boxed{\begin{matrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_1 \end{matrix}} & * \\ 0 & \boxed{\begin{matrix} \lambda_2 & & * \\ & \ddots & \\ 0 & & \lambda_2 \end{matrix}} \end{pmatrix} ,$$

where the first diagonal block is r by r, and the second one is s by s. We shall perform matrix similarities on S using Gaussian elimination matrices $X_{i,j}(\alpha) := (\alpha I_n + \alpha E_{i,j})$ of the form (6.3) for indices $1 \leq i \leq r$ and $r+1 \leq j \leq r+s = n$ and reduce S to a matrix of type $\widehat{S}$ with zeros in its (1,2) off-diagonal block.

Here $E_{i,j}$ is the elementary n by n matrix with zeros in each of its n^2 positions, except for a 1 in position (i,j). As stated in Lemma 2 of Section 6.2, $(X_{i,j}(\alpha))^{-1} = X_{i,j}(-\alpha)$ for all i, j, and α. We start the process with α in position $i = r$ and $j = r+1$ and form $(X_{r,r+1}(\alpha))^{-1}\, S X_{r,r+1}(\alpha) \;=$

$$\left(\begin{array}{c|c} I_r & \begin{matrix} & 0 \\ -\alpha & \end{matrix} \\ \hline O_{s,r} & I_s \end{array}\right) \cdot \begin{pmatrix} \boxed{\begin{matrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_1 \end{matrix}} & \begin{matrix} & * \\ s_{r,r+1} & \end{matrix} \\ O_{r,s} & \boxed{\begin{matrix} \lambda_2 & & * \\ & \ddots & \\ 0 & & \lambda_2 \end{matrix}} \end{pmatrix} \cdot \left(\begin{array}{c|c} I_r & \begin{matrix} & 0 \\ \alpha & \end{matrix} \\ \hline O_{s,r} & I_s \end{array}\right) =$$

$$
= \left(\begin{array}{c|c} I_r & \begin{matrix} & 0 \\ -\alpha & \end{matrix} \\ \hline O_{s,r} & I_s \end{array}\right) \cdot \left(\begin{array}{cc} \boxed{\begin{matrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_1 \end{matrix}} & \begin{matrix} \dagger & & * \\ \dagger & & \\ \alpha\lambda_1 + s_{r,r+1} & & \end{matrix} \\ O_{s,r} & \boxed{\begin{matrix} \lambda_2 & & * \\ & \ddots & \\ 0 & & \lambda_2 \end{matrix}} \end{array}\right) =
$$

$$
= \left(\begin{array}{cc} \boxed{\begin{matrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_1 \end{matrix}} & \begin{matrix} \dagger & & * \\ \dagger & & \\ \alpha\lambda_1 + s_{r,r+1} - \alpha\lambda_2 & \ddagger & \ddagger \end{matrix} \\ O_{r,s} & \boxed{\begin{matrix} \lambda_2 & & * \\ & \ddots & \\ 0 & & \lambda_2 \end{matrix}} \end{array}\right) = \widetilde{S}\ .
$$

Here we have formed the product of the two matrix factors S and $X_{r,r+1}(\alpha)$ on the right first by updating the first $r-1$ entries in the $(r+1)^{st}$ column, denoted by $\dagger$, and by updating the $(r, r+1)$ entry $s_{r,r+1}$ of S to become $\alpha\lambda_1 + s_{r,r+1}$. Multiplying the resulting product matrix on the left by $(X_{r,r+1}(\alpha))^{-1} = X_{r,r+1}(-\alpha)$ updates the $(r, r+1)$ entry further to $\widetilde{s}_{r,r+1} := \alpha\lambda_1 + s_{r,r+1} - \alpha\lambda_2$ and changes the last $s-1$ entries of the r^{th} row. These updates are denoted by $\ddagger$ in $\widetilde{S}$. Note that no other entries of the original Schur form S are affected by the similarity.

If we choose $\alpha := \dfrac{s_{r,r+1}}{\lambda_2 - \lambda_1}$, then the $(r, r+1)$ entry $\alpha(\lambda_1 - \lambda_2) + s_{r,r+1}$ in $\widetilde{S} := (X_{r,r+1}(\alpha))^{-1}\, S X_{r,r+1}(\alpha)$ becomes zero.

If $s > 1$ we repeat this process with the intend to zero out the $(r, r+2)$ entry of $\widetilde{S}$ via a matrix similarity $X_{r,r+2}(-\alpha)\ \widetilde{S}\ X_{r,r+2}(\alpha)$ for a different value of α in the $(r, r+2)$ position of $X_{r,r+2}(\alpha)$: $X_{r,r+2}(-\alpha)\ \widetilde{S}\ X_{r,r+2}(\alpha)\ =$

$$
\left(\begin{array}{c|c} I_r & \begin{matrix} 0 & & 0 \\ & & \\ 0 & -\alpha & 0 \end{matrix} \\ \hline O_{s,r} & I_s \end{array}\right) \cdot \left(\begin{array}{cc} \boxed{\begin{matrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_1 \end{matrix}} & \begin{matrix} * & * & * \\ * & * & * \\ 0 & \widetilde{s}_{r,r+2} & * \end{matrix} \\ O_{s,r} & \boxed{\begin{matrix} \lambda_2 & & * \\ & \ddots & \\ 0 & & \lambda_2 \end{matrix}} \end{array}\right) \cdot \left(\begin{array}{c|c} I_r & \begin{matrix} 0 & & 0 \\ & & \\ 0 & \alpha & 0 \end{matrix} \\ \hline O_{s,r} & I_s \end{array}\right) =
$$

$$= \left(\begin{array}{c|ccc} & 0 & & 0 \\ I_r & & & \\ & 0 & -\alpha & 0 \\ \hline O_{s,r} & & I_s & \end{array} \right) \cdot \left(\begin{array}{ccc} \boxed{\begin{matrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_1 \end{matrix}} & \begin{matrix} * & \dagger & * \\ * & \dagger & * \\ 0 & \alpha\lambda_1 + \widetilde{s}_{r,r+2} & * \end{matrix} \\ O_{s,r} & \boxed{\begin{matrix} \lambda_2 & & * \\ & \ddots & \\ 0 & & \lambda_2 \end{matrix}} \end{array} \right) =$$

$$= \left(\begin{array}{cc} \boxed{\begin{matrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_1 \end{matrix}} & \begin{matrix} * & \dagger & * \\ * & \dagger & * \\ 0 & \alpha\lambda_1 + \widetilde{s}_{r,r+2} - \alpha\lambda_2 & \ddagger \end{matrix} \\ O_{s,r} & \boxed{\begin{matrix} \lambda_2 & & * \\ & \ddots & \\ 0 & & \lambda_2 \end{matrix}} \end{array} \right) = \widetilde{\widetilde{S}} \; .$$

Here the only entries in $\widetilde{\widetilde{S}}$ that differ from those in $\widetilde{S}$ lie above and to the right of the $(r, r+2)$ position while the $(r, r+2)$ entry changes to $\widetilde{\widetilde{s}}_{r,r+2} = \alpha\lambda_1 + \widetilde{s}_{r,r+2} - \alpha\lambda_2$. If we pick $\alpha := \dfrac{s_{r,r+2}}{\lambda_2 - \lambda_1}$ in this step, then the $(r, r+2)$ entry in $\widetilde{\widetilde{S}} = (X_{r,r+2}(\alpha))^{-1}\, \widetilde{S} X_{r,r+2}(\alpha)$ becomes zero.

Note the pattern here: We have chosen specific constants α and elementary matrices $X_{i,j}(\alpha)$ as similarity transformations for the Schur normal form S in sequence for $i \leq r$ and $r+1 \leq j \leq r+s$. Namely we start at the lower left corner position $(r, r+1)$ in the $(1, 2)$ block of S and zero out the $(r, r+1)$ entry in S while updating the entries of S that lie in the $r+1^{st}$ column above position $(r, r+1)$ and the entries in the r^{th} row of S to the right of position $(r, r+1)$ via a matrix similarity with $X_{r,r+1}(\alpha)$.
Next we repeat the process for position $(r, r+2)$ with a similarity by $X_{r,r+2}(\beta)$ that does not affect the zero created in position $(r, r+1)$. In this way we can sweep each row from left to right in the (1,2) block, from the last, the r^{th} row on up to the first, each time using appropriate elementary matrix similarities $X_{i,j}(\alpha_{i,j})$ for $r \leq i \leq 1$ and $r+1 \leq j \leq r+s$. And thus we achieve the block diagonal form $\widehat{S}$.

This process readily generalizes to Schur normal forms S of matrices A with several distinct eigenvalues. Combining this with Theorem 14.1 proves our main result:

Theorem 14.2: (**The Jordan Normal Form**)

Every matrix $A_{nn} \in \mathbb{R}^{n,n}$ (or $\mathbb{C}^{n,n}$) is similar over the complex numbers $\mathbb{C}$ to a block diagonal matrix consisting of certain Jordan blocks for its eigenvalues.
In other words, for each A_{nn} there exists a Jordan basis or a nested principal vector basis of $\mathbb{C}^n$ collected in the columns of X_{nn}, so that

$$\begin{aligned} X^{-1}AX \;=\;& \operatorname{diag}\left(J(\lambda_1, n_{1,1}), ..., J(\lambda_1, n_{1,\ell_1}), J(\lambda_2, n_{2,1}), ..., J(\lambda_2, n_{2,\ell_2}), ...,\right. \\ & \left. ..., J(\lambda_k, n_{k,1}), ..., J(\lambda_k, n_{k,\ell_k})\right) \ . \end{aligned}$$

Here $n_{j,i} \geq 1$, $\ell_j \geq 1$, and $\sum_{j=1}^{k} \sum_{i=1}^{\ell_j} n_{j,i} = n$. Each sequence of Jordan block dimensions $n_{m,1}, ..., n_{m,\ell_m}$ is uniquely determined for the eigenvalue $\lambda_m \in \mathbb{C}$ of A.

To prove this theorem, we simply invoke subsection (a) on each of the diagonal Schur blocks $\begin{pmatrix} \lambda_i & & * \\ & \ddots & \\ 0 & & \lambda_i \end{pmatrix}$ that has been obtained via elementary similarities from the Schur normal form S of A in $\widehat{S}$.

Remark 2: The Jordan normal form of a matrix is generally not unique since the order of the eigenvalues can be changed at will, as can the order of the sequence of Jordan block dimensions $n_{j,i}$ for each repeated eigenvalue λ_j of A.
The Jordan normal form of a matrix is generally not diagonal. It contains one or several non zero co–diagonal entries of 1 in general.
The Jordan normal form represents A sparsely with respect to a principal vector chain basis, called the Jordan basis of A.
If A is diagonalizable, then the Jordan normal form of A is diagonal and A has an eigenvector basis, and vice versa. ◀

Example 4: Find the Jordan normal form J for $A = \begin{pmatrix} 1 & 0 & -1 & -1 \\ 1 & 2 & 1 & 1 \\ 0 & 0 & 2 & 0 \\ 1 & 0 & 1 & 3 \end{pmatrix}$.

Starting with $x = \begin{pmatrix} -2 \\ 1 \\ -1 \\ 2 \end{pmatrix}$ for example, we compute the vector iteration matrix

$$\begin{pmatrix} | & | & & | \\ x & Ax & \dots & A^4x \\ | & | & & | \end{pmatrix} = \begin{pmatrix} -2 & -3 & -4 & -4 & 0 \\ 1 & 1 & 0 & -4 & -16 \\ -1 & -2 & -4 & -8 & -16 \\ 2 & 3 & 4 & 4 & 0 \end{pmatrix}$$

and its RREF as $\begin{pmatrix} 1 & 0 & -4 & -16 & -48 \\ 0 & 1 & 4 & 12 & 32 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$. Thus $A^2x = 4Ax - 4x$ and the minimal polynomial of A contains the factor $p(x) = x^2 - 4x + 4 = (x-2)^2$ that we have read off the first linearly dependent column, the third column, in the RREF. As the minimal polynomial of A contains a quadratic factor, A is not diagonalizable according to Chapter 9.

To find the Jordan block structure of A, we analyze $\ker(A - 2I_4)^2$ next:

Note that $A - 2I = \begin{pmatrix} -1 & 0 & -1 & -1 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 \end{pmatrix}$ and $(A-2I)^2 = O_2$. Thus $p(x) = (x-2)^2 = p_A(x)$ is indeed A's minimal polynomial and $p_2 = \dim P_2(2) = \dim(\ker(A-2I)^2) = 4$ while $p_1 = \dim P_1(2) = \dim E(2) = \dim(\ker(A-2I)) = 3$. This makes $d_1 = p_1 = 3 > d_2 = p_2 - p_1 = 4 - 3 = 1$. Thus the Jordan normal form J of A has $1 = d_2$ Jordan block $J(2,2)$ of size two and $2 = d_1 - d_2$ Jordan blocks $J(2,1)$ of size one, all for the single repeated eigenvalue $\lambda = 2$ of A. The Jordan normal form of A is

$$J = \operatorname{diag}(J(2,2),\ J(2,1),\ J(2,1)) \ .$$

Incidentally, the Jordan normal form of our matrix $A_{4,4}$ could have been deduced by a counting argument alone: $A_{4,4}$ has the single eigenvalue $\lambda = 2$ with geometric multiplicity 3; just look at $A - 2I$ above which has rank 1. Hence there must be precisely one principal vector for $\lambda = 2$ of index 2 in order that the Jordan dot diagram has 4 dots:

grade 2 :	$\underline{\bullet}$!		
grade 1 :	$\bullet$	$\bullet$	$\bullet$
Jordan block sizes :	2	1	1

Or alternatively, one could argue differently again: The eigenvalue 2 of A has the index 2 since $p_A(x) = (x-2)^2$. Therefore the matrix A of size 4 could possibly have a Jordan normal form consisting of 2 Jordan blocks of size 2 for its single eigenvalue, or its Jordan form could have one Jordan block of size 2 and two more 1–dimensional Jordan blocks, all for $\lambda = 2$. The eigenspace $E(2)$ for A has dimension 3 since $\operatorname{rank}(A - 2I) = 1$. Hence the latter must be the Jordan normal form of A, since in the former case A could only have two linearly independent eigenvectors. ◀

(c) *Practicalities*

For ease with developing the Jordan normal form theory, we have so far introduced principal vector chains $x^{(j)}$, $x^{(j-1)} = (A - \lambda I)x^{(j)}$, ... from the top grade principal vector(s) $x^{(j)}$ of A_{nn} on down. This assumes implicitly that we know the index j of an eigenvalue λ of A and the sequence of principal subspace dimensions $p_i = \dim P_i(\lambda)$ beforehand. There is an easier way than to construct all nested kernels $E(\lambda) = P_1(\lambda) \subset P_2(\lambda) \subset ...$ and to compute their dimensions p_i explicitly. We can start instead with an eigenvector basis for an eigenvalue λ of A. From the Jordan dot diagrams, we know that there are precisely $\dim E(\lambda)$ many Jordan blocks for λ in the Jordan normal form of A. What we have to ascertain is their dimensions, i.e., the length of each principal vector chain for λ from the eigenspace generators of A themselves.

If $x^{(1)}$ is a known eigenvector for A and λ, and if $x^{(2)}$ is a solution of the singular inhomogeneous linear system $(A - \lambda I)x = x^{(1)}$, then $x^{(2)}$ is a principal vector of grade 2 for A and λ. Consequently the principal vector chain that has $x^{(1)}$ as its eigenvector can be reconstructed from $x^{(1)}$ by solving the linear systems $(A - \lambda I)x = x^{(i)}$ successively for $x =: x^{(i+1)}$ and $i = 1, 2, ...$, until the linear system becomes inconsistent and thus has no solution. This allows us to find the index of an eigenvalue λ and the length of its principal vector chains, as well as the sizes of its Jordan blocks by linear means from λ and a known eigenvector. We have to choose the generating eigenvectors $x^{(1)}$ carefully in this process, though, so that we obtain useful principal vector chains.
The selection process will be made clear in Example 5. This method simplifies finding a similarity transformation matrix X that sends A to its Jordan normal form.

Example 5: We use the matrix $A = \begin{pmatrix} 1 & 0 & -1 & -1 \\ 1 & 2 & 1 & 1 \\ 0 & 0 & 2 & 0 \\ 1 & 0 & 1 & 3 \end{pmatrix}$ of Example 4 and try to find its principal vector chains from its eigenvectors.
For the eigenvalue $\lambda = 2$ of A we first compute a basis for the kernel of $A - \lambda I$:

$$\ker(A - 2I_4) \ : \quad \begin{array}{rrrr|ll} -1 & 0 & -1 & -1 & 0 & \\ 1 & 0 & 1 & 1 & 0 & + row_1 \\ 0 & 0 & 0 & 0 & 0 & \\ 1 & 0 & 1 & 1 & 0 & + row_1 \\ \hline -1 & 0 & -1 & -1 & 0 & \\ 0 & 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & 0 & \end{array}$$

Thus the three vectors $x_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$, $x_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}$, and $x_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}$ form a basis for the eigenspace for $\lambda = 2$ and A. Note that neither of these three eigenvectors leads to a solvable linear system of the form $(A-2I)x = x_i$ for $i = 1, 2, 3$. This should be checked by the students.

Next we check whether there is a linear combination $y := \alpha x_1 + \beta x_2 + \gamma x_3 = \begin{pmatrix} \beta+\gamma \\ \alpha \\ -\beta \\ -\gamma \end{pmatrix} \in E(2) = P_1(2)$ of the three eigenvectors for which $(A-2I)x = y$ is solvable.

$$\begin{array}{cccc|cl} -1 & 0 & -1 & -1 & \beta+\gamma & \\ 1 & 0 & 1 & 1 & \alpha & \quad + row_1 \\ 0 & 0 & 0 & 0 & -\beta & \\ 1 & 0 & 1 & 1 & -\gamma & \quad + row_1 \\ \hline -1 & 0 & -1 & -1 & \beta+\gamma & \\ 0 & 0 & 0 & 0 & \alpha+\beta+\gamma & \\ 0 & 0 & 0 & 0 & -\beta & \\ 0 & 0 & 0 & 0 & \beta & \end{array}$$

This system is solvable precisely when $\beta = 0$ and $\alpha = -\gamma$. If we set $\alpha := 1$ then $x^{(1)} := x_1 - x_3 = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \end{pmatrix}$ is an eigenvector that belongs to a principal vector chain of order at least 2 for A and $\lambda = 2$. The corresponding principal vector $x^{(2)}$ of grade 2 solves $(A-2I)x = x^{(1)}$. We find $x^{(2)}$ from the scheme

$$\begin{array}{cccc|cl} -1 & 0 & -1 & -1 & -1 & \\ 1 & 0 & 1 & 1 & 1 & \quad + row_1 \\ 0 & 0 & 0 & 0 & 0 & \\ 1 & 0 & 1 & 1 & 1 & \quad + row_1 \\ \hline -1 & 0 & -1 & -1 & -1 & \\ 0 & 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & 0 & \end{array}$$

as $x^{(2)} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ for example. Now, the next level linear system $(A-2I)x = x^{(2)} = e_1$ is inconsistent, because its first and second row equations are contradictory. Therefore there are no higher grade principal vectors than of grade 2 for the eigenvalue $\lambda = 2$.

If we write the principal vector matrix X in the form $X = \begin{pmatrix} | & | & | & | \\ x^{(2)} & x^{(1)} & x_1 & x_2 \\ | & | & | & | \end{pmatrix}$, by using the first found eigenvectors x_1 and x_2 in columns 3 and 4 and the principal vector chain $x^{(2)}$, $x^{(1)}$ in columns 1 and 2 of X for example, then

$$X^{-1}AX = J = \text{diag}(J(2,2), J(1,1), J(1,1)) .$$

And we have found an explicit matrix similarity that transforms A to its Jordan normal form J.

◀

We have generalized the concept of diagonalizability of a matrix and have obtained a sparse normal form representation of matrices that are not diagonalizable.

14.1.P Problems

1. Write down a 5 by 5 matrix that is not diagonalizable.

2. Write down a 7 by 7 matrix that is not diagonal but diagonalizable.

3. Find a 12 by 12 triangular matrix that is diagonalizable, and one that is not.

4. Find the Jordan normal form of $\begin{pmatrix} -4 & 4 \\ -1 & 0 \end{pmatrix}$.

5. Find the Jordan normal form of $A = \begin{pmatrix} 0 & 0 & 1 \\ -3 & -1 & -3 \\ -1 & 0 & -2 \end{pmatrix}$, given that -1 is an eigenvalue of A. (Hint: What is the algebraic multiplicity of -1?)

6. Find the Jordan normal form of $B = \begin{pmatrix} -3 & -7 & -3 \\ 2 & 5 & 2 \\ -1 & -3 & -1 \end{pmatrix}$, given that 1 and 0 are eigenvalues of B.

7. Write out the Jordan dot diagram for A in Example 4 and for Problem 5.

8. Decide whether the matrix $A = \begin{pmatrix} -4 & 9 \\ -1 & 2 \end{pmatrix}$ is diagonalizable or not.
   What is its Jordan normal form?

9. Show that the real matrix $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ is similar to the complex matrix $\begin{pmatrix} \lambda & 0 \\ 0 & \overline{\lambda} \end{pmatrix}$ where $\lambda = a + bi \in \mathbb{C}$.

10. Determine whether the two matrices $A = \begin{pmatrix} 2 & 5 & 0 & -2 \\ -3 & -6 & 0 & 2 \\ -7 & -12 & -1 & 5 \\ -3 & -5 & 0 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & -4 & -2 & 0 \\ -1 & 1 & 1 & 0 \\ -1 & -11 & -4 & 1 \\ 8 & -3 & -4 & -2 \end{pmatrix}$ are similar.
    (Note that both matrices have the 4 fold eigenvalue $\lambda = -1$.)

11. Find the Jordan normal form of $C = \begin{pmatrix} -6 & 12 & -22 \\ 2 & -2 & 5 \\ 3 & -5 & 10 \end{pmatrix}$, given that -1 and 2 are eigenvalues of C. What is the third eigenvalue of C? (Think 'trace'.)

12. Determine whether $A = \begin{pmatrix} 4 & -2 & -1 \\ -6 & 8 & 3 \\ 16 & -16 & -6 \end{pmatrix}$ has the Jordan normal form $J = \begin{pmatrix} 2 & & \\ & 2 & 0 \\ & 1 & 2 \end{pmatrix}$, $K = \begin{pmatrix} 2 & & \\ 1 & 2 & \\ 0 & 1 & 2 \end{pmatrix}$, or neither.
    What other Jordan normal form could A possibly have, given that it has the three fold eigenvalue $\lambda = 2$?

13. What are all possible Jordan diagrams for a matrix A with a single 4 fold real eigenvalue λ.

14. Repeat the previous problem for a matrix B with the 4 fold eigenvalue $\lambda = 4$ of
    (a) index 3;

(b) index 1;

(c) index 4.

15. Write down all possible Jordan block dimensions for a matrix $A_{n,n}$ with a single 6 fold eigenvalue λ.

16. For each of the possible Jordan forms in the previous 3 problems, determine the algebraic and the geometric multiplicities of the eigenvalue λ.

17. If A has one real eigenvalue with geometric multiplicity 3, algebraic multiplicity 8, and index 4, how many differing Jordan block dimensions can there be for this eigenvalue of A? What Jordan diagrams are possible for this data?

18. Repeat the previous problem for an eigenvalue λ with

(a) geometric multiplicity 4, algebraic multiplicity 8, and index 5;

(b) geometric multiplicity 6 and algebraic multiplicity 8;

(c) equal geometric multiplicity, algebraic multiplicity, and index, if possible.

19. Find the possible range of indices of an eigenvalue λ which has algebraic multiplicity 12 and geometric multiplicity 7.

20. Find a matrix $X_{5,5}$ with

$$X \begin{pmatrix} 2 & 1 & & & 0 \\ 0 & 2 & & & \\ & & 1 & 1 & 0 \\ & & 0 & 1 & 1 \\ 0 & & 0 & 0 & 1 \end{pmatrix} X^{-1} =$$

$$= \begin{pmatrix} 1 & 0 & 0 & & 0 \\ 1 & 1 & 0 & & \\ 0 & 1 & 1 & & \\ & & & 2 & 0 \\ 0 & & & 1 & 2 \end{pmatrix}.$$

21. (a) If the eigenvalue λ of A has an algebraic multiplicity exceeding its geometric one by 1, what Jordan blocks are possible for λ and A?

(b) If the eigenvalue λ of A has an algebraic multiplicity two larger than its geometric one, what Jordan blocks are possible for λ and A?

(c) If the algebraic multiplicity of the eigenvalue λ of A is 6 and if this exceeds its geometric multiplicity by 3, how many different Jordan structures are possible for λ and A, disregarding interchanges of Jordan blocks?

22. Show that for a symmetric matrix $A = A^T \in \mathbb{R}^{n,n}$, we have $\ker(A - \lambda I) = \ker(A - \lambda I)^k$ for all $\lambda \in \mathbb{R}$ and all $k \geq 1$.
(Hint: start with $k = 2$.)

23. (a) Show that the inverse of a nonsingular Jordan block $J(\lambda, k)$ has the form

$$\begin{pmatrix} 1/\lambda & 0 & & \dots & 0 \\ -1/\lambda^2 & 1/\lambda & & & \\ 1/\lambda^3 & -1/\lambda^2 & 1/\lambda & & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ (-1)^{k+1}/\lambda^k & \dots & 1/\lambda^3 & -1/\lambda^2 & 1/\lambda \end{pmatrix}.$$

(b) If $J = \operatorname{diag}(J_i)$ is the Jordan normal form of a nonsingular matrix A, what is the Jordan normal form of A^{-1}?

24. True or false:

(a) If J_{kk} is a Jordan block, then J^2 is a Jordan block.

(b) The matrix $A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ has the eigenvalue $\lambda = 1$.

(c) The matrix $B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 2 \end{pmatrix}$ has the eigenvalue $\lambda = 0$.

(d) The matrix A of part (b) satisfies $A^2 - 2I = O_3$.

(e) The matrix A of part (b) satisfies $A^2(A - 2I) = O_3$

(f) The matrix A of part (b) satisfies $A(A - 2I) = O_3$

(g) The matrix B of part (c) satisfies $B(B^2 - 3B + 2I) = O_3$

(h) The eigenvalue -1 of $J = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ has the algebraic multiplicity 2.

(i) The eigenvalue 3 of J in part (h) has the algebraic multiplicity 2.

(j) 0 is an eigenvalue of J in part (h).

(k) The matrix J in part (h) has the eigenvalues $\lambda = 3$ and $\mu = -1$, both of geometric multiplicity equal to 1.

(l) The matrix J in part (h) has the same minimal and characteristic polynomial.

(m) The matrix J in part (h) is diagonalizable over $\mathbb{C}$, but not over $\mathbb{R}$.

25. Show that $A = \begin{pmatrix} -4 & -4 \\ 9 & 8 \end{pmatrix}$ is not diagonalizable. What is the Jordan normal form of A?

26. Show that a matrix of the form $A = \begin{pmatrix} 1 & 0 & 0 \\ * & 2 & 0 \\ 1 & * & 1 \end{pmatrix}$ cannot be diagonalized, no matter what the entries marked by $*$ are.

27. (a) Show that for a fixed λ all matrices of the form $A = \begin{pmatrix} \lambda & \alpha \\ 0 & \lambda \end{pmatrix}$ are similar for every $\alpha \neq 0$. What if $\alpha = 0$?

(b) Show that for a fixed λ all matrices of the form $A = \begin{pmatrix} \lambda & \alpha_1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & \alpha_{k-1} \\ 0 & & & \lambda \end{pmatrix}$ are similar for any choice of α_i as long as $\prod \alpha_i \neq 0$. What if some $\alpha_j = 0$?

(c) Are the matrices $B = \begin{pmatrix} \lambda & 2 \\ 0 & \lambda \end{pmatrix}$ and $C = \begin{pmatrix} \lambda & 0 \\ -201 & \lambda \end{pmatrix}$ similar or not? State your reasons.

28. Find the Jordan normal form of $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

29. Let A be an n by n matrix whose k^{th} column is a multiple ae_k of the k^{th} unit vector e_k.

(a) Show that A has the eigenvalue a. What is the corresponding eigenvector?

(b) Give several examples of 4 by 4 matrices A with second column $\begin{pmatrix} 0 \\ -3 \\ 0 \\ 0 \end{pmatrix}$ that are diagonalizable and likewise examples that are not diagonalizable.

(c) In part (b), which entries of your example matrices A seem to determine whether A is diagonalizable or not?

30. (a) If A has the Jordan normal form $J = \text{diag}(J(\lambda,3), J(\lambda,2), J(\lambda,1), J(\mu,4), J(\mu,4), J(3,2))$ for $\lambda \neq \mu$, what are the dimensions of the eigenspaces $E(\lambda)$ and $E(\mu)$, depending on whether $\lambda = 3$, $\mu = 3$, or neither.

(b) What are the characteristic and minimal polynomials for J in each case of (a)? Repeat the question for A.

31. (a) Can a complex matrix $A \notin \mathbb{R}^{n,n}$ have both real and nonreal eigenvalues?

(b) Can a complex matrix $A \notin \mathbb{R}^{n,n}$ have only real eigenvalues?

(c) Can parts (a) or (b) hold for a 2 by 2 matrix A?

Hint: Try to construct a complex matrix $A_{2,2}$ with real trace and negative determinant, if possible.

32. What happens to the appearance of the Jordan block J if on p. W-5 we look at the product $A_{nn} \cdot \begin{pmatrix} | & | & & | \\ x^{(1)} & x^{(2)} & \dots & x^{(k)} \\ | & | & & | \end{pmatrix}_{nk}$ instead for the same principal vector chain for A and λ?

Teacher's Problem–Making Exercise

T 14. To construct dense integer matrices A_{nn} with integer principal vector chains and a predetermined Jordan structure is a charm: Start with an integer Jordan normal form matrix J_{nn}, composed of integer Jordan blocks J_ℓ on its block diagonal. The diagonal Jordan blocks J_ℓ of J can have the standard form $J_\ell = J(\lambda_\ell, k) = \begin{pmatrix} \lambda_\ell & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & \lambda_\ell \end{pmatrix}_{kk}$ for an integer λ_ℓ or the real Jordan normal form (see Section 14.2 below) $J_\ell = J(a_\ell, b_\ell, 2r) = \begin{pmatrix} a_\ell & b_\ell & & & & 0 \\ -b_\ell & a_\ell & & & & \\ 1 & 0 & a_\ell & b_\ell & & \\ 0 & 1 & -b_\ell & a_\ell & & \\ & & \ddots & & \ddots & \\ & 0 & & & \ddots & \ddots \end{pmatrix}_{2r,2r}$ for two integers a_ℓ, b_ℓ with $b_\ell \neq 0$. Next construct an n by n unimodular matrix X as in Teacher Problem T 6 of Section 6.1.P from two unit triangular matrices, one lower and the other one upper triangular, i.e., $X = RL$ or $X = LR$. Then the matrix $A := XJX^{-1}$ has integer entries. Moreover A has the Jordan normal form J for the eigenvalues $\lambda_\ell \in \mathbb{R}$ and $a_\ell + b_\ell i \in \mathbb{C}$ and integer principal vector chains contained in the respective columns of X.

Example : For $n = 5$ and the Jordan normal form $J = \operatorname{diag}\left(\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & & \\ 1 & 1 & \\ & 1 & 1 \end{pmatrix} \right)$ with eigenvalues 1 (3 fold) and $1 \pm 2i$, we use $X = RL =$

$$\begin{pmatrix} -1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & -1 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 2 & 1 & -1 & 0 & 0 \\ -1 & 1 & 1 & 1 & 0 \\ 2 & -1 & 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 0 & -1 & 1 & 1 \\ -1 & 4 & 0 & 0 & -1 \\ 0 & 2 & -1 & -1 & -1 \\ -1 & 0 & -1 & -2 & -1 \\ -2 & 1 & 0 & -1 & -1 \end{pmatrix}$$

and compute $A = XJX^{-1} = \begin{pmatrix} 0 & 5 & -4 & 4 & -6 \\ 7 & 1 & -8 & 1 & 14 \\ 5 & 1 & -5 & 2 & 8 \\ 3 & -2 & 0 & 0 & 6 \\ 3 & -3 & 0 & -2 & 9 \end{pmatrix}$. The dense integer matrix A has the same eigenvalues 1 (3 fold) and $1 \pm 2i$ as J. The corresponding eigenvectors are $\begin{pmatrix} 3 \\ -1 \\ 0 \\ -1 \\ -2 \end{pmatrix} + \begin{pmatrix} 0 \\ 4 \\ 2 \\ 0 \\ 1 \end{pmatrix} i$ for $\lambda_1 = 1 + 2i$ and $\begin{pmatrix} 3 \\ -1 \\ 0 \\ -1 \\ -2 \end{pmatrix} - \begin{pmatrix} 0 \\ 4 \\ 2 \\ 0 \\ 1 \end{pmatrix} i$ for $\lambda_2 = 1 - 2i$. The respective real and imaginary part vectors of these complex eigenvectors appear in columns 1 and 2 of X that correspond to the first diagonal block $\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$ in

J. The eigenvector $\begin{pmatrix} 1 \\ -1 \\ -1 \\ -1 \\ -1 \end{pmatrix}$ for $\lambda_3 = 1$ appears in the last column of X. For $\lambda_3 = 1$ with index equal to three, a complete principal vector chain $x^{(3)}$, $x^{(2)}$, $x^{(1)}$ appears in descending order in the columns 3, 4, and 5 of X.

If we form the matrix $B = XJ^TX^{-1} = \begin{pmatrix} 1 & -1 & -1 & 1 & 0 \\ -8 & 3 & 6 & 2 & -18 \\ -6 & -1 & 8 & -1 & -10 \\ -3 & -2 & 6 & -2 & -2 \\ -3 & 1 & 3 & 0 & -5 \end{pmatrix}$ instead, then this matrix has the same eigenvalues as A and as J and its Jordan normal form has the same Jordan structure. The corresponding eigenvectors and principal vector chains are a bit harder to find from the construction data: We have $X^{-1}BX = J^T$ and $EJ^TE = J$ where $E = \text{diag}(E_i)$ is conformally partitioned as $J = \text{diag}(J_i)$ and each $E_i = \begin{pmatrix} & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & 0 \end{pmatrix} = E_i^{-1}$. – [The students might want to check the last assertion on transforming a block matrix of transposed upper Jordan blocks $J^T = \text{diag}(J_i^T)$ to its lower block form $J = \text{diag}(J_i)$.] – Thus

$$\text{diag}(E_i)X^{-1}BX\text{diag}(E_i) \;=\; E^{-1}J^TE \;=\; J\;.$$

And the eigenvectors and principal vectors of B appear in the respective columns of X $\text{diag}(E_i)$ instead. ◀

14.2 Theory (The Real Jordan Normal Form, the Companion Matrix Normal Form, and Symmetric Matrix Products)

We study normal forms that can be achieved over the reals for non diagonalizable real matrices, as well as symmetric matrix products.

According to Chapters 7 through 12, a matrix normal form $N = X^{-1}AX$ represents the standard matrix $A = A_{\mathcal{E}}$ of a linear transformation with respect to a particularly chosen basis $\mathcal{X}$. Normal forms are designed to reveal certain aspects of the linear transformation $x \mapsto Ax$. We have described several normal forms for matrices, such as the diagonal form under similarity in Sections 9.1, 9.1.D, 11.1, and 11.2, the singular value decomposition of Chapter 12, as well as the complex Jordan normal form of Section 14.1. Among these, the SVD is the only normal form that is real for all real matrices A.

The similarity normal forms, such as a diagonal or a Jordan normal form representation of a given matrix $A \in \mathbb{R}^{n,n}$, generally involve complex matrix representations for real linear mappings $A : \mathbb{R}^n \to \mathbb{R}^n$, and thus they may not be all that revealing and desirable for real problems. The main culprit here is the Fundamental Theorem of Algebra which places the roots of a polynomial, such as the minimal or the characteristic polynomial of a real matrix A in $\mathbb{C}$. Thus the Jordan normal forms of Theorems 14.1 and 14.2 can generally only be achieved for a real matrix by using a complex Jordan basis of $\mathbb{C}^n$. For many problems, such as for finding real solutions of linear differential equations for example, it is, however, important to find sparse real representations of real system matrices A. In Section 9.3 we have partially dealt with this issue. Now we extend those results by using the Jordan normal form of Theorem 14.2.

If $A \in \mathbb{R}^{n,n}$ has the eigenvalue $\lambda = a + bi \notin \mathbb{R}$ for a necessarily complex eigenvector $z = u + iw$ with $u, w \in \mathbb{R}^n$, then $Az = \lambda z$ and $A\overline{z} = \overline{Az} = \overline{\lambda z} = \overline{\lambda}\,\overline{z}$ since $\overline{A} = A$ for real A. Therefore the complex number $\overline{\lambda} = a - bi \neq \lambda$ is also an eigenvalue of A for the eigenvector $\overline{z} = u - iw$ if λ is for z. The main idea that establishes a real Jordan form representation for real matrices A lies in combining the two complex conjugate eigenvalues λ and $\overline{\lambda}$ of $A \in \mathbb{R}^{n,n}$ and their corresponding eigenvectors and principal vectors chains into one.

Lemma 4: If $A \in \mathbb{R}^{n,n}$ and $Az = \lambda z$ for $\lambda = a + bi \notin \mathbb{R}$ and $z = u + iw \in \mathbb{C}^n$ with $u, w \in \mathbb{R}^n$ and $a, b \in \mathbb{R}$, then the two vectors u and w are linearly independent. Moreover $Au = au - bw$ and $Aw = bu + aw$. ◀

Proof: By linearity and complex arithmetic we see that

$$Az = A(u + iw) = Au + iAw = \lambda z = (a + bi)(u + iw) = au - bw + i(bu + aw) \ .$$

And both equations $Au = au - bw$ and $Aw = bu + aw$ follow by comparing the real and imaginary parts in this identity.

To prove linear independence of u and $w \in \mathbb{R}^n$, we may assume without loss of generality that $u = \alpha w$ for $\alpha \in \mathbb{R}$. Then $u + iw = (\alpha + i)w$ and

$$Az = A(u + iw) = (\alpha + i)Aw = \lambda z = \lambda\alpha w + i\lambda w = \lambda(\alpha + i)w.$$

Thus $Aw = \lambda w$ with $w \in \mathbb{R}^n$. But the left hand side Aw of this equation lies in $\mathbb{R}^n$, while the right hand side λw does not, a contradiction that makes the real and imaginary parts u and w of a complex eigenvector z of $A \in \mathbb{R}^{n,n}$ linearly independent. ■

If we link the two complex conjugate eigenvalues $\lambda \neq \overline{\lambda}$ and the corresponding eigenvectors $z, \overline{z}$ of a real matrix A, we observe that

$$A \begin{pmatrix} | & | \\ z & \overline{z} \\ | & | \end{pmatrix} = \begin{pmatrix} | & | \\ z & \overline{z} \\ | & | \end{pmatrix} \begin{pmatrix} \lambda & 0 \\ 0 & \overline{\lambda} \end{pmatrix} \ ,$$

i.e., we see a partial complex diagonalization $\begin{pmatrix} \lambda & 0 \\ 0 & \overline{\lambda} \end{pmatrix}$ of A with respect to its two complex eigenvectors z and $\overline{z}$. Using the real and imaginary parts vectors u and w of the complex eigenvector $z = u + iw$ instead, we arrive at an analogous partial real representation of A with respect to u and $w \in \mathbb{R}^n$. Namely

$$A \begin{pmatrix} | & | \\ u & w \\ | & | \end{pmatrix} = \begin{pmatrix} | & | \\ u & w \\ | & | \end{pmatrix} \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \tag{14.4}$$

according to Lemma 4. Note that the real matrix $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ is similar to the complex matrix $\begin{pmatrix} \lambda & 0 \\ 0 & \overline{\lambda} \end{pmatrix}$; see Problem 9 in Section 14.1.P. We can repeat this capture of two complex conjugate eigenvalues λ and $\overline{\lambda}$ of a real matrix A in one 2 by 2 real matrix in (14.4) for principal vectors of grade k and establish linear independence of their real and imaginary parts vectors as well.

Lemma 5: If $z^{(k)} = u^{(k)} + iw^{(k)}$ with $u^{(k)}, w^{(k)} \in \mathbb{R}^n$ is a complex principal vector of grade k for the eigenvalue $\lambda = a + bi \notin \mathbb{R}$ and $A \in \mathbb{R}^{n,n}$, then

(a) $\overline{z^{(k)}} = u^{(k)} - iw^{(k)}$ is a principal vector of grade k for the eigenvalue $\overline{\lambda}$ of A, and

(b) the vectors $u^{(k)}$ and $w^{(k)} \in \mathbb{R}^n$ are linearly independent. ◀

Proof: Part (a) follows readily by complex conjugation of the two defining equations $(A - \lambda I)^k z^{(k)} = 0$ and $(A - \lambda I)^{k-1} z^{(k)} \neq 0$.

In (b), we establish the result for $k = 2$ only. The case of eigenvectors, i.e., $k = 1$, has been settled in Chapter 9. The higher dimensional cases follow along similar lines as our proof for $k = 2$ below.

If the real and imaginary parts $u^{(2)}$ and $w^{(2)} \in \mathbb{R}^n$ of the principal vector $z^{(2)} = u^{(2)} + iw^{(2)} \neq 0$ of grade two are linearly dependent for $\lambda = a + bi \notin \mathbb{R}$ and $A \in \mathbb{R}^{n,n}$, then we may assume without loss of generality that $u^{(2)} = \alpha w^{(2)} \neq 0$ for $\alpha \in \mathbb{R}$.

In this case, $w^{(2)}$ is a real principal vector of grade 2 for A and λ, since $(A - \lambda I)^2 z^{(2)} = (\alpha + i)(A - \lambda I)^2 w^{(2)} = 0$ with $\alpha \in \mathbb{R}$. Thus for $k = 2$ and one complex principal vector of grade 2 with linearly dependent real and imaginary parts, the real vector $x := w^{(2)} \neq 0$ is also a principal vector of grade 2.

If we write out $(A - \lambda I)^2 x = (A^2 - 2\lambda A + \lambda^2 I)x = 0$, we note that $A^2 x = (2\lambda A - \lambda^2 I)x$ is a real vector. Thus $A^2 x = \overline{A^2 x} = (2\overline{\lambda} A - \overline{\lambda}^2 I)x$, since x, A, and I are all real. Consequently

$$0 = A^2 x - \overline{A^2 x} = (2(\lambda - \overline{\lambda})A - (\lambda^2 - \overline{\lambda}^2)I)x = (4biA - 4abiI)x = 4bi(A - aI)x \,,$$

or $(A - aI)x = 0$ since $\lambda - \overline{\lambda} = 2bi \neq 0$ and $\lambda^2 - \overline{\lambda}^2 = 4abi$.

If $Ax = ax$ for $x \in \mathbb{R}^n$ and $a \in \mathbb{R}$, then the identity $0 = (A - \lambda I)^2 x = (A^2 - 2\lambda A + \lambda^2 I)x = a^2 x - 2(a + bi)ax + (a^2 - b^2 + 2abi)x$ makes its real parts vector equal to zero. Namely $(a^2 - 2a^2 + a^2 - b^2)x = -b^2 x = 0$. This forces $b = 0$, unless $x = 0$. But if $b = 0$, then $\lambda = a \in \mathbb{R}$, a contradiction. And if $x = w^{(2)} = 0$, then $z^{(2)} = 0$ gives another contradiction. Thus the real and imaginary parts of a grade 2 principal vector for a complex eigenvalue of a real matrix must be linearly independent. ■

The following follows directly from Lemma 5 (a).

Corollary 2: If A is a real square matrix, then the Jordan blocks for nonreal eigenvalues occur in pairs in the Jordan normal form J of A. Specifically if $\lambda = a + bi \notin \mathbb{R}$ is an eigenvalue of $A \in \mathbb{R}^{n,n}$ and if the Jordan block $J(\lambda, m)$ occurs in J, so does the block $J(\overline{\lambda}, m)$. ◀

If $\{x^{(\ell)} = u^{(\ell)} + iw^{(\ell)}\}$ denotes a chain of principal vectors of order k for $\ell = k, ..., 1, 0$ and real vectors $u^{(\ell)}$ and $w^{(\ell)}$ for the nonreal eigenvalue $\lambda = a + bi$ of $A \in \mathbb{R}^{n,n}$ (with $x^{(0)} := 0$ for convenience), then for any $1 \le \ell \le k$ we have

$$\begin{aligned} A\begin{pmatrix} | & | \\ u^{(\ell)} & w^{(\ell)} \\ | & | \end{pmatrix} &= \begin{pmatrix} | & | \\ u^{(\ell)} & w^{(\ell)} \\ | & | \end{pmatrix}\begin{pmatrix} a & b \\ -b & a \end{pmatrix} + \begin{pmatrix} | & | \\ u^{(\ell-1)} & w^{(\ell-1)} \\ | & | \end{pmatrix} \\ &= \begin{pmatrix} | & | & | & | \\ u^{(\ell)} & w^{(\ell)} & u^{(\ell-1)} & w^{(\ell-1)} \\ | & | & | & | \end{pmatrix}\left(\begin{array}{cc} a & b \\ -b & a \\ \hline 1 & 0 \\ 0 & 1 \end{array}\right) . \end{aligned} \tag{14.5}$$

To see this, observe that $x^{(\ell-1)} = u^{(\ell-1)} + iw^{(\ell-1)} = (A - \lambda I)x^{(\ell)}$ by the definition of a principal vectors chain. And equation (14.5) becomes apparent by expansion and by sorting $x^{(\ell-1)}$'s real and imaginary part vectors as follows:

$$\begin{aligned} x^{(\ell-1)} &= (A - \lambda I)x^{(\ell)} = (A - aI - biI)(u^{(\ell)} + iw^{(\ell)}) \\ &= Au^{(\ell)} - au^{(\ell)} + bw^{(\ell)} + i(Aw^{(\ell)} - aw^{(\ell)} - bu^{(\ell)}) \\ &=: u^{(\ell-1)} + iw^{(\ell-1)} . \end{aligned}$$

Comparing the real and imaginary part vectors above, gives us the two identities

$$Au^{(\ell)} - au^{(\ell)} + bw^{(\ell)} = u^{(\ell-1)} \quad \text{and} \quad Aw^{(\ell)} - aw^{(\ell)} - bu^{(\ell)} = w^{(\ell-1)} ,$$

or

$$Au^{(\ell)} = au^{(\ell)} - bu^{(\ell)} + u^{(\ell-1)} \quad \text{and} \quad Aw^{(\ell)} = bu^{(\ell)} + aw^{(\ell)} + w^{(\ell-1)} .$$

This establishes (14.5) . Note that (14.5) is a real matrix equation that generalizes the real and imaginary parts eigenvector equation (14.4) to principal vectors.

For each complex eigenvalue $\lambda = a + bi \notin \mathbb{R}$ of a matrix $A \in \mathbb{R}^{n,n}$ there may be several Jordan blocks $J(\lambda, n_1), ..., J(\lambda, n_j)$ associated with $\lambda \in \mathbb{C}$. Each of these has an equal sized complex conjugate Jordan block $J(\overline{\lambda}, n_1), ..., J(\overline{\lambda}, n_j)$ in the complex Jordan normal form of A. The **real Jordan normal form** of a real matrix A uses the two corresponding Jordan blocks $J(\lambda, m)$ and $J(\overline{\lambda}, m)$ in tandem to represent A with respect to their joint principal real and imaginary parts vector subspaces by the $2m$ by $2m$ real matrix

$$J(a, b, 2m) \; = \left(\begin{array}{cc|cc ccc cc} a & b & 0 & 0 & & & & 0 & 0 \\ -b & a & 0 & 0 & & & & 0 & 0 \\ \hline 1 & 0 & a & b & & & & & \\ 0 & 1 & -b & a & & & & & \\ & & & \ddots & & \ddots & & & \\ & & & & \ddots & & \ddots & 0 & 0 \\ & & & & & & & 0 & 0 \\ \cline{8-9} 0 & 0 & & & & 1 & 0 & a & b \\ 0 & 0 & & & & 0 & 1 & -b & a \end{array} \right)$$

with m diagonal 2 by 2 blocks $\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \in \mathbb{R}^{2,2}$, $m-1$ subdiagonal 2 by 2 identity matrices, and zeros elsewhere. The matrix $J(a, b, 2m)$ is called the **real Jordan block** for the complex eigenvalue $\lambda = a + bi$ of A where $b \neq 0$.

Notation: We denote Jordan normal forms and Jordan blocks by the same letter J throughout. A specific Jordan block can carry either two or three variables as in $J(\lambda, 5)$ or in $J(a, b, 2k)$. Jordan blocks with 2 variables $J(\lambda, k)$ always denote the standard Jordan block $\begin{pmatrix} \lambda & & & 0 \\ 1 & \lambda & & \\ & \ddots & \ddots & \\ 0 & \dots & 1 & \lambda \end{pmatrix}$ with $\lambda \in \mathbb{C}$ or $\lambda \in \mathbb{R}$ and k a positive integer. Jordan blocks with three attached variables $J(a, b, \ell)$ always denote a real Jordan block as just described. Here a and b are always real numbers and ℓ is a positive even integer. ◀

Theorem 14.3: (**The Real Jordan Normal Form**)

Every real square matrix A is similar over the reals to its real Jordan normal form J. J is a block diagonal matrix comprised of real Jordan blocks. Namely if $\lambda \in \mathbb{R}$ is an eigenvalue of A, the Jordan blocks associated with it have the standard form $J(\lambda, m) = \begin{pmatrix} \lambda & & & 0 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ 0 & & 1 & \lambda \end{pmatrix} \in \mathbb{R}^{m,m}$ for various sizes m, while for $\lambda = a + bi \notin \mathbb{R}$ and $a, b \in \mathbb{R}$, the associated Jordan blocks have the 2 by 2 lower block triangular form

$$J(a,b,2r) = \begin{pmatrix} a & b & & & & 0 \\ -b & a & & & & \\ 1 & 0 & a & b & & \\ 0 & 1 & -b & a & & \\ & & \ddots & & \ddots & \\ & 0 & & & \ddots & \ddots \end{pmatrix}.$$

The individual eigenvalues and their respective Jordan block dimensions and types are determined uniquely by A. The order of the Jordan blocks that appear in J is not unique.

Each Jordan block $J(a,b,2r) = \begin{pmatrix} a & b & & & & \\ -b & a & & & & \\ 1 & 0 & a & b & & \\ 0 & 1 & -b & a & & \\ & & \ddots & & \ddots & \\ & & & & \ddots & \ddots \end{pmatrix}$ in Theorem 14.3 can be represented in an even sparser form as $\begin{pmatrix} a & b & & & & \\ -b & a & & & & \\ 0 & 1 & a & b & & \\ 0 & 0 & -b & a & & \\ & & \ddots & & \ddots & \\ & & & & \ddots & \ddots \end{pmatrix}$. Here the off diagonal 2 by 2 blocks I_2 of the real Jordan normal form J have been replaced by

$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ while all other entries are left unchanged in J. This sparse form shows that every real square matrix can be reduced to a real tridiagonal matrix via a real similarity.

Corollary 3: Every real square matrix is similar over $\mathbb{R}$ to a real tridiagonal matrix. ◀

Example 6: Find the real Jordan normal form of $A = \begin{pmatrix} 4 & 7 & -1 & -6 \\ -2 & -3 & 2 & 3 \\ 0 & 1 & 2 & -1 \\ 0 & -1 & 0 & 1 \end{pmatrix} \in \mathbb{R}^{4,4}$.

Using the vector iteration method of Section 9.1 with $x = \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}$ for example gives us the vector iteration matrix $\begin{pmatrix} | & & | \\ x & \dots & A^4x \\ | & & | \end{pmatrix} = \begin{pmatrix} 0 & 5 & 16 & 30 & 32 \\ 0 & -1 & -4 & -8 & -8 \\ 1 & 3 & 6 & 8 & 4 \\ -1 & -1 & 0 & 4 & 12 \end{pmatrix}$ with the RREF $\begin{pmatrix} 1 & 0 & 0 & 0 & -4 \\ 0 & 1 & 0 & 0 & 8 \\ 0 & 0 & 1 & 0 & -8 \\ 0 & 0 & 0 & 1 & 4 \end{pmatrix}$. Thus the minimal polynomial $p_A(\lambda) = x^4 - 4x^3 + 8x^2 - 8x + 4 = (x^2 - 2x + 2)^2$. Since it has degree $4 = n$, it coincides with the characteristic polynomial $f_A(\lambda) = \det(\lambda I - A)$ of A. Moreover since $p_A(\lambda)$ is a square of the irreducible real polynomial $x^2 - 2x + 2 = (x-1-i)(x-1+i)$, A is not diagonalizable. A has the double complex conjugate eigenvalues $1 \pm i$. Thus the real Jordan normal form of A is

$$J = \begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & 1 \end{pmatrix}$$

according to Theorem 14.3. The real Jordan normal form J can be achieved via the similarity $J = X^{-1}AX$ of A for $X = \begin{pmatrix} 3 & 1 & 3 & -1 \\ 0 & 0 & -1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix}$. The first two columns of X contain the real and imaginary parts of the principal vector $x^{(2)} = \begin{pmatrix} 3+i \\ 0 \\ 1+i \\ 1 \end{pmatrix}$ of grade 2 for A and $\lambda = 1+i$. Here $x^{(2)} \in \ker(A-(1+i)I)^2$, while $x^{(2)} \notin \ker(A-(1+i)I)$. The last two columns of X contain the real and imaginary part vectors of the eigenvector $x^{(1)} = (A - (1+i)I)x^{(2)} = \begin{pmatrix} 3-i \\ -1 \\ 1 \\ -i \end{pmatrix}$ for λ.

The students should verify all of the assertions by hand computations and row reducing $A - (1+i)I$ and $(A - (1+i)I)^2$. ◀

In the sequel we demonstrate how to represent a rational matrix $A \in \mathbb{Q}^{n,n}$ in a sparse normal form if its entries are rational numbers of the form $\frac{r}{s}$ for integers r and s with $s \neq 0$. (For the definition of rational numbers $\mathbb{Q}$, see Appendix A.) The **companion matrix normal form** of a rational matrix is a sparse rational matrix that is reminiscent of the Jordan normal form, but it uses different types of diagonal blocks. If p is a monic polynomial of degree n, i.e., if $p(x) = x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0$ with coefficients $a_j \in \mathbb{R}$, then the **companion matrix**

$$C(p) = \begin{pmatrix} 0 & & & -a_0 \\ 1 & & & \vdots \\ & \ddots & & \vdots \\ & & 1 & -a_{n-1} \end{pmatrix}_{nn}$$

has the characteristic polynomial $f_{C(p)}(\lambda) = \det(\lambda I - C(p)) = p(\lambda)$; see Problem 9 in Section 9.R or expand $\det(\lambda I - C(p))$ along its last column. The companion matrix $C(p)$ has the same characteristic polynomial as its complex and as its real Jordan normal form, since these three matrices are all similar to each other.

We start by investigating the real Jordan normal form for $C(p)$ in light of the possible factorizations of $p(x)$ over the reals before studying the same over the rational numbers. The real Jordan normal form J of $C(p)$ is lower block triangular according to Theorem 14.3. Specifically for a real root λ_k of $p(x)$, the real eigenvalue λ_k of $C(p)$ appears on the diagonal of J, repeated according to its algebraic multiplicity. For a complex root $\lambda_k = a_k + b_k i \notin \mathbb{R}$ of p, the two corresponding complex conjugate eigenvalues λ_k and $\overline{\lambda_k}$ of $C(p)$ give rise to a 2 by 2 real diagonal block $\begin{pmatrix} a_k & b_k \\ -b_k & a_k \end{pmatrix} \in \mathbb{R}^{2,2}$, possibly repeated; all according to Theorem 14.3. Clearly for each complex eigenvalue $\lambda_k = a_k + b_k i \notin \mathbb{R}$ we have $\det\left(\lambda I_2 - \begin{pmatrix} a_k & b_k \\ -b_k & a_k \end{pmatrix}\right) = (\lambda - a_k)^2 + b_k^2$. The lower co-diagonal entries of 1 or of I_2 in J for real or complex eigenvalues of $C(p)$ of index $j > 1$, respectively, can be ignored when forming $C(p)$'s characteristic polynomial since the characteristic polynomial of a block triangular matrix is equal to the product of the characteristic polynomials of its individual diagonal blocks. This follows from the determinant Proposition, part 13, of Chapter 8. Therefore

$$\begin{aligned} p(\lambda) &= f_{C(p)}(\lambda) = \det(\lambda I_n - C(p)) = \\ &= \prod_{\lambda_\ell \text{ real}} (\lambda - \lambda_\ell)^{n_\ell} \prod_{\substack{\lambda_j \text{ complex} \\ \lambda_j = a_j + b_j i}} \left((\lambda - a_j)^2 + b_j^2\right)^{n_j} \end{aligned}$$

with $n_\ell,\ n_j \geq 1$. This factors the given polynomial $p(\lambda)$ over the reals into degree one or degree 2 real polynomial factors, the latter of which have no real roots since $b_j \neq 0$. This proves the following well known result about factoring real polynomials p into **irreducible factors**. Here we call a real polynomial q **irreducible** over its field of coefficients $\mathbb{R}$ if it cannot be factored into real polynomials of lower degree, such as $q(\lambda) = (\lambda - a_k)^2 + b_k^2$ for $b_k \neq 0$.

Corollary 4: Every real polynomial $p(x)$ can be factored over the reals into first and second degree real polynomial factors. ◀

Example 7: Factor the two real polynomials $f(x) = x^5 + 1$ and $g(x) = x^5 - 1$ into the product of irreducible real polynomials.

We approach this problem by computing the eigenvalues λ_i and μ_j of the companion matrices $C(f) = \begin{pmatrix} 0 & & & -1 \\ 1 & & & 0 \\ & \ddots & & \vdots \\ & & 1 & 0 \end{pmatrix}_{5,5}$ and $C(g) = \begin{pmatrix} 0 & & & 1 \\ 1 & & & 0 \\ & \ddots & & \vdots \\ & & 1 & 0 \end{pmatrix}_{5,5}$ in MATLAB. They are $\lambda_1 = -1,\ \lambda_{2,3} = -0.309 \pm 0.951i,\ \lambda_{4,5} = 0.809 \pm 0.568i$ and $\mu_1 = 1,\ \mu_{2,3} = 0.309 \pm 0.951i,\ \mu_{4,5} = -0.809 \pm 0.588i$, respectively.

Thus over $\mathbb{C}$,

$$\begin{aligned} f(x) &= (x+1)(x-\lambda_2)(x-\overline{\lambda_2})(x-\lambda_4)(x-\overline{\lambda_4}) \\ &= (x+1)(x^2+0.618x+1)(x^2-1.618x+1) \\ &= (x+1)(x^2-(0.5-\frac{\sqrt{5}}{2})x+1)(x^2-(0.5+\frac{\sqrt{5}}{2})x+1)\ , \end{aligned}$$

while

$$\begin{aligned} g(x) &= (x-1)(x^2-0.618x+1)(x^2+1.618x+1) \\ &= (x+1)(x^2+(0.5-\frac{\sqrt{5}}{2})x+1)(x^2+(0.5+\frac{\sqrt{5}}{2})x+1)\ . \end{aligned}$$

Here we have expressed both f and g as a product of irreducible real factor polynomials. ◀

If we attempt to factor rational polynomials $p(x)$ with coefficients in $\mathbb{Q}$ into irreducible rational polynomial factors instead, we must realize that there is no equivalent result to Corollary 4 for the maximal factor degree of real factorizations. Namely, arbitrary degree rational polynomials may be irreducible over $\mathbb{Q}$, such as $p(x) = x^n - 2$ is for any $n \geq 1$. Note that the two rational polynomials in Example 7 factor irreducibly over $\mathbb{Q}$ into $f(x) = (x+1)(x^4 - x^3 + x^2 - x + 1)$ and $g(x) = (x-1)(x^4 + x^3 + x^2 + 1)$. This can be verified by long division and by looking back at the the real factors of f and g in Example 7.

Now we turn to rational sparse matrix representations of rational matrices $A_{nn} \in \mathbb{Q}^{n,n}$. We assume that $f_A(\lambda) = \det(\lambda I - A) = \prod_{i=1}^{k} p_i(x)^{n_i}$ and $p_A(\lambda) = \prod_{i=1}^{k} p_i(x)^{m_i}$ with $m_i \le n_i$ for each i. Furthermore we assume that the polynomials $p_i(x)$ are irreducible over $\mathbb{Q}$. Then there exists a basis of $\mathbb{Q}^n$ that represents the linear transformation $x \mapsto Ax$ of $\mathbb{Q}^n$ as a block diagonal matrix. This matrix contains a companion matrix for each irreducible factor p_i of f_A (or p_A) as its diagonal block, with or without accompanying co-diagonal identity matrices as follows:

Theorem 14.4: (**The Companion Matrix Normal Form**)

Let $A \in \mathbb{Q}^{n,n}$ be a matrix with rational entries. Assume that the characteristic polynomial of A is $f_A(\lambda) = \det(\lambda I - A) = \prod_{i=1}^{k} p_i^{n_i}$ and that the minimum polynomial $p_A(\lambda) = \prod_{i=1}^{k} p_i^{m_i}$ and $f_A(x)$ are both factored irreducibly over $\mathbb{Q}$.

Then $m_i \le n_i$ for each $i = 1, ..., k$ and A is similar over $\mathbb{Q}$ to a block diagonal matrix $C \in \mathbb{Q}^{n,n}$ comprised of k diagonal blocks C_i. In turn, each diagonal block C_i is in lower block diagonal form

$$C_i = \begin{pmatrix} C(p_i) & & & 0 \\ I_{d_i} & C(p_i) & & \\ & \ddots & \ddots & \\ 0 & & I_{d_i} & C(p_i) \end{pmatrix}_{\ell_i \cdot d_i \times \ell_i \cdot d_i},$$

comprised of $\ell_i \le m_i$ companion matrices $C(p_i)$ for the irreducible factor p_i of f_A (or p_A) on its block diagonal and copies of the identity matrix I_{d_i} of dimension $d_i = \text{degree}(p_i)$ on its lower block co–diagonal.

The matrix $C = \text{diag}(C_i)$ is called the **companion matrix normal form** of A over $\mathbb{Q}$.

The order of the individual blocks C_i in C is not unique.

Example 8: The companion matrix $A = \begin{pmatrix} 0 & 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$ for the polynomial $f(x) = x^5 + 1$ of Example 7 has the following normal forms, depending on the base field that we consider:

Over $\mathbb{C}$, A is similar to the complex diagonal matrix

$$\begin{pmatrix} -1 & & & & \\ & -0.309+0.951i & & & \\ & & -0.309-0.951i & & \\ & & & 0.809+0.568i & \\ & & & & 0.809-0.568i \end{pmatrix}$$

according to Example 7. This is the complex Jordan normal form of A. Over $\mathbb{C}$, A is diagonalizable because it has five distinct complex eigenvalues.

Over $\mathbb{R}$, A is similar to the block diagonal real matrix

$$\begin{pmatrix} -1 & & & & \\ & -0.309 & 0.951 & & \\ & -0.951 & -0.309 & & \\ & & & 0.809 & 0.568 \\ & & & -0.568 & 0.809 \end{pmatrix},$$

which is A's real Jordan normal form.

Finally over $\mathbb{Q}$, A is similar to its companion matrix normal form

$$\begin{pmatrix} -1 & & & & \\ & 0 & 0 & 0 & -1 \\ & 1 & 0 & 0 & 1 \\ & 0 & 1 & 0 & -1 \\ & 0 & 0 & 1 & 1 \end{pmatrix}$$

since $f(x) = x^5 + 1 = (x+1)(x^4 - x^3 + x^2 - x + 1)$ with both polynomial factors irreducible over $\mathbb{Q}$. ◀

Matrix normal forms allow us to classify all matrices that represent the same underlying linear transformation, except for different bases.

Proposition: Two n by n matrices A and B are similar if and only if the two matrices have the same complex Jordan normal form, the same real Jordan normal form, or the same companion matrix normal form.

Here we call two matrix normal forms identical, if the two forms contain the same collection of normal form blocks, disregarding their specific order. ◀

Finally, the Jordan normal form allows us to prove a classical result, namely that every real or complex matrix can be written as the product of two symmetric matrices, real or complex, respectively.

Corollary 5: Every real or complex matrix A_{nn} can be written as the product of two symmetric (real or complex, respectively) matrix factors $A = S_1 S_2$ with $S_i^T = S_i$ and S_1 nonsingular. ◀

Note that not all complex matrices can be written as the product of two hermitian matrices, see Problems 8 – 12 below. Symmetry is a stronger matrix attribute than being hermitian in the complex matrix case.

Proof: Every Jordan block $J(\lambda, k) = \begin{pmatrix} \lambda & & & 0 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ 0 & & 1 & \lambda \end{pmatrix}$ can be factored as

$$J(\lambda, k) = \begin{pmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{pmatrix}_{k,k} \begin{pmatrix} 0 & & 1 & \lambda \\ & \iddots & \iddots & \\ 1 & \iddots & & \\ \lambda & & & 0 \end{pmatrix}_{kk}$$

into the product of two real (or complex, if $\lambda \notin \mathbb{R}$) symmetric matrices. Likewise every real Jordan block

$$J(a, b, 2m) = \begin{pmatrix} a & b & & & & & & \\ -b & a & & & & & & \\ 1 & 0 & \ddots & & & & & \\ 0 & 1 & & \ddots & & & & \\ & & \ddots & & \ddots & & & \\ & & & & 1 & 0 & a & b \\ & & & & 0 & 1 & -b & a \end{pmatrix}_{2m,2m}$$

for a complex eigenvalue $a + bi \notin \mathbb{R},\ a,\ b \in \mathbb{R}$ factors symmetrically as

$$J(a, b, 2m) = \begin{pmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{pmatrix}_{2m,2m} \begin{pmatrix} & & & & 0 & 1 & -b & a \\ & & & & 1 & 0 & a & b \\ & & & \iddots & & \iddots & & \\ 0 & 1 & & & \iddots & & & \\ 1 & 0 & \iddots & & & & & \\ -b & a & & & & & & \\ a & b & & & & & & \end{pmatrix}.$$

This can be readily checked. Consequently any Jordan normal form matrix J, real or complex, can be written as the product $J = S_1 S_2$ of two block diagonal matrices S_1 and S_2 that carry symmetric diagonal blocks of the form $\begin{pmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{pmatrix}$ in S_1, and $\begin{pmatrix} 0 & & & 1 & \lambda \\ & & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \\ & & \cdot^{\cdot^{\cdot}} & & \\ 1 & \cdot^{\cdot^{\cdot}} & & & \\ \lambda & & & & 0 \end{pmatrix}$ or $\begin{pmatrix} & & & & 0 & 1 & -b & a \\ & & & & 1 & 0 & a & b \\ & & & \cdot^{\cdot^{\cdot}} & & \cdot^{\cdot^{\cdot}} & & \\ 0 & 1 & & & \cdot^{\cdot^{\cdot}} & & & \\ 1 & 0 & \cdot^{\cdot^{\cdot}} & & & & & \\ -b & a & & & & & & \\ a & b & & & & & & \end{pmatrix}$ in S_2. By construction, both S_1 and S_2 are symmetric while S_1 is nonsingular.

If $X^{-1}AX = J$ is the Jordan normal form of A and if $J = S_1S_2$, then $A = XJX^{-1} = XS_1S_2X^{-1} = XS_1X^T \left(X^T\right)^{-1} S_2X^{-1} = T_1T_2$ for the nonsingular symmetric matrix $T_1 = XS_1X^T = T_1^T$ and the symmetric matrix $T_2 = \left(X^T\right)^{-1} S_2X^{-1} = T_2^T$. ■

The symmetric factorization that exists for every square matrix allows us to classify real diagonalizable matrices via their symmetric factors.

Corollary 6: A real matrix A_{nn} is diagonalizable over $\mathbb{R}$ if and only if A can be factored into the product of one positive definite real symmetric matrix and a real symmetric matrix. ◀

Proof: If A is diagonalizable over $\mathbb{R}$, i.e., if $X^{-1}AX = J = \operatorname{diag}(\lambda_i)$ for $\lambda_i \in \mathbb{R}$ and $X \in \mathbb{R}^{n,n}$, then

$$A \;=\; XIX^T \cdot (X^T)^{-1}JX^{-1} \;=\; S_1S_2$$

with $S_1 = XIX^T = XX^T = S_1^T$ positive definite according to the definition in Section 11.3 and $S_2 = X^{-T}\operatorname{diag}(\lambda_i)X^{-1}$ real symmetric. Here we have abbreviated $(X^T)^{-1}$ by X^{-T}.

Conversely, if $A = S_1S_2$ with $S_i = S_i^T \in \mathbb{R}^{n,n}$ and S_1 positive definite, then there is a real orthogonal matrix U according to Section 11.1 with $U^TS_1^{-1}U = \operatorname{diag}\left(\frac{1}{\mu_i}\right)$ for the positive eigenvalues $\mu_i > 0$ of S_1 since S_1^{-1} has the eigenvalues $\frac{1}{\mu_i}$. With $M := \operatorname{diag}(\sqrt{\mu_i}) = M^T \in \mathbb{R}^{n,n}$ we then have

$$M^TU^TS_1^{-1}UM \;=\; \operatorname{diag}(\sqrt{\mu_i})\operatorname{diag}\left(\frac{1}{\mu_i}\right)\operatorname{diag}(\sqrt{\mu_i}) \;=\; I\ ,$$

or $X^TS_1^{-1}X = I$ for $X := UM \in \mathbb{R}^{n,n}$. This implies that $S_1^{-1} = X^{-T}X^{-1}$, or $S_1 = XX^T$. For this specific matrix X we look at

$$X^{-1}AX \;=\; X^{-1}S_1S_2X \;=\; X^{-1}(XX^T)X^{-T}X^TS_2X \;=\; IX^TS_2X \;=\; X^TS_2X\ .$$

Thus A is similar to the real symmetric matrix X^TS_2X which in turn is orthogonally diagonalizable over $\mathbb{R}$ according to Chapter 11. Therefore A itself is diagonalizable over the reals. ■

This result for diagonalizable matrices relies upon our understanding of nondiagonalizable ones; giving us a circular closing to the subject.

14.2.P Problems

1. Find the real Jordan normal form of $A = \begin{pmatrix} 1 & -6 & -3 & -1 \\ 4 & 7 & 12 & -3 \\ -3 & -3 & -7 & 2 \\ -2 & -5 & -6 & 1 \end{pmatrix}$. (Hint: Decide first whether to use vector iteration as in Chapter 9.1 or determinants as in Chapter 9.1.D to find the eigenvalues of A.)

2. What is the companion matrix for $p(x) = x^7 - x^5 + 4x^2 - 3x + 12$?

3. Write out all possible real Jordan normal forms for a 13 by 13 real matrix A with the eigenvalues $\lambda = 2 + 3i$ (3 fold), $\mu = 1 - 2i$ (double), and $\nu = 7$. What is the algebraic multiplicity of ν? What are the possible geometric multiplicities of λ?

4. Determine the real Jordan normal form of the companion matrix $C = \begin{pmatrix} 0 & & & & 0 \\ 1 & & & & 6 \\ & 1 & & & -5 \\ & & 1 & & -5 \\ 0 & & & 1 & 5 \end{pmatrix}$. Is C diagonalizable over $\mathbb{R}$ or not? What are C's eigenvalues?

5. Let $J = \begin{pmatrix} \lambda & & & 0 \\ 1 & \lambda & & \\ & \ddots & \ddots & \\ 0 & & 1 & \lambda \end{pmatrix}$ be a Jordan block of size k by k.

   (a) Find an eigenvector of J for λ.

   (b) Find the index of the eigenvalue λ of J.

   (c) Find a principal vector of grade k for J and λ.

   (d) Find a principal vector of grade $1 < j < k$ for J and λ.

6. We call a nonzero row vector $y \in \mathbb{C}^n$ a **left eigenvector** of a matrix A for the eigenvalue λ, if $yA = \lambda y$. Likewise a row vector z with $z(A - \lambda I)^j = 0$ and $z(A - \lambda I)^{j-1} \neq 0$ is called a **left principal vector** of grade j for A and λ.

   (a) For the Jordan block $J = \begin{pmatrix} \lambda & & & 0 \\ 1 & \lambda & & \\ & \ddots & \ddots & \\ 0 & & 1 & \lambda \end{pmatrix}$ of size k by k, find a left eigenvector.

   (b) For the Jordan block $J = J(\lambda, k)$ find the left principal vectors for all possible grades.

7. Let $A = \begin{pmatrix} 0 & & & -a_0 \\ 1 & & & \vdots \\ & \ddots & & \vdots \\ & & 1 & -a_{n-1} \end{pmatrix} = C(p)$ be the companion matrix for $p(x) = x^2 + a_{n-1}x^{n-1} + ... + a_1 x + a_0$.
   Show: If z is a root of the polynomial p, then the row vector $v = \begin{pmatrix} 1 & z & z^2 & \dots & z^{n-1} \end{pmatrix}$ is a left eigenvector for the eigenvalue z of A.

8. Show that $A = \begin{pmatrix} 2 & i \\ -i & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathbb{C}^{2,2}$ can be expressed as the product of two symmetric complex matrices $S_i = S_i^T \in \mathbb{C}^{2,2}$. Find S_1 and S_2.

9. Show that $B = \begin{pmatrix} 2i & -2 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathbb{C}^{2,2}$ can be expressed as the product of two symmetric complex matrices $S_i = S_i^T \in \mathbb{C}^{2,2}$.

10. Show that $A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} 4 & -1 & 1 \\ -1 & -1 & -1 \\ 1 & -1 & 0 \end{pmatrix}$ can be diagonalized over $\mathbb{R}$.

11. Factor $A = \begin{pmatrix} -2 & 1 \\ 4 & -4 \end{pmatrix}$ into the product of two real symmetric matrices, if possible.

12. (a) Show that $A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ is the product of $Y = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}$ and $S = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$.

(b) Which of the matrices S or Y is symmetric? Which is positive definite?

(c) How does this square with Corollary 5?

(d) Can you find a symmetric factorization of A in part (a) into the product of two real symmetric matrices with one positive definite factor?

(e) If the answer to (d) is affirmative, find such a symmetric factorization of A.

13. Show that the two real Jordan blocks $J(a, b, 2)$ and $J(a, -b, 2)$ are always similar.

14.3 Applications (Differential Equations and MATLAB)

(a) *Linear systems of differential equations*

In Section 9.3 we have solved **systems of differential equations** via linear algebra. Specifically for diagonalizable real system matrices A_{nn} we have described the solutions $x(t) \in \mathbb{R}^n$ of the DE $x'(t) = Ax(t)$ in 9.3.

In this section we assume that $J = U^{-1}AU$ is the Jordan normal form, real or complex, diagonal or not, of $A \in \mathbb{R}^{n,n}$ and that U is nonsingular. As in equation (9.7), we rewrite the DE $x' = Ax = UJU^{-1}x$ as

$$U^{-1}x'(t) \;=\; U^{-1}Ax(t) \;=\; JU^{-1}x(t)\ .$$

By setting $v(t) := U^{-1}x(t)$ we only need to solve the nearly separated system of DEs $v'(t) = Jv(t)$. If such a solution v is known, then clearly $x(t) := Uv(t)$ solves the original system of DEs $x'(t) = Ax(t)$.

If the Jordan normal form $J = \operatorname{diag}(J_1, ..., J_k)$ for A consists of k Jordan blocks J_i, then the problem $v'(t) = Jv(t)$ separates into k subproblems, one for each of the Jordan blocks J_i in J. The diagonalizable case $J_i = (\lambda_i) \in \mathbb{C}^{1,1}$ has been treated in Section 9.3. By Theorems 14.1 and 14.2, we assume that a Jordan block $J_i = \begin{pmatrix} \lambda & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & \lambda \end{pmatrix}$ for A has size ℓ by ℓ for $\ell > 1$ and $\lambda \in \mathbb{C}$ and try to solve $w'(t) = J_i w(t)$ for $w(t) = \begin{pmatrix} w_1(t) \\ \vdots \\ w_\ell(t) \end{pmatrix}$.

Since

$$w'(t) \;=\; \begin{pmatrix} w_1'(t) \\ \vdots \\ w_\ell'(t) \end{pmatrix} \;=\; \begin{pmatrix} \lambda & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & \lambda \end{pmatrix} \begin{pmatrix} w_1(t) \\ \vdots \\ w_\ell(t) \end{pmatrix}\ , \qquad (14.6)$$

we immediately note that $w_1'(t) = \lambda w_1(t)$ from row one and that $w_1(t) = ce^{\lambda t}$ according to elementary calculus. Next $w_2'(t) = w_1(t) + \lambda w_2(t)$ in row 2 of (14.6) is solved by $w_2(t) := cte^{\lambda t} = tw_1(t)$ since $w_2'(t) = ce^{\lambda t} + \lambda cte^{\lambda t} = w_1(t) + \lambda w_2(t)$ according to the product rule of differentiation. For arbitrary $j > 1$ we set $w_j(t) := c\dfrac{t^{j-1}}{(j-1)!}e^{\lambda t} = \dfrac{t}{j-1}w_{j-1}(t)$ and observe that

$$w_j'(t) = c\frac{t^{j-2}}{(j-2)!}e^{\lambda t} + \lambda c\frac{t^{j-1}}{(j-1)!}e^{\lambda t} = w_{j-1}(t) + \lambda w_j(t)$$

satisfies the j^{th} equation of (14.6) . Therefore for one Jordan block J_i of size ℓ by ℓ with $\ell > 1$ and the eigenvalue $\lambda \in \mathbb{C}$, the general solution to $w'(t) = J_i w(t)$ is given in vector form by

$$w(t) = ce^{\lambda t}\begin{pmatrix} 1 \\ t \\ t^2/2 \\ \vdots \\ \dfrac{t^{\ell-1}}{(l-1)!} \end{pmatrix} \in \mathbb{C}^n \ . \tag{14.7}$$

Next, a solution $v(t)$ to the Jordan normal form DE $v'(t) = Jv(t)$ with $J = \text{diag}(J_1, ..., J_k)$ can be concatenated from the k individual solutions to the Jordan block DEs $w'(t) = J_i w(t)$, $i = 1, ..., k$, that have just been described. Finally $x(t) := Uv(t)$ is the solution to the original problem $x'(t) = Ax(t)$ if $U^{-1}AU = J = \text{diag}(J_1, ..., J_k)$.

For real system matrices A with complex eigenvalues, the solution that we have just described involves complex functions for all eigenvalues λ_i of A that do not lie in $\mathbb{R}$. This is unfortunate because a complex solution $x(t)$ gives us little applicable information for the underlying real world problem. In Section 9.3, namely in equations (9.9), (9.10), and in Example 10 there, we have dealt with the case of one complex eigenvalue $\lambda = a + bi$ with $b \neq 0$ of index 1 for $A \in \mathbb{R}^{2,2}$. If we use the Real Jordan Normal Form Theorem 14.3 for complex eigenvalues of $A \in \mathbb{R}^{n,n}$, we can generalize the results from Section 9.3 to obtain real solutions for complex eigenvalues of A of index greater than 1.

If $J = \text{diag}(J_i)$ is the complex Jordan normal form of a real matrix A and if λ is a complex eigenvalue $\lambda = a + bi \notin \mathbb{R}$ for $a, b \in \mathbb{R}$ with index m exceeding 1, then for each associated Jordan block $J(\lambda)$ in J there is a corresponding equal sized Jordan block $J(\overline{\lambda})$ among the diagonal blocks J_i in J according to Section 14.2. Two such equal sized Jordan blocks $J(\lambda)$ and $J(\overline{\lambda})$ were mated into one real block in the real Jordan normal form of Theorem 14.3. Namely, the two blocks were fused into the real Jordan block $J_j =$

$$J(a,b,2m) = \begin{pmatrix} a & b & & & & & \\ -b & a & & & & & \\ 1 & 0 & \ddots & & & & \\ 0 & 1 & & \ddots & & & \\ & & \ddots & & \ddots & & \\ & & & 1 & 0 & a & b \\ & & & 0 & 1 & -b & a \end{pmatrix}_{2m,2m}$$

. The case of index$(\lambda) = m = 1$ was treated in Section 9.3. In this section we try to find real solutions to $x' = Ax$ for eigenvalues $\lambda = a + bi$ of A whose indices exceed 1. In light of our earlier thoughts, the problem of solving $v'(t) = Jv(t)$ with $J = \operatorname{diag}(J_k(\lambda_i),\ J_p(a_j, b_j, 2m_j))$ with real eigenvalues λ_i and complex eigenvalues $\lambda_j = a_j + b_j i$, some of whose indices exceed one, again separates into individual Jordan block problems. Having just solved $w' = J(\lambda)w$ for one real or complex eigenvalue λ and one Jordan block, we need to solve $w'(t) = J(a,b,2m)w(t)$ for w with one 'real Jordan block' of size $2m$ by $2m$ with $m > 1$ and $b \neq 0$ now. To do so we look at the real Jordan block analogue to (14.6) :

$$w'(t) = \begin{pmatrix} w_1'(t) \\ \vdots \\ w_{2m}'(t) \end{pmatrix} = \begin{pmatrix} a & b & & & & & \\ -b & a & & & & & \\ 1 & 0 & \ddots & & & & \\ 0 & 1 & & \ddots & & & \\ & & \ddots & & \ddots & & \\ & & & 1 & 0 & a & b \\ & & & 0 & 1 & -b & a \end{pmatrix} \begin{pmatrix} w_1(t) \\ \vdots \\ w_{2m}(t) \end{pmatrix} . \qquad (14.8)$$

According to (9.10), the top 2 by 2 system $\begin{pmatrix} w_1'(t) \\ w_2'(t) \end{pmatrix} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \begin{pmatrix} w_1(t) \\ w_2(t) \end{pmatrix}$ in (14.8) has the real solutions $w_1(t) = ce^{at}\sin(bt)$ and $w_2(t) = ce^{at}\cos(bt)$ for $t \in \mathbb{R}$. This can be quickly verified by applying the differentiation rules for the exponential, the sine, and the cosine.

Let us try to solve the system of DEs (14.8) with $2m = 4$ first in order to find the general pattern for $w_1'(t)$, ..., $w_{2m}'(t)$ that will solve (14.8) . We consider

$$\begin{pmatrix} w_1'(t) \\ w_2'(t) \\ w_3'(t) \\ w_4'(t) \end{pmatrix} = \begin{pmatrix} a & b & 0 & 0 \\ -b & a & 0 & 0 \\ 1 & 0 & a & b \\ 0 & 1 & -b & a \end{pmatrix} \begin{pmatrix} w_1(t) \\ w_2(t) \\ w_3(t) \\ w_4(t) \end{pmatrix} . \qquad (14.9)$$

Here the first two component functions w_1 and w_2 are known as $w_1(t) = ce^{at}\sin(bt)$ and $w_2(t) = ce^{at}\cos(bt)$. Consequently w_3 and w_4 must satisfy the DEs

$$w_3' = w_1 + aw_3 + bw_4 \quad \text{and} \quad w_4' = w_2 - bw_3 + aw_4 .$$

Taking our lead from the complex case, we multiply w_1 and w_2 by t and set $w_3(t) := cte^{at}\sin(bt) = tw_1(t)$ and $w_4(t) := cte^{at}\cos(bt) = tw_2(t)$. Using the product rule of differentiation twice on w_3 we obtain

$$w_3'(t) \;=\; c[e^{at}\sin(bt) + ate^{at}\sin(bt) + bte^{at}\cos(bt)] \;=\; w_1 + aw_3 + bw_4$$

precisely as desired. Likewise for w_4'. This extends to real Jordan blocks $J(a, b, 2m)$ of arbitrary dimensions $2m$: Simply set the two components $w_{2\ell-1}$ and $w_{2\ell}$ of the solution $w(t) = \begin{pmatrix} w_1(t) \\ \vdots \\ w_{2m}(t) \end{pmatrix}$ of (14.8) equal to

$$\begin{aligned} w_{2\ell-1} &:= c\frac{t^{\ell-1}}{(\ell-1)!}e^{at}\sin(bt) \;=\; \frac{t}{\ell-1}w_{2\ell-3} \qquad \text{and} \\ w_{2\ell} &:= c\frac{t^{\ell-1}}{(\ell-1)!}e^{at}\cos(bt) \;=\; \frac{t}{\ell-1}w_{2\ell-2} \end{aligned} \tag{14.10}$$

for any $1 < \ell \leq \text{index}(\lambda) = m$.
Next we can synthesize the real solution $v(t)$ of $v'(t) = \text{diag}(J_k(\lambda_i),\; J_p(a_j, b_j, 2m_j))v(t)$ from its Jordan block DE parts and ultimately we find the real solution $x(t) := Uv(t)$ via the real principal vector matrix U that transforms $A \in \mathbb{R}^{n,n}$ to its real Jordan normal form $J = \text{diag}(J_k(\lambda_i),\; J_p(a_j, b_j, 2m_j)) = U^{-1}AU$ as before.

Example 9: Let us find a real solution to the linear system of differential equations $x'(t) = \begin{pmatrix} 4 & 7 & -1 & -6 \\ -2 & -3 & 2 & 3 \\ 0 & 1 & 2 & -1 \\ 0 & -1 & 0 & 1 \end{pmatrix} x(t) = Ax(t)$. Here the system matrix A that has been explored in Example 6 of Section 14.2.

A has the real Jordan normal form $J = \begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & 1 \end{pmatrix}$ for its complex conjugate eigenvalues $\lambda = 1 \pm i$, both of index 2. This form is achieved by the real matrix similarity $J = X^{-1}AX$ with $X = \begin{pmatrix} 3 & 1 & 3 & -1 \\ 0 & 0 & -1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix}$, see Example 6. Thus we have to deal with a special case of equation (14.9) for $\lambda = a + bi = 1 + i$ here:

$$\begin{pmatrix} w_1'(t) \\ w_2'(t) \\ w_3'(t) \\ w_4'(t) \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} w_1(t) \\ w_2(t) \\ w_3(t) \\ w_4(t) \end{pmatrix} .$$

This equation is solved by $w(t) = c\begin{pmatrix} e^t\sin t \\ e^t\cos t \\ te^t\sin t \\ te^t\cos t \end{pmatrix}$ according to (14.10) .

To obtain a real solution to $x'(t) = Ax(t)$, we transform the real Jordan nor-

mal form solution w to $x(t) = Xw(t) = c \begin{pmatrix} 3 & 1 & 3 & -1 \\ 0 & 0 & -1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} e^t \sin t \\ e^t \cos t \\ te^t \sin t \\ te^t \cos t \end{pmatrix} =$
$ce^t \begin{pmatrix} 3\sin t + \cos t + t(3\sin t - \cos t) \\ -t \sin t \\ \sin t + \cos t + t \sin t \\ \sin t - t \cos t \end{pmatrix}$.
It should be verified that this actually solves $x' = Ax$ by using calculus differentiation rules on the left and matrix times vector multiplication on the right, see Problem 1 below. ◀

Remark 3: Throughout this chapter we have chosen Jordan blocks J that are lower triangular (block) matrices. Some prefer to express Jordan normal forms via the upper triangular Jordan blocks $\begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix} = J_k^T = \begin{pmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{pmatrix} J_k \begin{pmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{pmatrix}$.
For such Jordan blocks a partial solution vector $w(t)$ to the linear DE $w' = J_k w$ should be replaced by its 'flipped brother' $\begin{pmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{pmatrix} w(t)$, concatenated with other flipped partial solutions, and then transformed to obtain the solution $x(t)$ of $x'(t) = Ax(t)$ via the different order principal vector matrix Y that transforms A to its upper triangular matrix representation J^T. This variation becomes obvious from the proof of Corollary 5 in Section 14.2. See also our example in Teacher Problem T 14 in Section 14.1.P. ◀

(b) $\mathbf{m^{th}}$ *order linear differential equations with constant coefficients*

Systems $x' = Ax$ of n linear differential equations involve the first derivative of the component functions in $x(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix} : \mathbb{R} \to \mathbb{R}^n$. Now we study one differential equation for one function $x(t) : \mathbb{R} \to \mathbb{R}$ that involves, however, several order derivatives of x.

To solve the $\mathbf{m^{th}}$ **order differential equation** with constant coefficients

$$x^{(m)}(t) \; + \; a_{m-1}x^{(m-1)}(t) \; + \; ... \; + \; a_1 x'(t) \; + \; a_0 x(t) \; = \; 0 \qquad (14.11)$$

for $x(t) : \mathbb{R} \to \mathbb{R}$ we introduce m auxiliary functions

$$y_1(t) := x(t),\ y_2(t) = x'(t),\ \ldots,\ \text{ and }\ y_m(t) = x^{(m-1)}(t)\ .$$

These functions allow us to rewrite (14.11) in terms of the vector valued auxiliary function $y(t) := \begin{pmatrix} y_1(t) \\ \vdots \\ y_m(t) \end{pmatrix}$ as follows:

$$\begin{aligned} y'(t) \;&=\; \begin{pmatrix} y_1'(t) \\ \vdots \\ y_m'(t) \end{pmatrix} \;=\; \begin{pmatrix} y_2(t) \\ \vdots \\ y_m(t) \\ -a_{m-1}y_m(t) - \ldots - a_0 y_1(t) \end{pmatrix} \\ &=\; \begin{pmatrix} 0 & 1 & 0 & & 0 \\ & & \ddots & & \\ & & & \ddots & \\ & & & & 1 \\ -a_0 & & \cdots & & -a_{m-1} \end{pmatrix} \begin{pmatrix} y_1(t) \\ \vdots \\ y_m(t) \end{pmatrix} \\ &=\; C(p)^T y(t)\ . \end{aligned}$$

This is a linear DE system with m differential equations. It has the transposed companion matrix $C(p)^T$ as its system matrix for the polynomial $p(r) = r^m + a_{m-1}r^{m-1} + \cdots + a_1 r + a_0$. Thus solving the m^{th} order differential equation (14.11) with constant coefficients reverts to solving the m by m system of linear DEs $y'(t) = C(p)^T y(t)$. This system can be solved by using the Jordan normal form approach, real if desired, on the system matrix $C(p)^T$ with the coefficients of p read off the m^{th} order equation (14.11) .

Example 10: To solve the third order differential equation $x'''(t) - x''(t) - 5x'(t) - 3x(t) = 0$ for $x(t) : \mathbb{R} \to \mathbb{R}$, we solve the following associated 3 by 3 system of linear DEs instead:

$$\begin{pmatrix} y_1' \\ y_2' \\ y_3' \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 3 & 5 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}\ .$$

The system matrix is the transposed companion matrix for $p(r) = r^3 - r^2 - 5r - 3$ read off directly from the given third order DE. $p(r)$ has the roots $\lambda_1 = -1$ (double) and $\lambda_2 = 3$ for the corresponding eigenvectors $x_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$ ($\lambda_1 = -1$) and $x_3 = \begin{pmatrix} 1 \\ 3 \\ 9 \end{pmatrix}$ ($\lambda_2 = 3$). We note that $C(p)^T$ is not diagonalizable, having but one eigenvector x_2 for its repeated eigenvalue -1. To find a principal vector of grade 2 for $\lambda_1 = -1$, we look at $\ker\big((A+I)^2\big)$. The matrix $(A+I)^2 = \begin{pmatrix} 1 & 2 & 1 \\ 3 & 6 & 3 \\ 9 & 18 & 9 \end{pmatrix}$ has $x_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ in its kernel. Next, $(A+I)x_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = x_2$ is the eigenvector of $C(p)^T$ found earlier for $\lambda_1 = -1$. Thus taking the three vectors $x_1,\ x_2$, and x_3 into a basis of $\mathbb{R}^3$ and forming their column matrix X, we know that $J = \begin{pmatrix} -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 3 \end{pmatrix} = X^{-1}C(p)^T X$ is the real Jordan normal form of $C(p)^T$. The solution of the Jordan normal form DE is $w(t) = \begin{pmatrix} ce^{-t} \\ cte^{-t} \\ ke^{3t} \end{pmatrix}$ for arbitrary independent constants c and k according to subsection (a). Since $X = \begin{pmatrix} | & & | \\ x_1 & \dots & x_3 \\ | & & | \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 3 \\ -1 & 1 & 9 \end{pmatrix}$ transforms $C(p)^T$ to J, the solution to the equivalent companion matrix form DE $y' = C(p)^T y$ is given by $y(t) = Xw(t) = \begin{pmatrix} ce^{-t}(1+t) + ke^{3t} \\ -cte^{-t} + 3ke^{3t} \\ ce^{-t}(t-1) + 9ke^{3t} \end{pmatrix}$.

And y's first component $x(t) = y_1(t) = ce^{-t}(1+t) + ke^{3t}$ solves the given third order differential equation.

Students should verify this result directly for the original third order differential equation using the rules of calculus. ◀

(c) ***Numerical computations of Jordan forms and MATLAB***

MATLAB – and for that matter, all mathematical software – is nearly *incapable* of computing the Jordan normal form, real or complex, of any non diagonalizable matrix A_{nn}, unless n is very small ($n \leq 4$ or 5) if A is dense, or unless A of larger size does break up into several such small dense diagonal blocks and unless A has integer or rational coefficients. This is due to two separate problems with finding a non diagonal Jordan normal form from A.

For one, non diagonalizable matrices must have repeated eigenvalues. All but the smallest sized, integer or rational entry matrices require numerical approximation to find their eigenvalues. There are no "algebraic formulas" such as the 'quadratic formula' for second degree polynomials in Appendices A and B for polynomials of degrees exceeding 4. Therefore, when the eigenvalues of a matrix A_{nn} have been computed approximately, who would be willing or able to decide correctly that the two computed eigenvalues $\lambda_1 = 1.000000000123$ and $\lambda_2 = 0.9999999976$, say, are in reality distinct eigenvalues of A, or that they approximate the 'true' double eigenvalue $\lambda = 1.0000000000438756$ of A?

Even if the first problem of eigenvalue multiplicity could be settled satisfactorily for a given matrix A, the next problem is to reliably determine the dimensions of the principal subspaces $P_j(\lambda) = \ker(A - \lambda I)^j$ for a repeated eigenvalue λ of A. Even the most simple question in this realm, namely whether a repeated eigenvalue λ has index 1 or larger, is generally impossible to decide since the rank of $A - \lambda I$ cannot be computed precisely. For sizable matrices A_{nn}, not even the SVD of Chapter 12 helps much, since it requires a *personal decision* whether the next to last computed singular value $\sigma_{n-1} = 1.23 \cdot 10^{-7}$ of $A - \lambda I$ for example should be taken to be zero, or whether only $\sigma_{n-1} = 2.0237 \cdot 10^{-15}$ should be, giving λ an index greater than 1. Of course most often, not even the smallest singular value σ_n of $A - \lambda I$ will be computed exactly as zero for any approximately computed eigenvalue λ of A, see the problem section.

For properly constructed low dimensional integer matrices $A_{n,n}$ with integer Jordan normal forms and bases (see the Teacher's Problem T 14 in Section 14.1.P), we can, however, use MATLAB successfully to simplify some of the tedious tasks of finding eigenvalues and Jordan bases for A:
First, we can compute the eigenvalues λ_i of A in MATLAB via `eig(A)`. If there are no repeats, the Jordan normal form of A is $\text{diag}(\lambda_i) \in \mathbb{C}^{n,n}$. And the real Jordan normal form of $A \in \mathbb{R}^{n,n}$ is $\text{diag}(\lambda_\ell, J(Re(\lambda_k), Im(\lambda_k), 2))$ where the λ_ℓ denote the real eigenvalues of A and the λ_k are selected from each pair of complex conjugate eigenvalues $\lambda_j = \overline{\lambda_j}$ of A. An eigenvector basis for A can best be found from `rref(A-λ*eye(n))`. For a real eigenvalue λ we then solve the homogeneous system $A - \lambda I = 0$ according to Chapter 3 by hand. Invoking `null(A-λ*eye(n))` instead in MATLAB gives us an orthonormal basis for the kernel and is generally useless for our purpose to work within the integers. For a complex eigenvalue we also solve $A - \lambda I = 0$ from the complex RREF that was computed by MATLAB as before. The complex eigenvalue and its conjugate serve in the eigenvector basis for A. If we look for the real Jordan normal form instead for a real matrix A, then we take the real and the imaginary parts vectors into the Jordan basis for A.
If A has a repeated eigenvalue λ with algebraic multiplicity $k > 1$ according to the result of `eig(A)`, we need to compute the RREFs of $(A - \lambda I)^\ell$ for $\ell = 1, 2, 3, ..., k$ via MATLAB until the number of free variables in these does not increase any longer. This gives us the index m of λ. Next we pick a vector $x^{(m)} \in P_m(\lambda) - P_{m-1}(\lambda)$ and iterate down to the eigenvector $x^{(1)} = (A - \lambda I)^{m-1} x^{(m)}$. This defines a maximal length principal vector chain for λ. We might have to repeat this process for other principal vector chains for λ or A, but remember that the size of A is rather small in our examples that are fit for hand computations. The computed principal vector chain(s) form part of the possibly complex Jordan basis of A. If $A \in \mathbb{R}^{n,n}$ and $\lambda \in \mathbb{C}$, then the real and imaginary parts vectors of each $x^{(j)}$ above should be chosen instead to obtain the desired real Jordan basis of A. These vectors generate the real Jordan block $J(Re(\lambda), Im(\lambda), 2m)$ associated with the pair λ and $\overline{\lambda}$ of index m.
This process leads to the standard or the real Jordan normal form of any properly constructed $A \in \mathbb{R}^{n,n}$ over the integers and to its respective Jordan basis $\mathcal{X}$.

If the Jordan form J of A has been computed as a first step in solving the linear differential equation $x'(t) = Ax(t)$, we solve the Jordan normal form DE $w'(t) = Jw(t)$ using our formulas (14.7) for real eigenvalues and (14.10) for complex ones. And from X, the column vector matrix of the Jordan basis that transforms A to $J = X^{-1}AX$, we then find the solution $x'(t) = Xw(t)$ for $x'(t) = Ax(t)$. These latter steps are best carried out by hand. This process is relatively easy if A is properly constructed according to Teacher's problem T 14; see our exercises below and the solution set.

However in general, unless A has well separated eigenvalues and thus is diagonalizable, the Jordan structure of a general matrix A cannot be derived reliably from numerical computations. The main purpose of this chapter is to give us a theoretical tool that lets us glimpse the similarity invariants of matrices. The Jordan normal form helps us classify matrices and lets us study low dimensional linear DE systems and find their theoretical solutions.
There is, however, some justice. Probabilisticly speaking, every random matrix A_{nn} must have n distinct eigenvalues and therefore every random square matrix is diagonalizable with probability 1. The non diagonalizable matrices form a set of measure zero in matrix space $\mathbb{R}^{n,n}$ and our picture of matrix space in Figure 11.1 of Section 11.2 surely gives too much territory to the non diagonalizable matrices such as J on the right side of the Figure. For random matrices $A \in \mathbb{R}^{n,n}$ or $\mathbb{C}^{n,n}$, the assignment

$$A \longrightarrow \Lambda(A) := \{\lambda_i \mid \lambda_i \text{ is an eigenvalue of } A\} \subset \mathbb{C}$$

of A to its eigenvalue set $\Lambda(A)$ must create random (except for real axis symmetry, if A is real) n points as images, potentially with repetitions. Repetition of an eigenvalue, however, is unlikely since when making n marks randomly in the complex plane, it is unlikely to mark the same spot twice. Incidentally, for the same reason it is highly unlikely that a random matrix will be singular, or have the precise eigenvalue zero.

There is one drawback to these heuristic probability considerations in what amounts to our willful 'human touch'. Through our studies we have become acquainted with many matrices with repeated eigenvalues such as I_n or O_n. We can write down non diagonalizable Jordan normal forms with ease. And dense matrices whose eigenvalues do not repeat seemingly take an effort to construct. We simply are not good random matrix generators. But nature is.

14.3.P Problems

1. (a) Verify that the solution $x(t)$ computed in Example 9 satisfies the DE $x'(t) = Ax(t)$.
   (b) Repeat for Example 10.
2. For a number of randomly generated 5 by 5 matrices, test whether the smallest singular value $\sigma_n(\lambda_i)$ of $A - \lambda_i I$ is actually zero in MATLAB for any of the eigenvalues $\lambda_1, ..., \lambda_n$ of A.
3. Generate ten random 10 by 10 and 100 by 100 matrices each and determine whether the matrix is singular or not. Do so by using (a) the `rref`, (b) the `eig`, and (c) the `svd` functions of MATLAB. Do you obtain consistent results?
4. Repeat the previous problem for ten random 10 by 10 and 100 by 100 matrices each, that you doctor so that their last row is the sum of the first three rows of the given random matrix. Do you obtain consistent results across the three MATLAB functions in Problem 3?

5. Repeat Problem 3 for the matrices A and B of the Example in Teacher Problem T 14 in Section 14.1.P. Use the MATLAB computed eigenvalues of A and B, as well as the theoretically accurate ones and compare.

6. Let $A_c := \begin{pmatrix} c & -1 & 0 \\ 1 & c & 1 \\ 0 & 1 & c \end{pmatrix}$.
   (a) For which value of c does MATLAB compute the eigenvalues of A_c as repeated and for which values of c as distinct?
   (b) For which value of c are the eigenvalues of A_c repeated or distinct?

7. Find the real solution of the DE $x'(t) = Ax(t)$ with $A = \begin{pmatrix} 1 & -6 & -3 & -1 \\ 4 & 7 & 12 & -3 \\ -3 & -3 & -7 & 2 \\ -2 & -5 & -6 & 1 \end{pmatrix}$ from Problem 1 of Section 14.2.P.

8. Find the real Jordan normal form of $B = \begin{pmatrix} 3 & -3 & -1 & 8 \\ 1 & -2 & -7 & 25 \\ -3 & -1 & -8 & 29 \\ 0 & -1 & -3 & 11 \end{pmatrix}$.
(Hint: B has the eigenvalue $\lambda = 1 - 2i$.)

9. Find the real solution of the DE $x'(t) = Bx(t)$ with B from the previous problem.

10. Convert the 4^{th} order DE $x^{(4)}(t) - 6x^{(3)} + 3x'(t) = 4x(t)$ into a system of linear DEs.

11. Solve the 5^{th} order DE $y^{(5)}(t) - 5y^{(4)}(t) + 5y^{(3)}(t) + 5y^{(2)}(t) - 6y'(t) = 0$.

12. Solve the 5^{th} order DE $x^{(5)}(t) + x^{(4)}(t) - 9x^{(3)}(t) - 5x^{(2)}(t) + 16x'(t) + 12x(t) = 0$.

14.R Review Problems

1. For the complex or real Jordan normal form J of a matrix A prove:
   (a) $\text{rank}(J) = \text{rank}(J^T)$;
   (b) $\det(J) = \det(A)$.
   (c) $f_J = f_{J^T}$ for the characteristic polynomial $f_{..}$.
   (d) $p_J = p_{J^T}$ for the minimum polynomial $p_{..}$.
   (e) $f_A = f_{A^T}$.
   (f) $p_A = p_{A^T}$.
   (g) $f_A = f_J$.
   (h) $p_A = p_J$.

2. Show that the number of Jordan blocks for one eigenvalue λ of A is equal to the geometric multiplicity of the eigenvalue λ.
How is the geometric multiplicity of an eigenvalue defined?

3. (a) Show that the index of an eigenvalue λ of A is equal to the size of the largest Jordan block for λ in the Jordan normal form of A.
   (b) Should the above result be modified when talking about the real Jordan normal form of A instead? How?

4. Rephrase and reprove the previous problem for the <u>real</u> Jordan normal form of $A \in \mathbb{R}^{n,n}$.

5. If $x \in \mathbb{R}^n$ is a principal vector of grade $k > 2$ for $\lambda \in \mathbb{R}$ and A, show that $y := (A - \lambda I)^2 x$ is a principal vector of grade $k - 2$ for λ and A.

6. For which real square matrices A do the Jordan normal form and the real Jordan normal form coincide?

7. Which matrices have a diagonal companion matrix normal form?

8. (a) Construct a matrix A with the characteristic polynomial $f_A(\lambda) = \lambda(\lambda + 2)^2(\lambda - 3)^5$. What is A's size?
   (b) Construct a matrix A with the minimum polynomial $p_A(x) = (x + 2)^2(x - 3)^5$. What are A's possible sizes?

9. Construct 7 by 7 and 12 by 12 matrices A and B, respectively, that have the minimum polynomial $x^3(x+2)^4$.

10. Find the minimum and the characteristic polynomials $p_A(x)$ and $f_A(x)$ for a matrix A with the Jordan normal form $J = \text{diag}(J(2,4),\ J(2,3),\ J(2,1),\ J(-1,3),\ J(-1,5),\ J(0,1))$.

11. What is the minimal size of a matrix B whose minimum polynomial is that of the matrix J in the previous problem? What is the possible maximal size of such a B?

12. Assume that $A \in \mathbb{R}^{n,n}$ has the Jordan normal form $J = \text{diag}(J_i) \in \mathbb{C}^{n,n}$.

    (a) Show that A^2 is similar to J^2.

    (b) For a Jordan block $J_i = J(\lambda, k) \in \mathbb{C}^{n,n}$ compute J_i^2.

    (c) What are the eigenvalues and eigenvectors of $J_i(\lambda, k)^2$?
    (Careful: there are two different answers for $\lambda = 0$ and $\lambda \neq 0$.)

13. Find the Jordan normal form of A^2 if A has the Jordan normal form $J = \text{diag}(J_i)$.
    (Hint: Make use of the previous problem.)

Standard Questions and Tasks:

1. What is a principal subspace and a principal vector for a square matrix?
2. What does the grade of a principal vector indicate?
3. What is the maximal grade that is possible for a principal vector and an eigenvalue of index j? Must such a maximal grade principal vector exist?
4. Find a principal vector chain of maximal order for a given eigenvalue of a matrix A.
5. When is a matrix A_{nn} diagonalizable, when is it not?
6. Find the Jordan normal form of a given matrix.
7. Write out the Jordan diagram for a given sequence of principal subspace dimensions for one eigenvalue λ of a given matrix A_{nn} and determine the Jordan blocks associated with λ .

Subheadings of Lecture Fourteen :

Basic Equations :

p. W-2	$P_k(\lambda) = \ker(A - \lambda I)^k$	(Principal subspace)
p. W-3	$x^{(k)} \in P_k(\lambda)$ with $x^{(k)} \notin P_{k-1}(\lambda)$	(Principal vector of grade k)
p. W-3	$x^{(k)},\ x^{(k-1)} = (A - \lambda I)x^{(k)},\ \ldots$ $\ldots,\ x^{(1)} = (A - \lambda I)^{k-1}x^{(k)}$	(Principal vector chain)

Basic Notions :

p. W-2	Index of an eigenvalue of A
p. W-2	Order of a principal vector chain or of a principal subspace $P_k(\lambda)$
p. W-3	Grade of a principal vector in $P_k(\lambda)$ and not in $P_{k-1}(\lambda)$

Index

Frank Uhlig

Department of Mathematics
Auburn University
Auburn, AL 36849–5310

tla@auburn.edu

http://www.auburn.edu/~uhligfd