Block matrices

Def. Let $A \in K^{m\times n}$ be a matrix, and $m = m_1 + \ldots + m_r$, $n = n_1 + \ldots + n_s$ decomposition of m and n into a sum of positive integers. We divide the matrix into horizontal bands of $m_1, m_2, \ldots$ rows, and then we divide these bands vertically to matrices of $n_1, n_2, \ldots$ columns. Then we get an $r \times s$ matrix whose elements are also matrices.

The sum of matrices of equal sizes and block decompositions:

$$\begin{bmatrix} A_{11} & \ldots & A_{1s} \\ \ldots & \ldots & \ldots \\ A_{r1} & \ldots & A_{rs} \end{bmatrix} + \begin{bmatrix} B_{11} & \ldots & B_{1s} \\ \ldots & \ldots & \ldots \\ B_{r1} & \ldots & B_{rs} \end{bmatrix} = \begin{bmatrix} A_{11}+B_{11} & \ldots & A_{1s}+B_{1s} \\ \ldots & \ldots & \ldots \\ A_{r1}+B_{r1} & \ldots & A_{rs}+B_{rs} \end{bmatrix}.$$

The product of two block matrices with matching sizes and block decompositions (that is, if $A \in K^{\ell\times m}$ and $B \in K^{m\times n}$, where m is decomposed the same way in the block structure of A and B)

$$A = \begin{bmatrix} A_{11} & \ldots & A_{1r} \\ \ldots & \ldots & \ldots \\ A_{p1} & \ldots & A_{pr} \end{bmatrix} \quad \cdot \quad \begin{bmatrix} B_{11} & \ldots & B_{1s} \\ \ldots & \ldots & \ldots \\ B_{r1} & \ldots & B_{rs} \end{bmatrix} = C$$

where $C_{ij} = \sum_{t=1}^{m} A_{it}B_{tj}$. (Since we have matching decompositions, the products $A_{it}B_{tj}$ exist and can be added for $t = 1, \ldots, r$)

Example: The product $AB = \begin{bmatrix} 0 & 0 & | & 1 & 0 \\ 0 & 0 & | & 0 & 1 \\ - & - & - & - & - \\ -1 & 0 & | & 1 & 0 \\ 0 & -1 & | & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 3 & -1 \\ - & - \\ 0 & 1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 2 & 1 \\ - & - \\ -1 & -1 \\ -1 & 2 \end{bmatrix}$

can be calculated easier, if we consider A and B as block matrices with 2×2 blocks:

$$AB = \begin{bmatrix} 0 & I \\ -I & I \end{bmatrix} \cdot \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} B_2 \\ -B_1 + B_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 2 & 1 \\ - & - \\ -1 & -1 \\ -1 & 2 \end{bmatrix}.$$

Corollary: The product of block diagonal matrices (that is, $n \times n$ matrices divided along the same decomposition of n, and having only zero matrices in their non-diagonal positions) can be calculated by multiplying the corresponding diagonal elements:

$$\begin{bmatrix} A_1 & 0 & \ldots & 0 \\ 0 & A_2 & \ldots & 0 \\ & & \ddots & \\ 0 & 0 & \ldots & A_n \end{bmatrix} \cdot \begin{bmatrix} B_1 & 0 & \ldots & 0 \\ 0 & B_2 & \ldots & 0 \\ & & \ddots & \\ 0 & 0 & \ldots & B_n \end{bmatrix} = \begin{bmatrix} A_1B_1 & 0 & \ldots & 0 \\ 0 & A_2B_2 & \ldots & 0 \\ & & \ddots & \\ 0 & 0 & \ldots & A_nB_n \end{bmatrix}.$$

The Jordan normal form

Def. Jordan block: $\begin{bmatrix} \lambda & 1 & \dots & 0 \\ 0 & \lambda & \ddots & 0 \\ 0 & \dots & \ddots & 1 \\ 0 & \dots & 0 & \lambda \end{bmatrix}$

(Note that its only eigenvalue is λ, however the eigenspace is only 1 dimensional).

Jordan matrix: a block diagonal matrix whose diagonal matrices are Jordan blocks.

Exercise: Calculate the characteristic polynomial, minimal polynomial and the dimension of the eigenspace for the 4×4 Jordan block corresponding to the eigenvalue 2.

Proposition: Let $J \in K^{n\times n}$ be a Jordan block with eigenvalue λ. Then $k_A(x) = (-1)^n(x-\lambda)^n$ and $m_A(x) = (x-\lambda)^n$.

Proof: Since J is an upper triangular matrix, the first statement is obvious. As for the second, let us notice that $N := A - \lambda I$ is a Jordan block with eigenvalue 0, and it acts on the basis vectors in the following way: $\mathbf{b}_n \mapsto \mathbf{b}_{n-1} \mapsto \cdots \mapsto \mathbf{b}_1 \mapsto \mathbf{0}$. Then $N^k : \mathbf{b}_i \mapsto \mathbf{b}_{i-k}$ for $i > k$ and $\mathbf{b}_i \mapsto \mathbf{0}$ for $i \leq k$. This means that N^k has only a skew row of $1's$ parallel to the diagonal, starting at the position $(1, k+1)$. Thus $N^{n-1} = E_{1n} \neq 0$, but $N^n = 0$, showing that the minimal polynomial of A is $m_A(x) = (x-\lambda)^n$.

Corollary: If the different diagonal elements of an $n \times n$ Jordan matrix J are $\lambda_1, \dots, \lambda_r$ with multiplicities $a_1, \dots, a_k$ respectively, then the characteristic polynomial of the matrix is $(-1)^n(x-\lambda_1)^{a_1}\cdots(x-\lambda_k)^{a_k}$, and its minimal polynomial is $(x-\lambda_1)^{b_1}\cdots(x-\lambda_k)^{b_k}$, where for each i, the largest λ_i-block is of size b_i (meaning that it is a $b_i \times b_i$ matrix). Furthermore, the dimension of the λ_i-eigenspace is the number of λ_i-blocks.

Proof: The statement about the characteristic polynomial is clear, since the Jordan matrix is an upper triangular matrix.

For a polynomial $p(x)$, $p(J) = 0$ if and only if $p(J_t) = 0$ for every diagonal block J_t. But for a λ_i-block J_t, the matrix $J_t - \lambda_j I$ is invertible for $\lambda_j \neq \lambda_i$, so for the largest λ_i-block (of size b_i) to become 0, $(x - \lambda_i)$ should be at least on the power of b_i by the previous proposition, and the b_i'th power is clearly sufficient.

Finally, $J - \lambda_i I$ will be in row echelon form if we move the zero rows to the bottom, and the number of zero rows is exactly the number of λ_i-blocks (let that be d_i). Then $\operatorname{rank}(J - \lambda_i I) = n - d_i$, and by the dimension theorem, $\dim V_{\lambda_i} = \dim \operatorname{Ker}(J - \lambda_i I) = d_i$.

Jordan's theorem

Let $A \in K^{n\times n}$, and suppose that the characteristic polynomial of A can be factored into a product of linear polynomials over K, that is, $k_A(x) = (-1)^n(x-\lambda_1)^{a_1}\cdots(x-\lambda_k)^{a_k}$. Then A is similar to a Jordan matrix, which is unique up to the order of the diagonal blocks. This matrix is called the **Jordan normal form** of the matrix

Every non-constant polynomial in $\mathbb{C}[x]$ can be written as a product of linear polynomials (by the fundamental theorem of algebra), so every matrix in $\mathbb{C}^{n\times n}$ has a Jordan normal form.

Calculating the Jordan normal form

The sizes of the blocks in the Jordan normal form can be calculated from the ranks of the powers of the matrices $A - \lambda_i I$ (where λ_i are the eigenvalues). However, if the multiplicity of each eigenvalue in the characteristic polynomial is not greater than 6, then the normal form can be determined from

$k_A(x) = (x - \lambda_1)^{a_1} \cdots (x - \lambda_k)^{a_k}$,
$m_A(x) = (x - \lambda_1)^{b_1} \cdots (x - \lambda_k)^{b_k}$,
$d_i := \dim V_{\lambda_i}$ for $i = 1, \ldots, k$.

Since these numbers are invariant under similarity of matrices, the corollary above shows that for each eigenvalue λ_i

the sum of the sizes of λ_i-blocks is a_i
the largest size of the λ_i-blocks is b_i
the number of the λ_i-blocks is d_i.

Exercises: Determine the Jordan normal form if the characteristic polynomial, the minimal polynomial and the dimension of the eigenspaces are given.

1.
$k(x) = (x-2)^5$
$m(x) = (x-2)^2$
$\dim V_2 = 3$
$5 = 2 + 2 + 1$

$$\begin{bmatrix} 2 & 1 & & & \\ & 2 & & & \\ & & 2 & 1 & \\ & & & 2 & \\ & & & & 2 \end{bmatrix}$$

2.
$k(x) = (x-2)^5$
$m(x) = (x-2)^3$
$\dim V_2 = 3$
$5 = 3 + 1 + 1$

$$\begin{bmatrix} 2 & 1 & & & \\ & 2 & 1 & & \\ & & 2 & & \\ & & & 2 & \\ & & & & 2 \end{bmatrix}$$

3. Determine the Jordan normal form and the minimal polynomial of the matrix A.

$$A = \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

$k_A(x) = (x-1)^3(x-2)$. The eigenspace for the eigenvalue 1 is the kernel of $A - I$. We can use the Gaussian method to bring $A - I$ to row echelon form and see that $\operatorname{rank}(A - I) = 2$, so $\dim V_1 = 4 - 2 = 2$. This means that there are two 1-blocks in the Jordan-matrix, and these can only be a 2×2 and a 1×1 1-block, and we must have a 1×1 2-block:

$$A \sim \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

By the maximal sizes of the blocks, $m_A(x) = (x-1)^2(x-2)$.

Applications of the Jordan normal form

1) One can determine whether two matrices are similar.

2) For $J = P^{-1}AP$, the power J^m can be calculated relatively easily, so we also get A^m.

3) As in the case of the diagonal form (but a diagonal form does not always exist, even in $\mathbb{C}^{n\times n}$!) it provides a better understanding of the transformation.

(Note that in 1) and 3) we don't have to determine the transition matrix P.)

An immediate consequence of the Jordan normal form is the following condition for diagonalizability.

Theorem: A matrix $A \in K^{n\times n}$ whose characteristic polynomial can be written as a product of linear polynomials over K is diagonalizable if and only if the minimal polynomial has no multiple roots.

Proof: A diagonal matrix is also a Jordan matrix, and the Jordan matrix is essentially unique, so the matrix is diagonalizable if and only if every Jordan block is a 1×1 matrix, i.e. the maximal size of the λ_i-blocks is 1 for each i.