ON THE DIMENSION OF LINEAR SPACES OF NILPOTENT MATRICES

G.W. MACDONALD, J.A. MACDOUGALL, AND L.G. SWEET

Abstract. We obtain bounds on the dimension of a linear space $S$ of nilpotent $n \times n$ matrices over an arbitrary field. We consider the case where bounds $k$ and $r$ are known for the nilindex and rank respectively, and find the best possible dimensional bound on the subspace $S$ in terms of the quantities $n$, $k$, and $r$. We also consider the case where information is known concerning the Jordan forms of matrices in $S$ and obtain new dimensional bounds in terms of this information. These bounds improve known bounds of Gerstenhaber. Along the way, we generalize and give a new proof of a result Mathes, Omladič, and Radjavi concerning traces on subspaces of nilpotent matrices. This is a key component in the proof of our result and may also be of independent interest.

1. Introduction

In has been half a century since Murray Gerstenhaber, in a sequence of four papers appearing in the Annals of Mathematics, established the seminal results concerning algebras and subspaces of nilpotent matrices over arbitrary fields. In the third paper in the sequence [2], in 1959, Gerstenhaber showed that if $S$ is a subspace of the vector space of $n \times n$ matrices over some field $F$, $S$ consists of nilpotent matrices, and the $F$ is sufficiently large, then the maximal dimension of $S$ is $\frac{n(n-1)}{2}$. In the last paper in the sequence [3], in 1962, Gerstenhaber gave an improved bound on the dimension of $S$ in terms of possible sizes of Jordan blocks in the Jordan forms of matrices in $S$.

There has been only moderate incremental progress in this area since. In 1985, Serežkin [6] removed the cardinality condition in showing that Gerstenhaber’s 1959 result is valid over any field.

Other progress has mainly been in consideration of special cases. In 1991, Mathes, Omladič and Radjavi [5] showed that if the field has at more than two elements and $S$ has maximal dimension, then $S$ is triangularizable. In 1993, Brualdi and Chavey [1] considered the case where all matrices in $S$ have rank bounded by $r$, and showed that the dimension of $S$ is bounded by $nr - \frac{(r+1)^2}{2}$. They also considered the case where all matrices in $S$ have nilindex bounded by $k$, and obtained...
In 2009, the second and third authors of this paper considered a case where conditions were placed on both the rank and the nilindex \[7\], and showed that if the maximal nilindex is two and the rank is bounded by \(r\), then the dimension of \(S\) is bounded by \(r(n-r)\).

In this paper we make two substantial contributions to the general problem of obtaining dimension bounds for subspaces of nilpotent matrices. First, we obtain a sharp bound on the dimension of a subspace of \(n \times n\) nilpotent matrices with rank bound \(r\) and nilindex bound \(k\), in terms of the quantities \(k\), \(r\) and \(n\). Secondly, we give an advance over Gerstenhaber’s 1962 Theorem by giving an improved bound on the dimension of \(S\) in terms of data obtained from the possible Jordan forms of matrices in \(S\).

We also present a new, self-contained, and simplified proof of Serežkin’s generalization of Gerstenhaber’s 1959 Theorem. Many elements of this proof are then generalized to prove the more advanced theorems. The only background knowledge required is of standard theorems and results from linear algebra.

Before proceeding, we review some standard definitions and terminology which will be used throughout the paper.

For a field \(F\), we let \(\text{char}(F)\) denote the characteristic of the field and \(\text{card}(F)\) denote the cardinality of the field. We let \(M_n(F)\) denote the \(n \times n\) matrices over \(F\), \(M_{m,n}(F)\) denote the \(m \times n\) matrices over \(F\) and \(U_n(F)\) denote the set of all strictly upper triangular matrices in \(M_n(F)\).

As usual, for a matrix \(A \in M_n(F)\), we say \(A\) is nilpotent if some power of \(A\) equals 0, and define the nilindex of \(A\) to be the smallest natural number \(k\) so that \(A^k = 0\). Also, let \(\text{tr}(A)\) denote the trace of \(A\) and let \(A^T\) denote the transpose of \(A\).

We also adopt the convention that all matrices shall be denoted by capital letters and their entries by the corresponding lower-case letters. So if \(A\) is in \(M_n(F)\), its \((i,j)\)-th entry is \(a_{ij}\) and we write \(A = [a_{ij}]\).

If \(V\) is a vector space over \(F\), we let \(\dim(V)\) denote its dimension.

For any rational number \(x\), \([x]\), the floor function of \(x\), is the largest integer which is less than or equal to \(x\), and \(\lceil x \rceil\), the ceiling function of \(x\), is the smallest integer which is greater than or equal to \(x\). We define \([0]\) = 1

In Section 2, we give a new proof of a result of Mathes, Radjavi and Omladič [5] concerning a simple trace condition that must be satisfied by any space of nilpotent matrices. In [5] this result is obtained for matrices over any field of characteristic 0 and used to find a new and elegant proof of Gerstenhaber’s 1959 Theorem in that case. We generalize the theorem to fields not of characteristic 0. We then state and prove a Dimension Slicing Lemma (which is a slight generalization of Lemma 1 of [7]) and apply our generalization of the Mathes, Radjavi and Omladič Theorem and our Dimension Slicing Lemma to give an elementary proof of Gerstenhaber’s 1959 Theorem over arbitrary fields.

In Section 3, we obtain a block-matrix generalization of the Theorem of Mathes, Omladič and Radjavi. It gives additional trace conditions which must be satisfied by subspaces of nilpotent matrices, and takes into account the maximum nilindex of matrices in the space. It will be a key component in the sharpening the dimensional bounds in our main theorems, and should also be of independent interest.

In Section 4, we establish a number of technical lemmas required to prove our main theorems.
In Section 5, we use our block-matrix generalization of the Theorem of Mathes, Omladič and Radjavi from Section 3, our Dimension Slicing Lemma from Section 2, technical Lemmas from Section 4, and induction on \(n\) (the size of the matrices) to prove our first major dimension-bounding theorem: we establish the bound on the dimension of a space of nilpotent matrices in terms of the nilindex bound \(k\) and the rank bound \(r\). We show that, if \(S\) is a subspace of \(M_n(F)\), \(\text{card}(F) > n\), and \(S\) consists of nilpotent matrices with nilindex bound \(k\) and rank bound \(r\), then

\[
\dim(S) \leq nr - \frac{r^2}{2} - \frac{r}{2} + \frac{q^2}{2}(k - 1) + \frac{q}{2}(-2r + k - 1)
\]

(where \(q = \left\lfloor \frac{r}{k-1} \right\rfloor\)).

A number of known results, including some of Brualdi and Chavey [1], are special cases of this result and obtained as corollaries.

In Section 6, we show that our bound is sharp by exhibiting examples of maximal dimension in all feasible cases. We then compare and contrast best previously known dimensional bounds on subspaces of nilpotent matrices with our result.

In Section 7, we define a spatial Jordan partition of a subspace of nilpotent matrices. This is a list of numbers determined by considering possible Jordan forms of matrices in our subspace and is related to the definition of Jordan partition used by Gerstenhaber. Again using our block-matrix generalization of the Theorem of Mathes, Omladič and Radjavi from Section 3, our Dimension Slicing Lemma from Section 2 and our technical Lemmas from Section 4, we obtain improved dimensional bounds in terms of this quantity.

While Gerstenhaber used techniques and theorems from algebraic geometry, and Brualdi and Chavey use techniques and theorems from combinatorics, all of our proofs use only linear algebra.

2. The Mathes-Omladič-Radjavi Theorem

Following is the result that appears as Corollary 1 in [5].

**Theorem 2.1** (Mathes, Omladič, and Radjavi). Suppose \(S\) is a linear space of nilpotent matrices over a field of characteristic 0. If \(A, B \in S\) and \(k \in \mathbb{N}\) then \(\text{tr}(A^kB) = 0\).

We now provide a very elementary proof of a slightly stronger result.

**Theorem 2.2.** Suppose \(S\) is a subspace of \(M_n(F)\) and \(S\) consists of nilpotent matrices. If \(A, B \in S\) and \(k \in \mathbb{N}\) and \(\text{card}(F) \geq n\) and \(\text{char}(F)\) is not a multiple of \(k + 1\) then \(\text{tr}(A^kB) = 0\).

**Proof.** Since \(A\) is nilpotent, \(A^n = 0\) so the result is obvious for \(k \geq n\). Assume \(1 \leq k \leq n - 1\). For any \(x \in F\), \(xA + B\) is nilpotent so \(\text{tr}\left((xA + B)^{k+1}\right) = 0\). But \(\text{tr}\left((xA + B)^{k+1}\right)\) is a polynomial of degree \(k\) (since the coefficient of \(x^n\) is \(A^n = 0\)) and since \(\text{card}(F) \geq n > k\) the coefficient of each term must be 0. In particular the coefficient of \(x^k\) is

\[
\text{tr}(A^kB + A^{k-1}BA + A^{k-2}BA^2 + \cdots + BA^k) = (k+1)\text{tr}(A^kB) = 0.
\]

Since \(k + 1\) is not a multiple of \(\text{char}(F)\) we must have \(\text{tr}(A^kB) = 0\) as required. \(\square\)
Corollary 2.3. Suppose $\mathcal{S}$ is a subspace of $M_n(\mathbb{F})$ and $\mathcal{S}$ consists of commuting nilpotent matrices. If $A, B \in \mathcal{S}$, $k, l \in \mathbb{N}$, $\text{card}(\mathbb{F}) \geq k + l$, and $\text{char}(\mathbb{F})$ is not a multiple of the binomial coefficient $\binom{k+l}{k}$, then $\text{tr}(A^k B^l) = 0$.

Proof. The proof is similar to the above. Consider $\text{tr}((xA + B)^{k+l})$. □

We can weaken the hypotheses of Theorem 2.1 even further by removing condition about the $\text{char}(\mathbb{F})$.

Theorem 2.4. Suppose $\mathcal{S}$ is subspace of $M_n(\mathbb{F})$ and $\mathcal{S}$ consists of nilpotent matrices. If $A, B \in \mathcal{S}$, $k \in \mathbb{N}$ and $\text{card}(\mathbb{F}) > k$ then $\text{tr}(A^k B) = 0$.

Proof. Using a similarity transformation we may assume that $A$ is in Jordan form. We may also assume $1 \leq k \leq n - 1$. For $B$ in $\mathcal{S}$ and $x \in \mathbb{F}$, the characteristic polynomial of $xA + B$ is $c_{xA+B}(\lambda) = \lambda^n = \sum_{i=0}^n c_{n-i}(x)\lambda^i$. The coefficient $c_{k+1}(x)$ is identically zero and is the sum of the $(k+1) \times (k+1)$ principal minors of $xA + B$. This is a polynomial of degree at most $k$ in $x$ whose leading term is $\pm \text{tr}(A^k B)$. Since $\text{card}(\mathbb{F}) > k$ the coefficient of each term must equal 0, so $\text{tr}(A^k B) = 0$. □

In the remainder of this paper we use a Dimension Slicing Lemma, which is a slight generalization of a lemma proved by two of the authors in a previous paper [7].

Given a subspace $\mathcal{S}$ of $M_{mn}(\mathbb{F})$, and a set of indices $I$ in $\{1, 2, \ldots, m\} \times \{1, 2, \ldots, n\}$, we associate two subspaces of $M_{mn}(\mathbb{F})$ to $\mathcal{S}$. The first is $W_\mathcal{S}$, which is constructed by taking all the elements of $\mathcal{S}$ and “zeroing out” the entries whose index is not in $I$. So

$$W_\mathcal{S} = \left\{ W \in M_{mn}(\mathbb{F}) : \text{ there exists } A \in \mathcal{S}, \text{ with } w_{ij} = \begin{cases} a_{ij} & \text{if } (i, j) \in I, \\ 0 & \text{if } (i, j) \notin I. \end{cases} \right\}$$

The second subspace is

$$U_\mathcal{S} = \{ S \in \mathcal{S} : s_{ij} = 0 \text{ whenever } (i, j) \in I \}.$$

Lemma 2.5 (Dimension Slicing Lemma). For $\mathcal{S}$ a subspace of $M_{mn}(\mathbb{F})$, $I$ a subset of $\{1, 2, \ldots, m\} \times \{1, 2, \ldots, n\}$, and $W_\mathcal{S}$ and $U_\mathcal{S}$ defined as above, we have that

$$\dim(\mathcal{S}) = \dim(W_\mathcal{S}) + \dim(U_\mathcal{S}).$$

Proof. Consider the linear transformation $P : M_{mn}(\mathbb{F}) \rightarrow M_{mn}(\mathbb{F})$ defined by $P(A) = B$ where

$$b_{ij} = \begin{cases} a_{ij} & \text{if } (i, j) \in I, \\ 0 & \text{if } (i, j) \notin I. \end{cases}$$

Restrict $P$ to $\mathcal{S}$ and apply the Rank-Nullity Theorem ([4]). Clearly the range of $P|_\mathcal{S}$ is $W_\mathcal{S}$ and the kernel of $P|_\mathcal{S}$ is $U_\mathcal{S}$, so the result follows. □

Using Theorem 2.4, Lemma 2.5 and a construction similar to one in [5] we can give a short proof of Serežkin’s extension Gerstenhaber’s Theorem [2] to arbitrary fields.

Theorem 2.6 (Gerstenhaber / Serežkin ). Suppose $\mathcal{S}$ is a subspace of $M_n(\mathbb{F})$ and $\mathcal{S}$ consists of nilpotent matrices, then $\dim(\mathcal{S}) \leq \frac{n(n-1)}{2}$. 
Theorem 2.1.\...
Theorem 3.1. Suppose \( S \) is a subspace of \( M_n(\mathbb{F}) \), \( S \) consists of nilpotent matrices and \( \text{card}(\mathbb{F}) \geq n \). Let \( Q \) be any matrix in \( S \) of maximal nilindex \( k = k_1 \) and let \( J = \oplus J_i \) be the Jordan normal form of \( Q \), where \( J_i \) is a \( k_i \times k_i \) matrix. If \( A \) is any matrix in \( S \) and \( A \) is written as a block matrix \( A = [A_{i,j}] \) (with respect to the same decomposition as \( Q \)), then
\[
\text{tr}(A_{i,j}J_j^l) = 0 \quad \text{for all } l \text{ such that } 0 \leq k - k_i \leq l \leq k_j - 1.
\]

Proof. Let \( x \in \mathbb{F} \). Then \((J + xA)^k = 0\) and since \( \text{card}(\mathbb{F}) \geq n \) all the coefficients of the matrix polynomial \((J + xA)^k\) must equal 0. The coefficient of \( x \) gives us
\[
J^{k-1}A + J^{k-2}AJ + J^{k-3}AJ^2 + \ldots + AJ^{k-1} = 0.
\]
Each entry of this sum provides a linear relationship on the entries of \( A \).

Now let \( B = (J + xA)^k \) and partition \( B \) as described above. Then
\[
B_{i,j} = J_i^{k-1}A_{i,j} + J_i^{k-2}A_{i,j}J_i + J_i^{k-3}A_{i,j}J_i^2 + \ldots + A_{i,j}J_i^{k-1} = 0.
\]
If \( J_i \neq 0 \) then \( J_iA_{i,j} \) can be found by pushing the rows of \( A_{i,j} \) down one row and inserting a row of zeroes on the top. Similarly if \( J_j \neq 0 \) then \( A_{i,j}J_j \) can be found by pushing the columns of \( A_{i,j} \) one column to the left and inserting a column of zeroes on the right.

Now for each \( i, j \) it is easy to see that the last row of \( B_{i,j} \) is made up of zeroes and sums of diagonals of \( A_{i,j} \). More precisely, the last row of \( B_{i,j} \) is
\[
\left[ \sum D_{k-k_1}, \sum D_{k-k_1+1}, \ldots, \sum D_{k_j-1}, 0, 0, \ldots, 0 \right] = 0
\]
where there are \( k - k_i \) zeroes at the end of the row.

If \( k = k_i \) then there are no zeroes, and we obtain that the sum along each diagonal of \( A_{i,j} \) above and including the main diagonal is zero. If \( k - k_i > k_j - 1 \) (so \( k > (k_i + k_j) - 1 \)) then the last row is all zeroes and we obtain no information. Also the argument is still valid if \( J_i = 0 \) and \( k_j = k \) or \( J_j = 0 \) and \( k_i = k \) but in these cases we only determine that the \((1, k_j)\) entry of \( B_{i,j} \), which is \( \sum D_{k_j-1} \) is 0.

Informally, this means that the sum of each of the \( k - (k_i + k_j) \) highest diagonals of \( A_{i,j} \) is zero. Finally we note that this result can also be expressed as
\[
\text{tr}(A_{i,j}J_j^l) = 0 \quad \text{for all } l \text{ such that } 0 \leq k - k_i \leq l \leq k_j - 1.
\]
(\text{In the case where } l = 0, \text{ we follow the usual convention that } J_j^0 \text{ is the } k_j \times k_j \text{ identity matrix.}) \quad \Box

4. Technical Lemmas

In this section we prove a number of technical lemmas required to prove our main theorems.

A key step will be to use our Dimension Slicing Lemma, Lemma 2.5 to break \( S \) into “diagonal” and “off-diagonal” subspaces. The main difficulty will be in bounding the dimension of the “off-diagonal” subspaces. For this we will again use Lemma 2.5, as well as Theorem 3.1 and some clever calculating. That is the content of the following technical lemma.

For \( 1 \leq k \leq n \), decompose \( \mathbb{F}^n \) as a direct sum \( \mathbb{F}^n = \mathbb{F}^k \oplus \mathbb{F}^{n-k} \). For each \( A \in M_n(\mathbb{F}) \), identify \( A \) with \( \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \).
Lemma 4.1. Suppose $S$ is a subspace of $M_n(\mathbb{F})$ so that $S$ consists of nilpotent matrices of bounded nilindex $k$, and there exists $N \in M_{n-k,n-k}(\mathbb{F})$ so that

$$J = \begin{bmatrix} J_k & 0 \\ 0 & N \end{bmatrix} \in S$$

(where $J_k$ is a $k \times k$ Jordan block). If $\text{card}(\mathbb{F}) \geq n$, then the subspace

$$\mathcal{X} = \left\{ \begin{bmatrix} 0 & A_{12} \\ A_{21} & 0 \end{bmatrix} : \text{there exists } A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \in S \right\}$$

has

$$\dim(\mathcal{X}) \leq (k-1)(n-k).$$

Proof. With no loss of generality, we may assume that $J$ is in Jordan form with first block $J_k$, since it can be transformed into such via a block diagonal similarity which will not change the dimension of $\mathcal{X}$.

Let $A \in S$ and let $x \in \mathbb{F}$. Then $(J + xA)^k = 0$ and since $\text{card}(\mathbb{F}) \geq n$ all the coefficients of the matrix polynomial $(J + xA)^k$ must equal 0. The coefficient of $x^2$ gives us

$$D = J^{k-2}A^2 + J^{k-1}A^2J + J^{k-1}AJA + \ldots A^2J^{k-2} = 0.$$

This sum consists of all words of length $k$ containing $k-2$ $J$'s and two $A$'s. Each entry of this sum provides a quadratic relationship on the entries of $A$. As before, $JA$ shifts rows down and $AJ$ shifts columns to the left and the total number of shifts in terms in this sum is always $k - 2$. Now consider the entry $d_{k,1}$. This is simply the sum of the dot product of the $k$-th row and the first column of each term of the above expression, if each term is written as a product of two matrices.

Let

$$A = \begin{bmatrix} A_1 & R \\ C & A_2 \end{bmatrix}$$

be the block matrix of $A$ with respect to the above decomposition. So $A_1$ is $k \times k$ and $A_2$ is $(n-k) \times (n-k)$. Now $C$ is $(n-k) \times k$, and we denote its columns by $c_1, c_2, \ldots, c_k$. These are vectors in $\mathbb{F}^{(n-k)}$. Also, $R$ is $(n-k) \times k$, and we denote its rows by $r_1^T, r_2^T, \ldots, r_k^T$, so $r_1, r_2, \ldots, r_k$ are also vectors in $\mathbb{F}^{(n-k)}$.

Since $J$ is block diagonal, each term of $d_{k,1}$ is a sum of terms of two types: products of pairs of entries from $A_1$ and products of pairs of entries, with one from $R$ and one from $C$. Furthermore, terms which are products of pairs of entries from $A_1$ must be of the form $a_{i,j}a_{m,n}$ where one of the two terms is in the strictly lower triangular part of $A$, and one is in the upper triangular part of $A$, otherwise it would either take more or less than $k-2$ shifts to get one term to the $l$-th entry of first column and the other term to $l$-th entry of the last row (or vice versa), which is necessary to appear in $d_{k,1}$. So with no loss of generality we may assume that these terms are of the form $a_{i,j}a_{m,n}$ where $i > j$ and $m \leq n$.

We claim that, in $d_{k,1}$, the sum over all products of pairs of entries from $A_1$ is zero; that is, $d_{k,1}$ is actually a sum of products of entries, with one from $R$ and one from $C$.

To see this, consider all terms in $d_{k,1}$ which contain a given $a_{i,j}$ where $i > j$. Such a term arises when $a_{i,j}$ is shifted into the $k$-th row or the first column. The first condition occurs when there is a word in the expression for $D$ of the form $(J^{k-1}A)B$ where $B$ is a word of length $i-1$ containing exactly one $A$ and $i-2$ $J$'s.
The second condition occurs when there is a word in the expression for $D$ of the form $B(AJ^{j-1})$ where $B$ is a word of length $k - j$ containing exactly one $A$ and $k - j - 3$ $J$’s. By taking all possible values of $B$ for the two cases it can be observed that the sum of all of the terms containing $a_{i,j}$ is actually $a_{i,j}(\sum D_{(i-j)-1})$. Since $(i - j) - 1 \geq 0$, Theorem 3.1 gives us that $\sum D_{(i-j)-1} = 0$.

What terms from $R$ and $C$ appear in the expression for $d_{k,1}$? No entry from $r_1^T$ (the first row of $R$) or $c_k$ (the last column of $C$) can appear, since it would take $k - 1$ shifts to move these terms to the last row of $R$ or first column of $C$ and we only have $k - 2$ shifts available. Note that from the term $J^{k-2}A^2$ we obtain $r_i^Tc_1$ since the second row of $R$ is shifted to the $k$-th row and then multiplied by the first column of $C$. From the term $J^{k-3}A^2J$ we obtain $r_i^Tc_2$ since it is the third row of $R$ which is shifted $(k - 3)$ times to the $k$-th row and it is the second column of $C$ which is shifted to the first column. Similarly, for each term of the form $J^{k-i}A^2J^{i-2}$, for $i = 2, 3, \ldots, k - 2$ we get a contribution $r_i^Tc_{i+1}$.

We can never obtain in this expression an entry from an $i$-th row of $R$ times an entry from a $j$-th column of $C$ where $j - i > 1$, since it would take $k - i$ down shifts to move the entries of $r_i^T$ to the $k$-th row and $j - 1$ left shifts to move the $j$-th column of $C$ to the first column but the total number of down shifts plus left shifts available is at most $k - 2$.

However, terms of the form $J^{k-i}A^iJ^{j-2}$ where $(j - i) + l = 0$ can contribute terms which include an entry of an $i$-th row of $R$ and $j$ column of $C$ where $j - i < 1$.

If we let

$$r = \begin{bmatrix} r_2 \\ r_3 \\ \vdots \\ r_k \end{bmatrix} \quad \text{and} \quad c = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_{k-1} \end{bmatrix}$$

then these are vectors in $\mathbb{F}^{(k-1)(n-k)}$ and the equation $d_{k,1} = 0$ can be expressed as $r^TXc = 0$, where $X$ is an $(k - 1)(n - k) \times (k - 1)(n - k)$ upper triangular matrix with ones on the diagonal and thus is invertible.

We now apply Lemma 2.5 to $X$, using the index set $I = \{k + 1, k + 2, \ldots, n\} \times \{1, 2, \ldots, k\}$. Then

$$W_X = \left\{ \begin{bmatrix} 0 & 0 \\ C & 0 \end{bmatrix} : \text{there exists } A \in \mathcal{S} \text{ with } A_{2,1} = C \right\} ,$$

$$U_X = \left\{ \begin{bmatrix} 0 & R \\ 0 & 0 \end{bmatrix} : \begin{bmatrix} 0 & R \\ 0 & 0 \end{bmatrix} \in \mathcal{S} \right\} ,$$

and

$$\dim(X) = \dim(W_X) + \dim(U_X).$$

Consider the mapping $T_1 : U_X \rightarrow \mathbb{F}^{(k-1)(n-1)}$ which maps

$$\begin{bmatrix} 0 & R \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} r_1^T \\ r_2^T \\ \vdots \\ r_k^T \\ 0 \ldots 0 \ldots 0 \end{bmatrix} \quad \text{to} \quad \begin{bmatrix} r_2 \\ \vdots \\ r_k \\ 0 \ldots 0 \end{bmatrix} = r$$
Lemma 4.2. For 

and the map \( T_2 : W_X \to \mathbb{F}^{(k-1)(n-k)} \) defined by mapping

\[
\begin{bmatrix}
0 & 0 \\
C & 0 \\
c_1 & \cdots & c_{k-1} & c_k & 0
\end{bmatrix}
\overset{T_2}{\to}
\begin{bmatrix}
c_1 \\
0 \\
0 \\
\vdots \\
0 \\
c_{k-1}
\end{bmatrix}
= \mathbf{c}.
\]

Both these maps are injective, by Theorem 3.1, as each entry in \( r_1 \) (resp. \( c_k \)) is the negative of a sum of entries in \( r_2, \ldots, r_k \) (resp. \( c_1, \ldots, c_{k-1} \)). Hence \( \dim(T_1(U_X)) = \dim(U_X) \) and \( \dim(T_2(W_X)) = \dim(W_X) \).

If \( \begin{bmatrix} 0 & Q \\ 0 & 0 \end{bmatrix} \in U_X \) and \( \begin{bmatrix} 0 & R \\ C & 0 \end{bmatrix} \in \mathcal{M} \) then \( \begin{bmatrix} 0 & R + Q \\ C & 0 \end{bmatrix} \in \mathcal{M} \). Letting \( \mathbf{r} = T_1(\begin{bmatrix} 0 & R \\ 0 & 0 \end{bmatrix}) \), \( \mathbf{q} = T_1(\begin{bmatrix} 0 & Q \\ 0 & 0 \end{bmatrix}) \) and \( \mathbf{c} = T_2(\begin{bmatrix} 0 & 0 \\ C & 0 \end{bmatrix}) \), we obtain that \( (\mathbf{r} + \mathbf{q})^T \mathbf{X} \mathbf{c} = 0 \). But we have \( \mathbf{r}^T \mathbf{X} \mathbf{c} = 0 \), so for all \( \mathbf{q} \in T_1(U_X) \) and \( \mathbf{c} \in T_2(W_X) \), we have \( \mathbf{q}^T \mathbf{X} \mathbf{c} = 0 \). So \( T_1(U_X) \) and \( XT_2(W_X) \) are orthogonal subspaces in \( \mathbb{F}^{(k-1)(n-k)} \).

Putting it all together (using that \( X \) is invertible) we obtain that

\[
\dim(\mathcal{M}) = \dim(U_X) + \dim(W_X) = \dim(T_1(U_X)) + \dim(T_2(W_X)) = \dim(\mathbb{F}^{(k-1)(n-k)}) = (k-1)(n-k)
\]

and the lemma is proven. \( \square \)

In Section 5, we will show that if \( \mathcal{S} \) is a subspace of \( M_n(\mathbb{F}) \) (where \( \text{card}(\mathbb{F}) > n \)) and \( \mathcal{S} \) consists of nilpotent matrices whose nilindex is less than or equal to \( k \) and whose rank less than or equal to \( r \), then

\[
\dim(\mathcal{S}) \leq nr - \frac{r^2}{2} - \frac{r}{2} + \frac{q^2}{2}(k-1) + \frac{q}{2}(-2r + k - 1)
\]

(where \( q = \left\lfloor \frac{r}{k-1} \right\rfloor \)).

We shall be using induction on \( n \) to prove this, but it will make it easier to verify the induction step if we know, apriori, that our dimension bound for \( \mathcal{S} \),

\[
b_{r,n}(k) = nr - \frac{r^2}{2} - \frac{r}{2} + \frac{q^2}{2}(k-1) + \frac{q}{2}(-2r + k - 1)
\]

is an increasing function of the nilindex bound \( k \) (holding the rank bound \( r \) and the matrix size \( n \) fixed). Recall that \( q = \left\lfloor \frac{r}{k-1} \right\rfloor \), so this assertion is not immediately obvious. It is true nonetheless, and this is a direct consequence of the following lemma.

**Lemma 4.2.** For \( r \in \mathbb{N} \), the function

\[
d(k) = \left\lfloor \frac{r}{k-1} \right\rfloor^2 (k-1) + \left\lfloor \frac{r}{k-1} \right\rfloor (-2r + k - 1)
\]

is an increasing function of \( k \) for \( k = 2, 3, \ldots, r + 1 \).
Proof. For a fixed $k \in \{2,3,\ldots,r\}$, we must show
\[ d(k+1)-d(k) = \left\lfloor \frac{r}{k} \right\rfloor ^2 (k) + \left\lfloor \frac{r}{k} \right\rfloor (-2r+k) - \left\lfloor \frac{r}{k-1} \right\rfloor ^2 (k-1) - \left\lfloor \frac{r}{k-1} \right\rfloor (-2r+k-1). \]
is non-negative.

Let $m = \left\lfloor \frac{r}{k-1} \right\rfloor$. Then $r \in M = \{mk(k-1), (m+1)k(k-1)\}$ and
\[ \{I_i = \{mk(k-1) + ik, mk(k-1) + (i+1)k : i = 0,1,\ldots,k-2\} \]
are two partitions of $M$. It is easy to see that $I_i \cap J_j = \emptyset$ unless either $j = i$ or $j = i+1$, so we must have that either: (1) $r \in I_i \cap J_j$ for some $i = 0,1,2,\ldots,k-1$; or (2) $r \in I_i \cap J_{i+1}$ for some $i = 0,1,2,\ldots,k-1$.

In case (1) we have that $mk(k-1) + ik \leq r < mk(k-1) + (i+1)(k-1)$ and
\[ \left\lfloor \frac{r}{k} \right\rfloor = m(k-1) + i, \quad \left\lfloor \frac{r}{k-1} \right\rfloor = mk + i. \]
In case (2) we have that $mk(k-1) + (i+1)(k-1) \leq r < mk(k-1) + (i+1)k$ and
\[ \left\lfloor \frac{r}{k} \right\rfloor = m(k-1) + i, \quad \left\lfloor \frac{r}{k-1} \right\rfloor = mk + i + 1. \]

Substituting these values into $d(k+1)-d(k)$ and using the bound for $r$ we obtain that in case (1)
\[ d(k+1)-d(k) \geq m^2(k^2 - k) + 2m(2i) + (i^2 + i), \]
while in case (2)
\[ d(k+1)-d(k) \geq m^2(k^2 - k) + 2m((k-i) + (mi-1)) + (i^2 + i). \]
In case (1), it is immediate that $d(k+1)-d(k) \geq 0$. In case (2), by considering the formula in the subcases where $m = 0$ or $i = 0$ separately we also obtain that $d(k+1)-d(k) \geq 0$. \qed

5. Dimensional bounds from nilindex and rank bounds

We now state and prove our result obtaining dimensional bounds from nilindex and rank bounds.

**Theorem 5.1.** Suppose $\mathcal{S}$ is a subspace of $M_n(\mathbb{F})$ and $\mathcal{S}$ consists of nilpotent matrices whose nilindex is less than or equal to $k$ and whose rank is less than or equal to $r$. If $\text{card}(\mathbb{F}) > n$ then, setting $q = \left\lfloor \frac{r}{k-1} \right\rfloor$, we have that
\[ \dim(\mathcal{S}) \leq nr - \frac{r^2}{2} - \frac{r}{2} + \frac{q^2}{2}(k-1) + \frac{q}{2}(-2r+k-1). \]

**Proof.** As mentioned, we shall prove this induction on $n$.

The base case is $n = 1$. Clearly, in this case, $\mathcal{S}$ must be the zero subspace, so $r = 0$ and $k = 1$. The bound expression clearly simplifies to zero so the Theorem is true when $n = 1$.

Next suppose the Theorem is true for all dimensions $1,2,3,\ldots,n-1$ and that $\mathcal{S}$ is as in the Theorem.
By Lemma 4.2 there is no loss of generality in assuming that there is a matrix $Q$ in $S$ which achieves the nilindex bound $k$, and with no loss of generality assume $Q$ is in its Jordan form (as a similarity applied to $S$ can achieve this without changing dimension, maximal nilindex or rank bound). So $Q = \oplus J_{k_i}$, where each $J_{k_i}$ is a Jordan block and the first Jordan block $J_{k_1} = J_k$ is the largest). Write $Q$ in $2 \times 2$ block matrix form

$$Q = \begin{bmatrix} J_k & 0 \\ 0 & N \end{bmatrix}.$$  

Choose two matrices $A$ and $B$ in $S$. With respect to this same decomposition, we express $A$ in $S$ as a $2 \times 2$ matrix

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

where $A_{11} \in M_{k,k}(F)$, $A_{22} \in M_{(n-k),(n-k)}(F)$, $A_{21} \in M_{(n-k),k}(F)$, and $A_{12} \in M_{k,(n-k)}(F)$.

Consider the index set $I = \{1,2,\ldots,k\} \times \{k+1,k+2,\ldots,n\} \cup \{k+1,k+2,\ldots,n\} \times \{1,2,\ldots,k\}$, and for this index set define $W_S$ and $U_S$ as in Lemma 2.5. Then $W_S$ is the subspace of $M_n(F)$ spanned that all the entries that appear in the $(1,2)$ and $(2,1)$ locations of the matrices in $S$, and $U_S$ will be the set of all the matrices in $S$ of the form

$$\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}.$$  

By Lemma 2.5,

$$\dim(S) = \dim(W_S) + \dim(U_S)$$

and by Lemma 4.1

$$\dim(W_S) \leq (k - 1)(n - k).$$

To find $\dim(U_S)$, we again apply Lemma 2.5 to $U_S$ with the index set $I = \{1,2,\ldots,k\} \times \{1,2,\ldots,k\}$. Then we obtain

$$W_{U_S} = \left\{ \begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix} : \text{there exists } A_2 \text{ so that } \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \in S \right\}$$

and

$$U_{U_S} = \left\{ \begin{bmatrix} 0 & 0 \\ 0 & A_2 \end{bmatrix} : \begin{bmatrix} 0 & 0 \\ 0 & A_2 \end{bmatrix} \in S \right\}.$$  

The elements of $W_{U_S}$ are $k \times k$ nilpotent matrices with no other rank or nilindex restrictions so the best bound we can get is $\dim(W_{U_S}) \leq \frac{(k-1)k}{2}$ by Theorem 2.6. Also, $U_{U_S}$ is essentially a set of $(n-k) \times (n-k)$ nilpotent matrices, and so is amenable to our induction hypothesis.

Summarizing, we now have that

1. $\dim(S) = \dim(W_S) + \dim(U_S)$
2. $\leq (k - 1)(n - k) + \dim(W_{U_S}) + \dim(U_{U_S})$
3. $\leq (k - 1)(n - k) + \frac{k(k - 1)}{2} + \dim(U_{U_S}).$

The dimension of $U_{U_S}$ is clearly equal to the dimension of

$$S' = \left\{ A' : \begin{bmatrix} 0 & 0 \\ 0 & A' \end{bmatrix} \in U_{U_S} \right\}.$$
We apply our induction hypothesis to this subspace $S'$. Clearly the matrix size is $n' = n - k < n$. If $A = \begin{bmatrix} 0 & 0 \\ 0 & A' \end{bmatrix} \in U_{S'}$, then for $Q = \begin{bmatrix} J_k & 0 \\ 0 & N \end{bmatrix}$ as above, $A + xQ$ is in $S$ for any $x \in F$ and so $k - 1 + rank(xN + A') \leq r$. Thus $rank(xN + A') \leq n - k + 1$ for all $x \in F$, which implies that $rank(A') \leq r - (k - 1) = r + 1 - k$ for all $A' \in S'$. So the new rank bound we use applying the induction hypothesis to $U_{S'}$ will be $r' = r + 1 - k$.

Clearly, $k'$, the maximal nilindex of matrices $A' \in S'$ can be no more than $k$ (since $k$ is the maximal index for $S$), but we must also have $k' \leq n - k$ or there would not be room for a second $k' \times k'$ Jordan block. Also since the rank of Jordan block of size $k$ is $k - 1$, we must have $(k - 1) + (k' - 1) \leq r$ or we would exceed the allowable rank, so $k' \leq r + 2 - k$.

Thus $k'$, the new maximal nilindex of $S'$ will satisfy

$$1 \leq k' \leq \min\{k, n - k, r + 2 - k\}.$$ 

Since Lemma 4.2 gives that our dimension formula increases as a function of $k$, with no loss of generality we may assume $k' = \min\{k, n - k, r + 2 - k\}$. Let $q' = \left\lceil \frac{r + k - 1}{r + 1 - k} \right\rceil$.

Consider the following three comprehensive possibilities.

**Case 1:** $k' = k \leq \min\{n - k, r + 2 - k\}$
Here, $q' = \left\lfloor \frac{r + 1 - k}{k - 1} \right\rfloor = \left\lfloor \frac{r}{k - 1} - 1 \right\rfloor = q - 1$.

Our induction hypotheses gives that

$$dim(S') \leq n' r' - \frac{r'^2}{2} - \frac{r'}{2} + \frac{q'^2}{2}(k' - 1) + \frac{q'}{2}(-2r' + k' - 1)$$

$$\leq (n - k)(r + 1 - k) - \frac{(r + 1 - k)^2}{2} - \frac{(r + 1 - k)}{2}$$

$$+ \frac{(q - 1)^2}{2}(k - 1) + \frac{(q - 1)}{2}(-2r - 2 + 2k + k - 1).$$

Then

$$dim(S) = dim(W_S) + dim(W_{U_S}) + dim(S')$$

$$\leq (k - 1)(n - k) + \frac{(k - 1)}{2} + (n - k)(r + 1 - k) - \frac{(r + 1 - k)}{2}$$

$$+ \frac{(r + 1 - k)}{2} + \frac{(q - 1)^2}{2}(k - 1) + \frac{(q - 1)}{2}(-2r - 2 + 2k + k - 1)$$

$$= nr - \frac{r^2}{2} - \frac{r}{2} + \frac{q^2}{2}(k - 1) + \frac{q}{2}(-2r + k - 1).$$

**Case 2:** $k' = r + 2 - k \leq n - k < k$

Here $q' = \left\lfloor \frac{r + 1 - k}{r + 2 - k} \right\rfloor = 1$. 


Our induction hypothesis gives that
\[
\dim(S') \leq n' r' - \frac{r'^2}{2} - \frac{r'}{2} + \frac{q'^2}{2} (k' - 1) + \frac{q'}{2} (-2r' + k' - 1)
\]
\[
\leq (n - k)(r + 1 - k) - \frac{(r + 1 - k)^2}{2} - \frac{(r + 1 - k)}{2} + \frac{(r + 1 - k)}{2}
\]
\[
+ \frac{(-2r - 2 + 2k + r + 2 - k - 1)}{2}
\]
\[
\leq (n - k)(r + 1 - k) - \frac{(r + 1 - k)^2}{2} + \frac{(r + 1 - k)}{2}.
\]
Then,
\[
\dim(S) = \dim(W_S) + \dim(W_{U_S}) + \dim(S')
\]
\[
\leq (k - 1)(n - k) + \frac{(k - 1)k}{2} + (n - k)(r + 1 - k)
\]
\[
- \frac{(r + 1 - k)^2}{2} + \frac{(r + k - 1)}{2}
\]
\[
\leq nr - \frac{r^2}{2} - 3r + k - 1.
\]
But since \(r + 2 - k < k\), we have that \(\frac{r}{k-1} < 2\) and so \(q = 1\) in this case. Thus
\[
\dim(S) \leq nr - \frac{r^2}{2} - \frac{3r}{2} + k - 1
\]
\[
= nr - \frac{r^2}{2} - \frac{r}{2} + \frac{q^2}{2} (k - 1) + \frac{q}{2} (-2r + k - 1).
\]

**Case 3: \(k' = n - k < \min\{k, r + 2 - k\}\)**

In this case, we have that \(n - k < r + 2 - k\). This implies that \(r + 2 > n\) and this is only possible if \(Q\) consists of a single \(n \times n\) Jordan block. Thus we have that the rank bound for \(S\) is \(r = n - 1\) and the maximal nilindex is \(k = n\). In this degenerate case, \(r' = 0\), \(n' = 0\) and \(k' = 1\). So \(\dim(S') = 0\) and we obtain that
\[
\dim(S) = \dim(W_S) + \dim(W_{U_S}) + \dim(S')
\]
\[
\leq (k - 1)(n - k) + \frac{(k - 1)k}{2} + 0
\]
\[
= \frac{n(n - 1)}{2}.
\]
It is easily verified that for \(r = n - 1\) and \(k = n\), our formula collapses to the Gerstenhaber bound from Theorem 2.6 as well, so in this case we also obtain that
\[
\dim(S) \leq nr - \frac{r^2}{2} - \frac{r}{2} + \frac{q^2}{2} (k - 1) + \frac{q}{2} (-2r + k - 1).
\]
This completes the proof of Theorem 5.1.

Suppose \(S\) is a subspace of \(M_n(\mathbb{F})\), which consists of nilpotents of bounded rank \(r\), but we are given no additional information on the maximal nilindex \(k\). By considering Jordan forms, it is clear that all we can say is that the largest the nilindex can be is \(k = r + 1\). As our bound in Theorem 5.1 increases with \(k\), we obtain the following corollary as a special case Theorem 5.1, which recaptures a result of Brualdi and Chavey [1].
Corollary 5.2. [1] Suppose $S$ is a subspace of $M_n(\mathbb{F})$ and $S$ consists of nilpotent matrices of rank less than or equal to $r$. If $\text{card}(\mathbb{F}) > n$ then

$$\dim(S) \leq nr - \frac{r^2}{2} - \frac{r}{2}.$$

Proof. By the comments preceding the Corollary, this follows by setting $k = r + 1$ (and so $q = 1$) in the bound in Theorem 5.1.

Another special case of note is when our subspace consists of square-zero matrices. In this case our Theorem recovers the result of [7].

Corollary 5.3. [7] Suppose $S$ is a subspace of $M_n(\mathbb{F})$ and $S$ consists of square-zero matrices of rank less than or equal to $r$. If $\text{card}(\mathbb{F}) > n$ then

$$\dim(S) \leq nr - r^2.$$

Proof. Apply Theorem 5.1, with $k = 2$.

One last special case that deserves consideration is when $S$ is a subspace of $M_n(\mathbb{F})$, which consists of nilpotents of maximal nilindex $k$, but we are given no additional information concerning a bound on the rank of matrices in $S$. By considering the Jordan form of matrices in $S$, and noting that each Jordan block is at most $k \times k$, we determine that there must be at least $\lceil \frac{n}{k} \rceil$ Jordan blocks in the Jordan form. (For a real number $x$, $\lceil x \rceil$ is the ceiling function of $x$, i.e. the smallest integer which is greater than or equal to $x$.) Each block has one-dimensional kernel, so the dimension of the kernel of a matrix in $S$ must be at least $\lceil \frac{n}{k} \rceil$. Hence, the ranks of matrices in $S$ are bounded by $r = n - \lceil \frac{n}{k} \rceil$. This gives the following cumbersome corollary.

Corollary 5.4. Suppose $S$ is a subspace of $M_n(\mathbb{F})$ and $S$ consists of nilpotent matrices of maximal nilindex $k$. If $\text{card}(\mathbb{F}) > n$ then

$$\dim(S) \leq n \left( n - \left\lceil \frac{n}{k} \right\rceil \right) - \frac{1}{2} \left( n - \left\lceil \frac{n}{k} \right\rceil \right)^2 - \frac{1}{2} \left( n - \left\lceil \frac{n}{k} \right\rceil \right) + \frac{1}{2} \left( \left\lceil \frac{n}{k} \right\rceil - 1 \right) \left( k - 1 \right)$$

$$+ \frac{1}{2} \left( \left\lceil \frac{n}{k} \right\rceil - 1 \right) \left( -2 \left( n - \left\lceil \frac{n}{k} \right\rceil \right) + k - 1 \right).$$

Proof. By the comments preceding the Corollary, this follows by applying Theorem 5.1, with

$$r = n - \left\lceil \frac{n}{k} \right\rceil \quad \text{and} \quad q = \left\lfloor \frac{n - \left\lceil \frac{n}{k} \right\rceil}{k - 1} \right\rfloor.$$

In the case where $k$ divides $n$, this formula simplifies significantly.
Corollary 5.5. Suppose $S$ is a subspace of $M_n(\mathbb{F})$ and $S$ consists of nilpotent matrices of maximal nilindex $k$. If $k$ divides $n$ and $\text{card}(\mathbb{F}) > n$ then
\[ \dim(S) \leq n^2 \left( \frac{k-1}{2k} \right) . \]

Proof. In this case, the bound for the rank is $r = n - \frac{n}{k}$, so $q = \frac{n}{k}$. Applying Theorem 5.1 with these values gives the result. □

The above two Corollaries consider a case contained in the paper of Brualdi and Chavey [1]. Corollary 3.6 of that paper says the following.

Corollary 5.6. [1] Let $k$ be an integer with $2 \leq k \leq n$. Let $W$ be a linear space of nilpotents in $M_n(\mathbb{F})$, each with index at most $k$. If $\mathbb{F}$ is sufficiently large, then
\[ \dim(W) \leq \frac{1}{2} \left( n^2 - \sum_{i=1}^{k} c_i^2 \right) , \]
where $\gamma = (c_1, c_2, \ldots, c_k)$ is a partition of $n$ into $k$ parts with parts differing by at most 1.

The definition of partition of $n$ used in Brualdi and Chavey [1] is a non-increasing sequence of natural numbers which sum to $n$. It is not too difficult to see that, in the case where $k$ divides $n$, the best bound is achieved by the constant partition $\gamma = \left( \frac{n}{k}, \frac{n}{k}, \ldots, \frac{n}{k} \right)$, in which case Corollary 5.6 gives the same bound as our Corollary 5.5. In the general case, it is possible that Corollary 5.6 gives the same bound as our Corollary 5.4.

6. Construction of Nilpotent Spaces of Maximal Rank

In this section we show that our bound in Theorem 5.1 is sharp by constructing subspaces of nilpotent spaces of maximal dimension in all feasible cases. A feasible case is a triple $(k, r, n)$ for which there exists a subspace $S_{k,r,n}$ in $M_n(\mathbb{F})$ consisting of nilpotent matrices of maximal rank $r$, and maximal nilindex $k$.

As mentioned previously, consideration of possible Jordan forms of matrices in $S_{k,r,n}$ leads immediately to the condition that $0 \leq k-1 \leq r \leq n-1$.

Also note that, for $k$ is the maximal nilindex, each matrix $A \in S_{k,r}$ in $M_n(\mathbb{F})$ must have at least $\left\lceil \frac{n}{k} \right\rceil$ blocks in its Jordan form. With each block there corresponds a zero column and so $A$ has at most $n - \left\lfloor \frac{n}{k} \right\rfloor$ non-zero columns. Thus, we obtain that $r \leq n - \left\lfloor \frac{n}{k} \right\rfloor$, or equivalently $k \geq \frac{n}{n-r}$.

These are the only conditions on the triple $(k, r, n)$ required for feasibility.

Theorem 6.1. For each triple $(k, r, n)$ of non-negative integers satisfying
\[ \frac{n}{n-r} \leq k \leq r + 1 \leq n \]
there exists a subspace $S_{k,r,n}$ in $M_n(\mathbb{F})$, consisting of nilpotent matrices of maximal rank $r$, and maximal nilindex $k$ and having dimension
\[ \dim(S_{k,r,n}) = nr - \frac{r^2}{2} - \frac{r}{2} + \frac{q^2}{2}(k-1) + \frac{q}{2}(-2r+k-1) . \]
Proof. Given \( n, k \) and \( r \) as above, let \( q \geq \left\lfloor \frac{r}{k-1} \right\rfloor \), and let \( s = r - q(k - 1) \). So \( r = q(k - 1) + s \) where \( 0 \leq s < k - 1 \). Then decompose \( \mathbb{F}^n \) as

\[
\mathbb{F}^n = \mathbb{F}^k \oplus \mathbb{F}^k \oplus \ldots \oplus \mathbb{F}^k \oplus \mathbb{F}^s \oplus \mathbb{F}^{n-s}.
\]

Let \( \mathcal{S}_{k,r,n} \) be the set of all matrices \( S \) in \( M_n(\mathbb{F}) \) whose \((q+2)\times(q+2)\) block matrices 

\[
[S_{ij}]_{i,j=1}^{q+2}
\]

with respect to the above decomposition, satisfy the following conditions:

1. For \( 1 \leq i, j \leq q \), \( S_{ij} \) is strictly lower-triangular \( k \times k \) matrix;
2. For \( 1 \leq i \leq q \), \( S_{i,q+1} \) is a \( k \times s \) matrix which has its \( k - 1 \) highest diagonals \((D_{s-k+1}, \ldots, D_{s-1})\) equal to zero;
3. For \( 1 \leq i \leq n \), \( S_{i,q+2} = 0 \);
4. For \( 1 \leq j \leq q \), \( S_{q+1,j} \) has its \( s + 1 \) highest diagonals \((D_{s-k}, \ldots, D_{s-1})\) equal to zero;
5. \( S_{q+1,q+1} \) is a lower triangular \( s \times s \) matrix;
6. For \( 1 \leq j \leq q \), \( S_{q+2,j} \) is an \((n-r-q)\times k\) matrix with last column consisting of zeroes;
7. \( S_{q+2,q+1} \) is an \((n-r-q)\times s\) matrix with last column consisting of zeroes.

In general it is clear that the rank is bounded by \( r \) (since there will be \( n-r \) columns of zeroes), and this rank is achieved. It is slightly less obvious that the maximal nilindex is \( k \). When you compute \( X^2 \) all the \((i,j)\) block entries move down or left creating more zeroes, except in the \((q,q+1)\) block where no new zero columns are created. But for each power after that, in every block a new column or diagonal of zeroes is created and since \( s < k - 1 \), we obtain that \( X^k = 0 \).

Now we calculate the dimension of \( \mathcal{S}_{k,r,n} \) by computing the dimension of each block.

We need to consider the cases where \( s = 0 \) and \( s > 0 \) separately. The \( s = 0 \) case is simpler and we leave that to the reader. In the case \( s > 0 \) we have the following dimension argument.

Clearly the blocks in the \((i,j)\)-th entries for \( 1 \leq i, j \leq q \) contribute \( q^2 \frac{k(k-1)}{2} \) dimensions. The \((q+1,q+1)\) block contributes \( \frac{s(s-1)}{2} \) dimensions. The blocks in the \((q+2,j)\) entries contribute \( (n-r-q)(kq + s - (q+1)) = (n-r-q)(r) \) dimensions, since there are \( n-r-q \) entries in each column and there are \( kq+s \) total columns but \( q \) of them are zero. By pairing off the \((q+1,i)\) entry with the \((i,q+1)\) entry (for \( i = 1, 2, \ldots, q \), we see that each paired block has exactly \((k-1)s\) arbitrary entries and there are \( q \) such blocks so they contribute \( qs(k-1) \) dimensions.

Thus

\[
\dim(\mathcal{S}_{k,r,n}) = q^2 \frac{k(k-1)}{2} + \frac{s(s-1)}{2} + (n-r-q)(r-1) + qs(k-1).
\]

Substituting \( s = r - q(k - 1) \) and simplifying, we obtain the formula in Theorem 5.1. \( \square \)

**Example 6.1.** If we apply Theorem 6.1, in the case where \( n = 12, r = 8 \) and \( k = 4 \), then \( \mathcal{S}_{4,8,12} \) consists of all matrices \( X \) of the following form, where entries
indicated by an * are arbitrary elements of \( \mathbb{F} \):

\[
X = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
* & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
* & * & 0 & * & * & 0 & * & 0 & * & 0 & * & 0 & * & 0 & 0 & * & 0 & * & 0 & * & 0 & 0 \\
* & * & * & 0 & * & * & * & 0 & * & * & 0 & * & * & 0 & * & * & 0 & * & * & 0 & * & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
* & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
* & * & 0 & * & * & 0 & * & 0 & * & 0 & * & 0 & * & 0 & 0 & * & 0 & * & 0 & * & 0 & 0 \\
* & * & * & 0 & * & * & * & 0 & * & * & 0 & * & * & 0 & * & * & 0 & * & * & 0 & * & 0 & 0 \\
* & * & * & 0 & * & * & * & 0 & * & * & 0 & * & * & 0 & * & * & 0 & * & * & 0 & * & 0 & 0
\end{bmatrix}.
\]

Theorem 5.1 and Theorem 6.1 give that \( \dim(S_{4,8,12}) = 51 \) which can be verified by counting the number of * in the above matrix.

If we only use the information about the rank bound and apply Corollary 5.2, we obtain a weaker bound that \( \dim(S_{4,8,12}) \leq 68 \), while if we only use information about the nilindex bound, and apply Corollary 5.5 (since \( k \) divides \( n \)), we obtain an improved but still less than optimal bound that \( \dim(S_{4,8,12}) \leq 54 \).

Note that in [3], Gerstenhaber also provides a general bound for the dimension of a subspace \( S \) of nilpotent matrices, in terms of all the possible sizes of Jordan blocks which occur the Jordan form of all the matrices in \( S \). This would, most likely, be difficult to determine in any particular case. One advantage of our formula in that it depends only on a rank bound and a nilindex bound; information that may be more accessible for a subspace of nilpotents \( S \) than the possible sizes of blocks in all possible Jordan decompositions of matrices in \( S \).

In cases where information about the possible structure of Jordan forms of matrices in the subspace is available, we can use our methods to give a dimensional bound in these cases as well, and this bound improves that of Gerstenhaber in many cases.

### 7. Improving Gerstenhaber’s General Theorem

In order to state the Gerstenhaber’s General Theorem, which relates the dimensional bound on a subspace of nilpotent matrices to information about the sizes of Jordan blocks in Jordan forms of matrices in \( S \), we need some preliminary terminology. We use the notation and terminology of [1].

For \( n \) a positive integer, we say \( \alpha = (a_1, a_2, \ldots, a_n) \) is a partition of \( n \) if \( a_1 \geq a_2 \geq \cdots \geq a_n \geq 0 \), and \( n = a_1 + a_2 + \cdots + a_n \).

The conjugate of the partition \( \alpha \) is the partition of \( n \) defined by \( \alpha^* = (a_1^*, a_2^*, \ldots, a_n^*) \) where \( a_i^* \) is the number of \( a_i \) in \( \alpha \) which are greater than or equal to \( j \).

We define a partial order \( \preceq \) on the set of partitions of \( n \) as follows: if \( \alpha = (a_1, a_2, \ldots, a_n) \) and \( \beta = (b_1, b_2, \ldots, b_n) \) are two partitions of \( n \), then \( \alpha \preceq \beta \) if

\[
a_1 + \cdots + a_j \leq b_1 + \cdots + b_j \quad \text{for all } j = 1, 2, \ldots, n.
\]

To each nilpotent matrix \( A \in M_n(\mathbb{F}) \) we associate a partition of \( n \) as follows: Let \( k_1 \geq k_2 \geq \cdots \geq k_l \geq 1 \) be the sizes of the Jordan blocks in the Jordan form of
$A$. Then $n = k_1 + k_2 + \cdots + k_l$ and so

$$jp(A) = (k_1, k_2, \ldots, k_l, 0, \ldots, 0)$$

is a partition of $n$ (we have adjoined $n-l$ zeros) which is called the Jordan partition of $n$.

If $S$ is a subspace of $M_n(F)$ consisting of nilpotent matrices, the Jordan partition of $S$ is the least upper bound (in the partial order mentioned above) of the set of all Jordan partitions of matrices in $S$.

The Gerstenhaber’s General Theorem [3] is the following.

**Theorem 7.1** (Gerstenhaber). Suppose $S$ is a subspace of $M_n(F)$ consisting of nilpotent matrices, and that $\gamma = (c_1, \ldots, c_n)$ is the conjugate of the Jordan partition of $S$. If $\text{card}(F)$ is sufficiently large, then

$$\dim(S) \leq \frac{1}{2} \left( n^2 - \sum_{i=1}^{n} c_i^2 \right)$$

As this bound involves only possible sizes of Jordan blocks of matrices in $S$, it can still be quite coarse. Consider the following two examples.

**Example 7.1.** For $n$ an even positive integer, and $k = \frac{n}{2}$, let

$$S_1 = \left\{ \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} : X_{11}, X_{12}, X_{21}, X_{22} \in U_k(F) \right\}$$

and let

$$S_2 = \left\{ \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} : X_{11}, X_{12}, X_{21}, X_{22} \in U_k(F), X_{11} = X_{22} \right\}$$

In both cases it is obvious that the maximal nilindex is $k$ and so it immediately follows that $jp(S_1) = jp(S_2) = (k, k, 0, \ldots, 0)$ and the conjugate of this partition is $\gamma = (2, \ldots, 2, 0 \ldots 0)$ where there are $k$ twos and $k$ zeros. Thus, for $i = 1, 2$, Gerstenhaber’s General Theorem gives the bound of

$$\dim(S_i) \leq \frac{1}{2} \left( n^2 - \sum_{i=1}^{k} 2^2 \right)$$

$$= \frac{1}{2} \left( n^2 - \frac{n}{2} (4) \right) = \frac{n(n-2)}{2} \text{ for } i = 1, 2.$$

From direct computation it is easy to see that this bound is sharp in the case of $S_1$, but the actual dimension of $S_2$ is three quarters of this bound.

We offer an finer bound which can distinguish between subspaces which have the same Jordan partition by introducing a spatial component.

**Definition 7.1.** For $S$ a subspace of $M_n(F)$ consisting of nilpotents, we define a spatial Jordan partition of $S$ to be

$$\zeta = (k_1, k_2, \ldots, k_l)$$

where $k_1 \geq k_2 \geq \cdots k_l \geq 1$ are inductively defined as follows:
(1) \(k_1\) is the maximal nilindex of matrices in \(S\). Fix an \(A_1\) in \(S\) which has this maximal index, and let \(F^n = M_1 \oplus N_1\), where \(A_1\) has block matrix
\[
\begin{bmatrix}
J_{k_1} & 0 \\
0 & N
\end{bmatrix}
\]
with respect to this decomposition. Let \(S_1\) be the subspace of all matrices in \(S\) whose block matrix with respect to this decomposition is of the form
\[
\begin{bmatrix}
0 & 0 \\
0 & X
\end{bmatrix}.
\]

(2) Inductively define \(k_{i+1}, M_{i+1}, N_{i+1}\) and \(S_{i+1}\) as follows: Choose \(A_{i+1} \in S_i\) of maximal nilindex \(k_{i+1}\) in \(S_i\). Let \(N_i = M_{i+1} \oplus N_{i+1}\), so that, with respect to the decomposition \(F^n = (M_1 \oplus \cdots M_i) \oplus M_{i+1} \oplus N_{i+1}\), \(A_i\) has block matrix
\[
\begin{bmatrix}
0 & 0 & 0 \\
0 & J_{k_{i+1}} & 0 \\
0 & 0 & N
\end{bmatrix},
\]
and let \(S_{i+1}\) be the subspace of all matrices in \(S\) whose block matrix with respect to this decomposition is of the form
\[
\begin{bmatrix}
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & X
\end{bmatrix}.
\]

Note that, if at any point \(S_i = \{0\}\) then \(k_j = 1\) for all \(j = i+1, \ldots, l\), as these correspond to one-dimensional Jordan blocks (i.e. zero matrices).

We are now ready to state our general theorem providing dimensional bounds on subspaces of nilpotent matrices in terms of spatial Jordan partitions.

**Theorem 7.2.** Suppose \(S\) is a subspace of \(M_n(F)\) consisting of nilpotent matrices, and that \(\zeta = (k_1, \ldots, k_l)\) is a spatial Jordan partition of \(S\). If \(\text{card}(F) > n\) then
\[
dim(S) \leq \sum_{i=1}^{l} (k_i - 1) \left( \frac{k_i}{2} + n - \sum_{j=1}^{i} k_j \right).
\]

**Proof.** The proof requires only our Dimensional Slicing Lemma (Lemma 2.5), our lemma for bounding off diagonal terms (Lemma 4.1) and the dimensional bound for a general subspace of nilpotents (Theorem 2.6).

Let \(A_i\) and \(S_i\), for \(i = 1, 2, \ldots, l\), be as in Definition 7.1.

With no loss of generality (by applying a similarity), we may assume that \(A_1\) is in its Jordan form, with Jordan blocks arranged in order of decreasing size. So, with respect to the decomposition \(F^n = F^{k_1} \oplus F^{n-k_1}\)
\[
A_1 = \begin{bmatrix} J_{k_1} & 0 \\ 0 & N \end{bmatrix}.
\]

We proceed similarly to the proof of Theorem 5.1. Consider the index set \(I = \{1, 2, \ldots, k_1\} \times \{k_1 + 1, k_1 + 2, \ldots, n\} \cup \{k_1 + 1, k_1 + 2, \ldots, n\} \times \{1, 2, \ldots, k_1\}\), and for this index set define \(W_S\) and \(U_S\) as in Lemma 2.5. Then \(W_S\) is the subspace of
$M_n(\mathbb{F})$ spanned that all the entries that appear in the $(1, 2)$ and $(2, 1)$ locations of the matrices in $S$, and $U_S$ will be the set of all the matrices in $S$ of the form

$$\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$$

By Lemma 2.5,

$$\dim(S) = \dim(W_S) + \dim(U_S)$$

and by Lemma 4.1

$$\dim(W_S) \leq (k_1 - 1)(n - k_1).$$

To find $\dim(U_S)$, we again apply Lemma 2.5 to $U_S$ with the index set $I = \{1, 2, \ldots, k_1\} \times \{1, 2, \ldots, k_1\}$. Then we obtain

$$W_{U_S} = \left\{ \begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix} : \text{there exists } A_2 \text{ so that } \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \in S \right\}$$

and

$$U_{U_S} = \left\{ \begin{bmatrix} 0 & 0 \\ 0 & A_2 \end{bmatrix} : \begin{bmatrix} 0 & 0 \\ 0 & A_2 \end{bmatrix} \in S \right\}.$$ The elements of $W_{U_S}$ are $k_1 \times k_1$ nilpotent matrices so by Theorem 2.6,

$$\dim(W_{U_S}) \leq \frac{(k_1 - 1)k_1}{2}.$$ Now, $U_{U_S}$ is a set of $(n - k_1) \times (n - k_1)$ nilpotent matrices, which is our set $S_1$. So

$$\dim(S) = \dim(W_S) + \dim(U_S)$$

$$\leq (k_1 - 1)(n - k_1) + \dim(W_{U_S}) + \dim(U_{U_S})$$

$$\leq (k_1 - 1)(n - k_1) + \frac{k_1(k_1 - 1)}{2} + \dim(S_1).$$

We now repeat this argument to bound $\dim(S_1)$ in terms of the new maximal nilindex $k_2$, the new dimension $n - k_1$ and the dimension of $S_2$ to obtain

$$\dim(S) \leq (k_1 - 1)(n - k_1) + \frac{k_1(k_1 - 1)}{2} + \dim(S_1)$$

$$\leq (k_1 - 1)(n - k_1) + \frac{k_1(k_1 - 1)}{2} + (k_2 - 1)(n - k_1 - k_2)$$

$$+ \frac{k_2(k_2 - 1)}{2} + \dim(S_2).$$

Repeating $l$ times and simplifying we obtain the required bound. $\square$

In Example 7.1 it is not to difficult to see that a spatial Jordan partition of $S_1$ is $\zeta_1 = (k, k)$ while a spatial Jordan partition of $S_2$ is $\zeta_2 = (k_1, 1, \ldots, 1)$ (there are $k$ ones). For the first subspace, our Theorem 7.2 gives the same bound as Gerstenhaber’s General Theorem

$$\dim(S_1) \leq \frac{n(n - 2)}{2}$$

but for the second subspace, our Theorem 7.2 gives the improved bound

$$\dim(S_2) \leq \frac{3n(n - 2)}{8}$$

which is sharp.
Informally, it could be said that Gerstenhaber’s General Theorem is unable to identify and account for multiplicity in the Jordan forms of matrices in the subspace, while our bound takes multiplicity into account and so it should give bound which is at least as good as Gerstenhaber’s in all cases. At present we are unable to determine if this is the case but are willing to make the following conjecture.

**Conjecture 7.3.** Suppose $S$ is a subspace of $M_n(F)$ consisting of nilpotent matrices and $\text{card}(F) > n$. Then the dimensional bound in Theorem 7.2 is less than or equal to the dimensional bound in Theorem 7.1.

Serežkin eventually improved Gertenhaber’s Theorem to remove any condition on the underlying field while retaining the same bound on the dimension of the subspace of nilpotent matrices. Is it possible the hypothesis that $\text{card}(F) > n$ could be dropped in Theorem 7.2 and still retain the dimensional bound in the conclusion? This is an open problem. We know of no counterexample in small fields, but do not have enough evidence or intuition to conjecture in either direction.

**References**


