BASIC THEORY AND APPLICATIONS OF THE JORDAN CANONICAL FORM

JORDAN BELL

Abstract. This paper gives a basic introduction to the Jordan canonical form and its applications. It looks at the Jordan canonical form from linear algebra, as an "almost" diagonal matrix, and compares it to other matrix factorizations. This article also looks at it from the viewpoint of abstract algebra, as a way to partition the set of square matrices with complex entries. It presents a proof due to A.F. Filippov that any square matrix is similar to a matrix in Jordan canonical form. Some background on the Jordan canonical form is given.

1. Introduction

A good use for diagonalizing a matrix is to view the linear transformation as the action of a matrix of only diagonal entries. It's often easier to understand what a diagonal matrix representation is doing, and for many computations a diagonal matrix is easier to work with than a dense matrix with many values off the diagonal. A good example of where a diagonal matrix is useful is for finding large powers of a matrix, which a diagonal matrix makes amazingly easy; e.g. finding, say, the 50-th power of a matrix made only of 2's on the diagonal. For a fairly easy to understand introduction to matrices and diagonalization, see D.C. Lay in [8]. However, even working with the complex numbers $\mathbb{C}$ as our field, not all matrices can be diagonalized. For example,

$$\begin{bmatrix} 0 & 3 \\ 0 & 0 \end{bmatrix},$$

is not diagaonalizable over $\mathbb{C}$.

The only eigenvalue of this matrix is 0 which has multiplicity 2 as a solution to the characteristic equation (that is, the algebraic multiplicity is 2), but this eigenvalue only yields 1 eigenvector (so the geometric

---

Date: August 31, 2004.

2000 *Mathematics Subject Classification.* 15A21.

Key words and phrases. Jordan canonical form, matrix factorization, partitions.

This paper was written during an NSERC USRA supervised by Dr. B. Stevens. The idea for this paper is from Dr. J. Poland.

multiplicity is 1). Since the geometric multiplicity is less than the algebraic multiplicity, this matrix is not diagonalizable. The *Jordan canonical form* can let us "almost" diagonalize this matrix (we won't form a matrix with only diagonal entries, but, as we will see, it will be fairly close).

2. Jordan blocks, Jordan basis and Jordan canonical form

A square matrix of the form

$$\begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda \end{bmatrix}$$

is called a *Jordan block*. Note that an $n \times n$ matrix is a Jordan block if it is upper triangular, has all diagonal entries equal, and has a 1 directly above each diagonal entry except the first.

A square matrix J, of the form

$$\begin{bmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_s \end{bmatrix},$$

where $J_1, \dots, J_s$ are all Jordan blocks, is said to be in Jordan canonical form. The "canonical" refers to its (possibly arbitrary) choice as a representative; it is representative of all matrices that are similar to it (we discuss this more in the abstract algebra section).

Definition 1 (Matrix direct sum)**.** *The* direct sum $\oplus$ *of two matrices* A $k \times l$ *and* B $m \times n$ *is an* $(k+m) \times (l+n)$ *matrix* C *with* A *in the top left corner and* B *in the bottom right corner; so* A *is in* $c_{1,1}$ *to* $c_{k,l}$ *and* B *is in* $c_{k+1,l+1}$ *to* $c_{k+m,l+n}$.

We can express a matrix in Jordan canonical form as the direct sum of Jordan blocks. So if we have s Jordan blocks then our Jordan canonical form is given by $J = J_1 \oplus \ldots \oplus J_s$ and more compactly as $J = \oplus_{i=1}^{s} J_i$.

Definition 2 (Jordan basis)**.** *Let* A *be an* $n \times n$ *matrix. An ordered set* $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$*, that is a basis for* $\mathbb{C}^n$ *is a Jordan basis for* A *if it has the* Jordan basis property*, that for each* $1 \leq i \leq n$*, we have either* $A\mathbf{b}_i = \lambda_i \mathbf{b}_i$ *or* $A\mathbf{b}_i = \lambda_i \mathbf{b}_i + \mathbf{b}_{i-1}$*, where* λ_i *is also the eigenvalue for* $\mathbf{b}_{i-1}$.

Our proof of Theorem 3 follows the explanations of G. Strang in [10] and Fraleigh and Beauregard in [5]. The original proof by A.F. Fillipov is available in Russian in [3].

We remember that $\mathbf{v}$ is an eigenvector of A if and only if $A\mathbf{v} = \lambda\mathbf{v}$. We can also have "almost" eigenvectors (we call these *generalized eigenvectors*), that after multiple mappings by A yield (along with other generalized eigenvectors) eigenvectors of A. Let us have $A\mathbf{x}_1 = 8\mathbf{x}_1$ and $A\mathbf{x}_2 = 8\mathbf{x}_2 + \mathbf{x}_1$, and $A\mathbf{x}_3 = 0\mathbf{x}_3$ and $A\mathbf{x}_4 = 0\mathbf{x}_4 + \mathbf{x}_3$. Then $\mathbf{x}_2$ and $\mathbf{x}_4$ are the generalized eigenvectors. We note that the set of all eigenvalues of A is known as the *spectrum* of A, that is $Sp(A)$.

Theorem 3 (Filippov's construction)**.** *Any $n \times n$ matrix A is similar to a matrix in Jordan canonical form.*

Proof. In the following we show that we can construct a Jordan basis with n entries for any $n \times n$ matrix, and that the entries of this basis can be used as columns for a matrix P such that $A = P^{-1}JP$, from which the Jordan canonical form J easily follows. We also note that we only need to show that matrices with 0 in their spectrum can have a Jordan basis constructed for them, because of the following, where we assume that any matrix with 0 in its spectrum is similar to a matrix in Jordan canonical form:

$$\begin{aligned} \lambda \in Sp(A) \Rightarrow 0 \in Sp(A - \lambda I) \Rightarrow A - \lambda I = P^{-1}JP \Rightarrow \\ P(A - \lambda I)P^{-1} = J \Rightarrow P^{-1}AP - \lambda I = J \Rightarrow \\ PAP^{-1} = J + \lambda I \Rightarrow A = P^{-1}(J + \lambda I)P. \end{aligned}$$

where $J + \lambda I$ is clearly in Jordan canonical form.

We procede by induction on the dimension r of the range of a linear transformation $T : \mathbb{C}^n \to \mathbb{C}^n$, constructing a Jordan basis with r members for the matrix representation of T. For $n = 1$, $A = [\lambda]$ and clearly A is already in Jordan canonical form, with a Jordan basis of, say, $\{1\}$. We make as our induction hypothesis that for all linear transformations with a range of dimension $r < n$, there exists a Jordan basis for their matrix representation.

Let us have an $n \times n$ matrix A as the matrix representation of a linear transformation with range of dimension r, so the rank of A is r. Thus by assumption there exists a Jordan basis for $ColA$, which will be $\mathcal{F} = \{\mathbf{f}_1, \ldots, \mathbf{f}_r\}$.

We introduce some notation now for use with eigenvectors and their generalized eigenvectors. We know that vectors in a Jordan basis are either eigenvectors or generalized eigenvectors, and for the i-th progression of generalized eigenvectors and one eigenvector, which we will think of as going from the generalized eigenvectors at the right and ending at the eigenvector in the left, let us represent these vectors as $\mathbf{v}_{i1}, \mathbf{v}_{i2}, \ldots, \mathbf{v}_{il(i)}$, where $l(i)$ denotes the length of the i-th progression.

There are $s \geq 0$ progressions of generalized eigenvectors and eigenvectors in $\mathcal{F}$ that end at an eigenvector with eigenvalue 0, and clearly for the i-th progression the first vector $\mathbf{f}_{i1}$ in it is in $NulA$. As well, the last vector in the i-th progression $\mathbf{f}_{il(i)}$ is clearly in $ColA$, so it is of the form $A\mathbf{g}_i = \mathbf{f}_{il(i)}$; let us have $\mathcal{G} = \{\mathbf{g}_1, \ldots, \mathbf{g}_s\}$.

We recall that there are s progressions with an eigenvector with eigenvalue 0 at the left end, and thus by elementary linear algebra we can extend these s vectors with $n - r - s$ vectors to form a basis for $NulA$. We express these as $\mathcal{H} = \{\mathbf{h}_1, \ldots, \mathbf{h}_{n-r-s}\}$.

Let us have $\mathcal{B} = \mathcal{F} \cup \mathcal{G} \cup \mathcal{H}$. We recall that there are s progressions in $\mathcal{F}$ associated with eigenvalue 0, each with a distinct vector in $\mathcal{G}$ at their right end; for convenience, we form a new set $\mathcal{E}$ that contains exactly the vectors in $\mathcal{F}$ associated with eigenvalue 0, and we remove these vectors from $\mathcal{F}$. Thus now the vectors in $\mathcal{F}$ are precisely those vectors that form progressions of generalized eigenvectors leading to eigenvectors associated with nonzero eigenvalues. We will say there are q such progressions, each made by some number of generalized eigenvectors and precisely one eigenvector per progression. Finally, each vector in $\mathcal{H}$ constitutes its own progression, as each is an eigenvector with eigenvalue 0. Thus we compose these progressions in $\mathcal{B}$ such that $\mathcal{B} = \{\mathbf{b}_{11}, \ldots, \mathbf{b}_{il(i)}\}$, and we can clearly compose the progressions such that $\mathcal{B}$ has the Jordan basis property. Now by Definition 2 we know that $\mathcal{B}$ is a Jordan basis for A if it is a basis for $\mathbb{C}^n$.

It is clear that the following equation has only the trivial solution if and only if the vectors in it are linearly independent:

$$\begin{aligned}(1)\qquad \epsilon_{11}\mathbf{e}_{11} + \ldots + \epsilon_{sl(s)}\mathbf{e}_{sl(s)} + \zeta_{11}\mathbf{f}_{11} + \ldots + \zeta_{ql(q)}\mathbf{f}_{ql(q)} + \eta_1\mathbf{g}_1 + \\ \ldots + \eta_s\mathbf{g}_s + \theta_1\mathbf{h}_1 + \ldots + \theta_{n-r-s}\mathbf{h}_{n-r-s} = \mathbf{0}.\end{aligned}$$

We recall that each $\mathbf{h}_i$ is in $NulA$, and thus by premultiplying (1) by A we have:

$$\begin{aligned}(2)\qquad \epsilon_{12}\mathbf{e}_{11} + \ldots + \epsilon_{sl(s)}\mathbf{e}_{s,l(s)-1} + (\lambda_1\zeta_{11} + \zeta_{12})\mathbf{f}_{11} + \ldots + \\ \lambda_q\zeta_{ql(q)}\mathbf{f}_{ql(q)} + \eta_1\mathbf{f}_{1l(1)} + \ldots + \eta_s\mathbf{f}_{sl(s)} = \mathbf{0}.\end{aligned}$$

which is a linear combination of r vectors from $\mathcal{E} \cup \mathcal{F}$, which we know is a basis for $\mathbb{C}^r$. Thus only zero weights will let (2) hold, so we can easily conclude the following:

$$\epsilon_{12} = \ldots = \epsilon_{s2} = \zeta_{1l(1)} = \ldots = \zeta_{ql(q)} = \eta_1 = \ldots = \eta_s = 0.$$

By substituting in these weights, we can make (1) into the following:

$$\begin{aligned}(3)\qquad &\epsilon_{11}\mathbf{e}_{11}+\ldots+\epsilon_{s1}\mathbf{e}_{s1}+\zeta_{11}\mathbf{f}_{11}+\ldots+\zeta_{1,l(1)-1}\mathbf{f}_{1,l(1)-1}+\\ &\zeta_{21}\mathbf{f}_{21}+\ldots+\zeta_{q,l(q)-1}\mathbf{f}_{q,l(q)-1}+\theta_1\mathbf{h}_1+\ldots+\theta_s\mathbf{h}_s=\mathbf{0}.\end{aligned}$$

Clearly the entries in $\mathcal{H}$ are independent from the entries in $\mathcal{F}$, and by assumption the entries in $\mathcal{H}$ are independent from the entries in $\mathcal{E}$, as we extended $\mathcal{E}$ with $\mathcal{H}$ to be a basis for $NulA$. Thus for (3) to hold, we must have the following:

$$\epsilon_{11}=\ldots=\epsilon_{s1}=\zeta_{11}=\ldots=\zeta_{1,l(1)-1}=\zeta_{21}=\ldots=\zeta_{q,l(q)-1}=\\ \theta_1=\ldots=\theta_s=0.$$

Hence all weights in (1) are equal to 0, so the vectors in (1) are linearly independent, so $\mathcal{B}$ is a basis for $\mathbb{C}^n$. Thus $\mathcal{B}$ is a Jordan basis for A. □

There are other proofs that any square matrix with complex entries is similar to a matrix in Jordan canonical form, aside from Filippov's proof. For example, Fletcher and Sorensen in [4] have a different proof that extends the Schur factorization (by using their results that every matrix in Schur form is similar to a block upper triangular of a certain form, and that each triangular block is similar to a matrix of a particular form) to achieve Jordan canonical form.

The Jordan canonical form was defined by the French mathematician Camille Jordan in 1870, in [7]. However, [2] notes that a similar form was used earlier by the German mathematician Karl Weierstraß.

3. Other matrix factorizations

When an $n \times n$ matrix A is similar to a matrix J in Jordan canonical form, we can express this as $A = P^{-1}JP$, where P has as columns n linearly independent vectors of a Jordan basis for A. This is a matrix factorization of A, as we are expressing A as a product of three matrices. There are several other useful factorizations of matrices. These factorizations are usually chosen for similar reasons why we factor a number into its prime factors, that factoring can make them easier to analyze. We compare the Jordan factorization of an $n \times n$ matrix, to the *LU factorization*, *QR factorization* and *Schur factorization*. We explain these factorizations and give an elementary look at their computational uses, but our discussion is mostly about the usefulness and interest of working with these factorizations in pure math.

3.1. **LU factorization.**

The LU factorization of an $m \times n$ matrix A is given by $A = LU$, with L $m \times m$ and U $m \times n$. L is permuted unit lower triangular (unit because all entries on the diagonal are 1's), and

U is in reduced echelon form. It is called "LU" factorization because L is permuted lower triangular, and the U matrix is a rectangular upper diagonal matrix.

The LU factorization is used to solve linear equations of the form $A\mathbf{x} = \mathbf{b}$ for $\mathbf{x}$. If there is an LU solution, there is a solution just by row reducing $\begin{bmatrix} A & \mathbf{b} \end{bmatrix}$. However, finding a solution with LU factorization often takes less operations, so it is computationally better. After we've found L and U, we first solve $L\mathbf{y} = \mathbf{b}$ and then solve $U\mathbf{x} = \mathbf{y}$. Each of these linear equations are fairly easy to solve, because both L and U are (possibly permuted) triangular.

For more discussion about the uses and efficiency of LU factorization, see section 3.2 of [6] by Golub and Van Loan; [8] also suggests looking at section 3.8 of [9] by Noble and Daniel.

We now give an algorithm to find L $m \times m$ and U $m \times n$ for A $m \times n$. By elementary linear algebra, A can be reduced to reduced echelon form U by left multiplication by elementary matrices. It will take some finite number of operations p to reduce A to U, hence we have $E_p \dots E_1 A = U \iff A = (E_p \dots E_1)^{-1} U \iff A = LU$ for $L = (E_p \dots E_1)^{-1} \iff L = E_1^{-1} \dots E_p^{-1}$. This gives us an L that is not in general lower triangular. However, it is permuted lower triangular, so some rearrangement of its rows will give us a lower triangular matrix.

Thus when we are solving $L\mathbf{x} = \mathbf{y}$ by reducing $\begin{bmatrix} L & \mathbf{y} \end{bmatrix}$ to $\begin{bmatrix} I & \mathbf{x} \end{bmatrix}$, instead of following the normal algorithm for reducing a matrix to reduced echelon form of adding multiples of a row to the rows below it, and then adding multiples of the rows above it, we add multiples of the row with the pivot of the first column to the other rows, and then add multiples of the row with the pivot of the second column to the other rows, etc. For example, if the pivot of column 1 was in row 3, and the pivot of column 2 was in row 4, and the pivot of column 3 was in row 10, etc., we would add multiples of row 3 to the other rows to eliminate all other entries of column 1, then add multiples of row 4 to the other rows to get rid of all the other entries in column 2, then add multiples of row 10 to the other rows to get rid of all the other entries in column 3, etc. At the end of these row operations, we will be done; that is, we will not have to use pivots to eliminate entries off the diagonal.

Finding the Jordan canonical form of a matrix requires quite a lot of operations. There are several methods of finding the Jordan canonical form of a matrix, but they all require finding things like null space, column space or eigenspaces. The LU form, on the other hand, only requires getting a matrix in echelon form, and then perhaps rearranging the rows of a permuted lower triagular L. So the LU factorization is computationally "better", in the sense of being easier to find, than the Jordan factorization. The Jordan canonical form is used for mostly different reasons than the LU factorization, however, the LU factorization does not have any clear uses for similarity. We only give the LU factorization to compare the matrix factorization it uses to the Jordan canonical form's matrix factorization.

3.2. **QR factorization.** We first give the definition of *orthonormal* that we use in QR factorization.

Definition 4 (Orthonormal)**.** *Let us have n vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$. We say that these vectors are orthonormal if the inner product of any choice of two of them is equal to 0, so $\langle \mathbf{v}_i, \mathbf{v}_j \rangle$=0 for $i \neq j$, and if the vector length of each is equal to 1, so $||\mathbf{v}_i|| = 1$.*

If we have an injective $m \times n$ matrix A (that is, A has n linearly independent column vectors), we can use QR factorization to give us $A = QR$, with Q $m \times n$ with columns an orthonormal basis for $ColA$ and R $n \times n$ an upper triangular invertible matrix with positive entries on the main diagonal.

If we have a basis $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ for a vector space W, the Gram-Schmidt process allows us to form an orthonormal basis $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$. For our matrix A, we can have $A = \begin{bmatrix}\mathbf{b}_1 & \ldots & \mathbf{b}_n\end{bmatrix}$.

As well, the Gram-Schmit process has, for $a_i \in R$:

$$d\mathbf{b}_p = a_1\mathbf{c}_1 + \ldots + a_p\mathbf{c}_p + 0\mathbf{c}_{p+1} + \ldots + 0\mathbf{b}_n. \tag{4}$$

Hence we can form an orthonormal basis $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$ from a basis for Col$A$. Let the ordered vectors in $\mathcal{C}$ be the ordered columns of Q, so $Q = \begin{bmatrix}\mathbf{c}_1 & \ldots & \mathbf{c}_n\end{bmatrix}$.

Since we have Equation 4, we can have $Q\mathbf{r}_i = \mathbf{b}_i$, with $\mathbf{r}_i$ a vector of weights for the column vectors of Q; these weights will all be 0 for the columns after i. So we can have:

$$\mathbf{r}_i = \begin{bmatrix}\rho_1 \\ \vdots \\ \rho_i \\ 0 \\ \vdots \\ 0\end{bmatrix}.$$

By the definition of the Gram-Schmidt process, $\mathbf{b}_i = Q\mathbf{r}_i$. Thus $A = \begin{bmatrix}Q\mathbf{r}_1 & \ldots & Q\mathbf{r}_n\end{bmatrix}$, as $\mathbf{b}_i$'s are the columns of A. Hence for $R =$

$\begin{bmatrix}\mathbf{r}_1 & \dots & \mathbf{r}_n\end{bmatrix}$, we have $A = QR$. We know that Q has as column vectors an orthonormal basis for $ColA$, and we know that R is upper triangular because of the "stairsteps" implied by Equation 4. To have R with only positive entries on the main diagonal, if for $\mathbf{r}_i$ we have $\rho_i < 0$, we multiply $\mathbf{c}_i$ by -1. This is the same as multiplying a_p by -1 and $\mathbf{c_p}$ by -1 in Equation 4 to keep $\mathbf{b}_p$ the same; multiplying $\mathbf{c}_p$ by -1 does not change the vector space W it spans, and $\mathcal{C}$ is still an orthonormal basis for W.

The QR factorization is very useful in least squares calculations, for finding the best approximation inside a vector space to something outside the vector space. For example, if a matrix A has linearly independent columns, we can find a unique least squares solution for $A\mathbf{x} = \mathbf{b}$, with $\hat{\mathbf{x}} = R^{-1}Q^T\mathbf{b}$. The QR method does not have immediate applications to similarity.

3.3. **Schur decomposition.** We should explain several notations before considering Schur decomposition. For $A \in \mathscr{M}^{n\times n}(\mathbb{C})$, A^H is the *Hermitian conjugate* of A, which takes the complex conjugate of every entry of A and then tranposes A. A matrix A is *unitary* if and only if $A^H = A^{-1}$; thus, for example, an orthogonal matrix with real entries is unitary. A *nilpotent matrix* A has $A^k = 0$ for some $k \geq 1$.

Theorem 5 (Schur's decomposition)**.** *If A is in $\mathscr{M}^{n\times n}(\mathbb{C})$ then there exists a unitary matrix Q such that $Q^HAQ = T = D + N$, where $D = \lambda_1 \oplus \dots \oplus \lambda_s$, with λ_i a Jordan block and N a nilpotent strictly upper triangular matrix.*

Q is unitary, $Q^H = Q^{-1}$, so for $Q^HAQ = T$, $A \sim T$. Thus both the Jordan canonical form and the Schur decomposition express similarity. The difference is that the Jordan form is more likely to give a nicer J, as it has only diagonal eigenvalue entries and sometimes 1's above those so it is sparse, while T is only upper diagonal, so it could be denser. However, the Schur decomposition has a very nice similarity transformation matrix Q, because Q is orthogonal.

4. Applications in group theory

We show that being similar to a Jordan canonical form matrix is a relation that partitions the set of square matrices with complex entries $\mathscr{M}^{n\times n}(\mathbb{C})$. Of course because only one equivalence class has the identity matrix, Jordan canonical form does not partitition the ring of matrices into subrings. We study the question of forming an algebraic structure from the equivalence classes made by partitioning $\mathscr{M}^n(\mathbb{C})$.

We give a lemma on partitioning a set into equivalence classes. A proof is in, eg, [1].

Lemma 6. *Let $*$ be an equivalence relation on a set X. Then the collection of equivalence classes $\mathscr{C}(x)$ forms a partition of X.*

So by Lemma 6, if we can show that $\sim$ is an equivalence relation on $\mathscr{M}^{n\times n}(\mathbb{C})$, it will show that $\sim$ partitions that set into subsets all similar to a particular Jordan canonical form matrix (so the Jordan form of the subset of each subset "represents" the members of that subset). Recall from Cantor's diagonal proof that the real numbers are uncountable, so the complex numbers, which contain the real numbers, are also uncountable. But note that for an $n \times n$ matrix $A = \oplus_{i=1}^{n} \lambda_i$, with $\lambda_i \in \mathbb{C}$, we can have any eigenvalues for A in the complex numbers. Since we have partitioned $\mathscr{M}^{n\times n}(\mathbb{C})$ based on eigenvalues, this gives us uncountably many similarity equivalence classes.

Note that for similarity $A \sim B$ expresses $A = P^{-1}BP$, for some P. But we can have $P = I$, and then $P^{-1} = I$. So we find that $\sim$ is reflexive, because we have an invertible P such that $A = P^{-1}AP$, as $A = A$. If $A \sim B$, then $A = P^{-1}BP$. But then $B = PAP^{-1}$, and we can just have $Q = P^{-1}$, which gives us $B = Q^{-1}AQ$, so $B \sim A$, so $\sim$ is symmetric. If $A \sim B$ and $B \sim C$, then $A = P^{-1}BP$ and $B = Q^{-1}CQ$. Note that $B = PAP^{-1}$, so $PAP^{-1} = Q^{-1}CQ$. Thus $A = P^{-1}Q^{-1}CQP$. Let $R = QP$. As P and Q are invertible, R, their product, is also invertible. So $R^{-1} = P^{-1}Q^{-1}$. Thus $A = R^{-1}CR$, so $\sim$ is transitive. Thus $\sim$ is an equivalence relation. Hence, by the partitioning lemma, $\sim$ partitions $M^{n\times n}(\mathbb{C})$ into subsets of matrices similar to some matrix in Jordan canonical form (and all similar to each other).

We can show that the product of two arbitrary matrices in a certain equivalence class is not in general in that equivalence class; similarly, the equivalence classes are not themselves closed under addition. As well, multiplication of elements in these classes is not well defined, as in general the product of two different choices of matrices as representatives of their equivalence class are not in the same equivalence class. For example, let us have the following:

$$A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, C = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Clearly both A and B are similar to each other, and B is the Jordan canonical form representative of their equivalence class. We find that $AB = C$, and that C is not similar to B^2. Thus we cannot make any groupoid with matrix multiplication as its operation out of similarity classes, because multiplication of similarity classes is not well defined.

5. Conclusion

The Jordan canonical form has many uses in linear and abstract algebra. When it can happen, diagaonalization is quite useful, and it is good to know that we can always get something very close to diagonal form with Jordan canonical form. Of the factorizations we considered, the Jordan canonical form seems most comparable to the Schur decomposition; it seems that for an original matrix A equal to $P^{-1}TP$, the Jordan canonical form gives the "best" T and the Schur decomposition gives the "best" P and P^{-1}. We also found that we can partition the set of square matrices $\mathscr{M}^{n\times n}(\mathbb{C})$ into equivalence classes of different Jordan canonical forms, but that there does not seem to be any natural algebraic structure that can be formed by similarity classes.

References

1. Norman J. Bloch, *Abstract algebra with applications*, Prentice-Hall, Englewood Cliffs, New Jersey, 1987.
2. A. J. Coleman, *The greatest mathematical paper of all time*, Math. Intelligencer **11** (1989), no. 3, 29–38. MR **90f:**01047
3. A. F. Filippov, *A short proof of the theorem on reduction of a matrix to Jordan form.*, Vestnik Moskov. Univ. Ser. I Mat. Meh. **26** (1971), no. 2, 18–19. MR 43 #4839
4. R. Fletcher and D. C. Sorensen, *An algorithmic derivation of the Jordan canonical form*, Amer. Math. Monthly **90** (1983), no. 1, 12–16. MR **84d:**15011
5. John B. Fraleigh and Raymond A. Beauregard, *Linear algebra*, third ed., Addison-Wesley, Reading, Mass., 1995.
6. Gene H. Golub and Charles F. Van Loan, *Matrix computations*, third ed., Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, Baltimore, MD, 1996. MR **97g:**65006
7. Camille Jordan, *Traité des substitutions et des équations algébriques*, Les Grands Classiques Gauthier-Villars. [Gauthier-Villars Great Classics], Éditions Jacques Gabay, Sceaux, 1989, Reprint of the 1870 original. MR **94c:**01039
8. David C. Lay, *Linear algebra and its applications*, third ed., Addison-Wesley, Reading, Mass., 2003.
9. Ben Noble and James W. Daniel, *Applied linear algebra*, third ed., Prentice-Hall, Englewood Cliffs, NJ, 1988.
10. Gilbert Strang, *Linear algebra and its applications*, second ed., Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1980. MR **81f:**15003

E-mail address: `jbell3@connect.carleton.ca`

School of Mathematics and Statistics, Carleton University, 1125 Colonel By Drive, Ottawa, K1S 5B6, Canada.