Chapter 9 Jordan Block Matrices

In this chapter we will solve the following problem. *Given a linear operator T find a* basis R of \mathbf{F}^n such that the matrix $M_R(T)$ is as simple as possible. Of course "simple" is a matter of taste. Here we consider a matrix simple if it is as close as possible to diagonal. Unfortunately, not every matrix is similar to a diagonal one. We will introduce and study the next best thing, the Jordan block matrix. But first we investigate the optimistic case: there exists a basis R such that $A = M_R(T)$ is a diagonal matrix, i.e. $a_{ij}=0$ if $i \neq j$. What is so special about this basis? Let us denote vectors of R by $v_1, v_2, ..., v_n$. Then, for each *i* we have $T(v_i) = a_{ii}v_i$.

Definition 9.1. A scalar λ is called an *eigenvalue* of T iff there exists a nonzero vector v such that T(v)= λ v. Every such vector is called an *eigenvector* of T belonging to λ .

<u>Theorem 1.1.</u> The matrix $M_R(T)$ is diagonal iff R is a basis consisting entirely of eigenvectors of T.

It seems that eigenvalues and eigenvectors have an important part to play in this theory. How do we find eigenvalues and eigenvectors for an operator? We simply follow the definition, except that we use the matrix of T in a basis R of our choice, in place of T itself. For example if R is the standard basis for \mathbf{F}^n and A is the matrix of T in R then T(x)=Ax. The equation $T(x)=\lambda x$ is then equivalent to $Ax=\lambda x=\lambda Ix$. This yields $(A-\lambda I)x=\Theta$. Thus λ is an eigenvalue for T (or for A) iff the homogeneous system of equations $(A-\lambda I)x=\Theta$ has a nonzero solution (beside the zero solution, which is always there). Every such solution is then an eigenvector of T belonging to the eigenvalue λ . Since solutions of a homogeneous system of linear equations form a vector space we get that all eigenvector itself) form a subspace of \mathbf{F}^n , called an eigenspace. Nonzero solutions exist if and only if the rank of A is less than *n* - the number of columns of A. This happens if and only if det $(A-\lambda I)=0$. It is easy to see that the function det $(A-\lambda I)$ is a polynomial of degree *n* in variable λ . It is called the *characteristic polynomial* of the operator T (and of the matrix A). Hence we have proved

Theorem 9.1. A scalar *t* is an eigenvalue for A iff *t* is a root of the characteristic polynomial of A. \Box

Theorem 9.2. If $\lambda_1, \lambda_2, ..., \lambda_n$ are pairwise different eigenvalues for T and for each i, $L_i = \{v_{i,1}, v_{i,2}, ..., v_{i,k(i)}\}$ is a linearly independent set of eigenvectors belonging to λ_i then $L_1 \cup L_2 \cup ... L_n$ is linearly independent.

Roughly speaking, the theorem states that *eigenvectors belonging to different eigenvalues are linearly independent*.

Proof. First, suppose that every L_i is a one-element set. For simplicity we will write $L_i = \{v_i\}$. We prove our case by induction. For n=1 there is nothing to prove. Suppose the theorem holds for some *n* and consider the condition $\sum_{i=1}^{n+1} a_i v_i = \Theta$. Applying T to both sides we get

$$\sum_{i=1}^{n+1} a_i T(v_i) = \sum_{i=1}^{n+1} a_i \lambda_i v_i = \Theta, \text{ while scaling both sides by } \lambda_{n+1} \text{ yields } \sum_{i=1}^{n+1} a_i \lambda_{n+1} v_i = \Theta. \text{ Subtracting}$$

one from the other we get $\sum_{i=1}^{n+1} (\lambda_i - \lambda_{n+1}) a_i v_i = \sum_{i=1}^n (\lambda_i - \lambda_{n+1}) a_i v_i = \Theta$ because for *i*=n+1,

 $\lambda_i - \lambda_{n+1} = 0$. By the induction hypothesis, this implies $(\lambda_i - \lambda_{n+1})a_i = 0$ for i = 1, 2, ..., n. Since the eigenvalues are assumed to be pairwise different, we get $a_1 = a_2 = ... = a_n = 0$. This and

$$\sum_{i=1}^{n+1} a_i v_i = \Theta \text{ imply that } a_{n+1} v_{n+1} = \Theta, \text{ hence } a_{n+1} = 0.$$

To conclude the proof suppose that $|L_i| = k(i)$ and $\sum_{i=1}^{n} \sum_{j=1}^{k(i)} a_{i,j} v_{i,j} = \Theta$. For each *i* denote

 $w_i = \sum_{j=1}^{k(i)} a_{i,j} v_{i,j}$. Each vector w_i , being a linear combination of eigenvectors belonging to the same eigenvalue λ_i is either equal to Θ or is itself an eigenvector belonging to λ_i . If some of them, say $w_{i(1)}, w_{i(2)}, \dots, w_{i(p)}$ are in fact eigenvectors then $\sum_{i=1}^{n} \sum_{j=1}^{k(i)} a_{i,j} v_{i,j} = \sum_{i=1}^{n} w_i = \sum_{t=1}^{p} w_{i(t)} = \Theta$ is a linear combination of eigenvectors belonging to different eigenvalues. Hence, by the first part of the proof, all coefficients of the linear combination $\sum_{t=1}^{p} w_{i(t)} = \Theta$ are zeroes, while everybody can clearly see that they are ones. This is impossible since in every field $1 \neq 0$. Hence all $w_i = \Theta$. Since each set L_i is linearly independent, all $a_{i,j}$ are equal to 0.

Example 9.1. Find a diagonal matrix for T(x,y,z) = (3x+y-z,2x+4y-2z,x+y+z).

The matrix A of T with respect to the standard basis of \mathbf{R}^3 is $\begin{bmatrix} 3 & 1 & -1 \\ 2 & 4 & -2 \\ 1 & 1 & 1 \end{bmatrix}$. The characteristic polynomial of A is $\Delta_A(\lambda) = -\lambda^3 + 8\lambda^2 - 20\lambda + 16 = (2-\lambda)^2(4-\lambda)$. For $\lambda = 2$ we get A- $\lambda \mathbf{I} = \mathbf{A} - 2\mathbf{I} = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 2 & -2 \\ 1 & 1 & -1 \end{bmatrix}$. The rank of A-2I is obviously 1 and we can easily find a basis for the solution space, namely $\{(1, -1, 0), (0, 1, 1)\}$. For $\lambda = 4$ we get A- $\lambda \mathbf{I} = \mathbf{A} - 4\mathbf{I} = \begin{bmatrix} -1 & 1 & -1 \\ 2 & 0 & -2 \\ 1 & 1 & -3 \end{bmatrix}$. The rank of A-4I is 2 and we can chose (1,2,1) as an eigenvector belonging to 4. The set $\mathbf{R} = \{(1, -1, 0), (-1, 0), (-1, 0), (-1, 0), (-1, 0), (-1, 0), (-1, 0)\}$.

(0,1,1),(1,2,1) is a basis for \mathbf{R}^3 and $M_R(T) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}$.

Unfortunately, not every matrix can be diagonalized. For some matrices (operators) there is no basis consisting of eigenvectors. In those cases the next best thing is the Jordan block matrix.

Definition 9.2. A *block matrix* is a matrix of the form

 $A = \begin{bmatrix} B_1 & 0 & \dots & 0 \\ 0 & B_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & B_s \end{bmatrix}$ where each B_i is a square matrix, the diagonal of B_i is a part of the

diagonal of A and all entries outside blocks B_{i} are zeroes.

Definition 9.3. A *Jordan block* of size k and with diagonal entry λ is the *k*×*k* matrix

	Γλ	1	0			0
B=	0	λ	1	0	•••	0
	:	·.	·.	·.	·.	:
	÷		·.	·.	·.	0
				·.	λ	1
	0				0	λ

Definition 9.4. A *Jordan block matrix* is a matrix of the form

$$J = \begin{bmatrix} B_1 & 0 & \dots & 0 \\ 0 & B_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & B_s \end{bmatrix}$$
where each B_i is a Jordan block.
$$Example 9.2. J = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}$$
is a Jordan block matrix with three Jordan blocks.

One of size 1 with diagonal entry 2, one of size 2 with diagonal entry 1, and one of size 3 with diagonal entry 2.

<u>**Theorem 9.3.**</u> (Jordan) Suppose $T: \mathbf{F}^n \to \mathbf{F}^n$ and T has *n* eigenvalues (counting multiplicities). Then there exists a basis R for \mathbf{F}^n such that $J=M_R(T)$ is a Jordan block matrix. **Proof substitute.** We will describe (working backwards) how to construct such a basis R= $\{v_1, \dots, v_n\}$ studying properties of its (hypothetical) vectors. Suppose that the size of B₁, the first Jordan block, is k and the diagonal entry is λ . Then, by the definition of the matrix of a linear operator with respect to R and by the definition of a Jordan block matrix, we get $T(v_1)$ $=\lambda v_1 + 0v_2 + ... + 0v_n = \lambda v_1$, i.e. v_1 is an eigenvector for λ . This can be expressed differently as $(T-\lambda id)v_1 = \Theta$. Since the second column of J consists of 1, λ and all zeroes, we get that $T(v_2)=1v_1+\lambda v_2$. In other words, $(T-\lambda id)v_2=v_1$. Extending this argument we get that for each v_1 from the first k vectors from R we have $(T-\lambda id)v_i = v_{i-1}$, except that for v_1 w have $(T-\lambda id)v_1 = \Theta$. Vectors $v_2, ..., v_k$ are called vectors *attached* to the eigenvector v_1 , of the order 1,2, ..., k-1, respectively. Now the structure of R becomes clear. It consists of several bunches of vectors. The number of the bunches is the number of Jordan blocks in J, each bunch is lead by an eigenvector and followed by its attached vectors of orders 1,2 and so on. It is all very nice in the hindsight, knowing what the matrix J looks like, but how do we actually find those vectors? Or at least how do we find the matrix J? Well, kids, here comes the story.

- Given a linear operator, you should find its matrix A in a basis of your choice, most likely your choice will be the standard basis S.
- (2) Having found the matrix A, find all eigenvalues solving the characteristic equation of T, i.e. det(A-λI)=0.

- (3) For each eigenvalue λ_i find the number of Jordan blocks with this particular eigenvalue as the diagonal entry. This is, of course, equal to the maximum number of linearly independent eigenvectors belonging to λ_i, in other words the dimension of the solution space of the system of equations (A-λ_iI)X=Θ, in other words *n*-rank(A-λ_iI).
- (4) Calculate sizes of Jordan blocks with λ_i as the diagonal entries using the formula: rank(A- $\lambda_i I$)^k-rank(A- $\lambda_i I$)^{k+1} is the number of λ_i -Jordan blocks of the size at least k+1. It follows from the fact that ranks of matrices (A- $\lambda_i I$)^k and (J- $\lambda_i I$)^k are the same (because the matrices are similar) and that

0	1	0	•••	0	0	0	1	0	•••	0	0		0	0	1	•••	0	0	
0	0	1		÷	0	0	0	1		÷	0		0	0	0	·.	÷	0	
0	0	·.	·.	0	0	0	0	·.	·.	0	0		0	0	·.	·.	1	0	1 41
0	0		۰.	1	0	0	0		·.	1	0	=	0	0		·.	0	1	, and with
0	0			0	1	0	0			0	1		0	0			0	0	
0	0			0	0	0	0	•••		0	0		0	0			0	0	

each next power of the matrix, the line of ones goes one position to the right losing one 1, until there is nothing left to loose. Hence ranks of all used-to-be- λ_i blocks of the matrix J- λ_i I go down by one with each consecutive multiplication by J- λ_i I, except for those, whose ranks have already reached 0. The blocks corresponding to other eigenvalues retain their ranks.

Hence we have:

n-rank(A- λ_i I) λ_i Jordan blocks of all sizes, rank(A- λ_i I)- rank(A- λ_i I)² λ_i Jordan blocks of size at least 2, rank(A- λ_i I)²- rank(A- λ_i I)³ λ_i Jordan blocks of size at least 3, and so on.

(5) Repeat steps (1)-(4) for each eigenvalue $\lambda_1, \lambda_2, ..., \lambda_n$.

To find the basis R, we find first attached vectors of the highest order, and then transform them by A- $\lambda_i I$ till we get eigenvectors. Assuming that the largest order of an attached vector is k-1, we find the vector v_k choosing a solution to $(A-\lambda_i I)^k X=\Theta$ that satisfies $(A-\lambda_i I)^{k-1} X \neq \Theta$. Then we put $v_{k-1}=(A-\lambda_i I)v_k$, $v_{k-2}=(A-\lambda_i I)v_{k-1}$, ..., $v_1=(A-\lambda_i I)v_2$. If we have more than one λ_i Jordan block then we have to take care when we choose our attached vectors so that they and their eigenvectors are linearly independent. This may be tricky and you must be extra careful here. **<u>Problem</u>**. Find a basis R for \mathbf{R}^4 such $M_R(T)$ is a Jordan block matrix J, where

$$T(x,y,z,t)=(-x-y-2z+t,2x+y+3z-t,2x+4z-t,2x-2y+5z)$$

Solution. First we form A - the matrix for T in the standard basis of \mathbf{R}^4 . Obviously

$$A = \begin{bmatrix} -1 & -1 & -2 & 1 \\ 2 & 1 & 3 & -1 \\ 2 & 0 & 4 & -1 \\ 2 & -2 & 5 & 0 \end{bmatrix}$$
. Next we calculate $det(A - \lambda I) = det\begin{bmatrix} -1 - \lambda & -1 & -2 & 1 \\ 2 & 1 - \lambda & 3 & -1 \\ 2 & 0 & 4 - \lambda & -1 \\ 2 & -2 & 5 & -\lambda \end{bmatrix} =$
$$(r_2 - r_3) = (1 - \lambda) det\begin{bmatrix} -1 - \lambda & -1 & -2 & 1 \\ 0 & 1 & -1 & 0 \\ 2 & 0 & 4 - \lambda & -1 \\ 2 & -2 & 5 & -\lambda \end{bmatrix} =$$
$$= (r_3 + c_2) = (1 - \lambda) det\begin{bmatrix} -1 - \lambda & -1 & -3 & 1 \\ 0 & 1 & 0 & 0 \\ 2 & 0 & 4 - \lambda & -1 \\ 2 & -2 & 3 & -\lambda \end{bmatrix} = (1 - \lambda) det\begin{bmatrix} -1 - \lambda & 0 & 1 - \lambda \\ 0 & 1 - \lambda & \lambda -1 \\ 2 & 3 & -\lambda \end{bmatrix} = (1 - \lambda) det\begin{bmatrix} 1 - \lambda & 0 & 1 - \lambda \\ 0 & 1 - \lambda & \lambda -1 \\ 2 & 3 & -\lambda \end{bmatrix} = (1 - \lambda)^3 det\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 2 & 3 & -\lambda \end{bmatrix} = (1 - \lambda)^4.$$
 Now we must
calculate rank(A-I) = rank
$$\begin{bmatrix} -2 & -1 & -2 & 1 \\ 2 & 0 & 3 & -1 \\ 2 & 0 & 3 & -1 \\ 2 & -2 & 5 & -1 \end{bmatrix} = (r_2 + r_1, r_3 + r_1, r_4 + r_1) =$$
$$rank\begin{bmatrix} -2 & -1 & -2 & 1 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -3 & 3 & 0 \end{bmatrix} = (r_3 - r_2, r_4 - 3r_2) = rank\begin{bmatrix} -2 & -1 & -2 & 1 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = 2.$$
 This means that J has

two blocks with diagonal entries 1, but sizes of the blocks may be 2×2 and 2×2 or 1×1 and 3×3 . Now we must calculate ranks of matrices A-I, $(A-I)^2$ and so on. It turns out that $(A-I)^2$ is the zero matrix, so its rank is 0. By part (4) of our algorithm we get that J has 2 blocks of size

at least 2 each, that is J has 2 blocks of size 2 by 2. Hence
$$J = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
. Thus our basis

R consists of two eigenvectors v_1 and v_3 and their attached vectors v_2 and v_4 . The attached

vectors are solutions of $(A-I)^2X=\Theta$, that do not satisfy $(A-I)X=\Theta$. Since $(A-I)^2$ is the zero matrix, the first system of equations is trivial ($\Theta=\Theta$), so the only condition v_2 and v_4 must satisfy is $(A-I)X\neq\Theta$. We can choose $v_2=(1,0,0,0)$ and $v_4=(0,1,0,0)$ getting $v_1=(-2,2,2,2)$ and $v_3=(-1,0,0,-2)$. These vectors form the columns of the change-of-basis matrix P such that J =

P⁻¹AP, i.e.
$$P = \begin{bmatrix} -2 & 1 & -1 & 0 \\ 2 & 0 & 0 & 1 \\ 2 & 0 & 0 & 0 \\ 2 & 0 & -2 & 0 \end{bmatrix}$$
.