Finding a Jordan basis for a nilpotent operator

Leonardo T. Rolla

April 15, 2020

1 The algorithm

We prove the following proposition:

For finite-dimensional G and nilpotent $N \in \mathcal{L}(G)$ there is a Jordan basis.

The proof is explicit and provides an algorithm to actually find a Jordan basis. We define the notion of *thread* as a chain of vectors like this:

$$u \mapsto v \mapsto w \mapsto \cdots \mapsto z \mapsto \ ?.$$

Here "$\mapsto$" means the image under N. The above example is an *open thread*. If we know that $Nz = \mathbf{0}$, we can have a *closed thread*

$$u \mapsto v \mapsto w \mapsto \cdots \mapsto z \mapsto \mathbf{0}.$$

In both cases, the *tip* of the thread is z and the *base* is u. The *length* of a thread is the number of non-zero vectors.

The object we will be dealing with all the time is a *collection of threads.* The *size* of a collection is the sum of the lengths of the threads.

We introduce two operations on collections of threads: *stretch* and *reduce.*

We can *stretch* a collection of threads if at least one of the threads is still open. The operation consists in applying N to the tip of open threads and

appending the result at the end. This causes the open threads to either increase in length or become closed.

We can *reduce* a collection of threads if the tips are linearly dependent. To do that, we sort the collection from longest to shortest, find which ones are linear combinations of previous ones, and subtract the corresponding linear combination at their base. The vectors to be used in this linear combination are those with the same distance form the tip as the base of the concerned thread. The result of this subtraction will propagate throughout the thread, and at some point the thread will become zero, either at the tip or maybe before. This causes the size of the collection to decrease.

This process must stop. Indeed, each time we stretch, we enlarge the longest thread or we close a thread, and the longest thread cannot be longer than $\dim G$; moreover, between two stretches we can only perform finitely many reductions. When the process stops, we have a collection of closed threads whose tips are linearly independent. We claim that, in this case, the whole collection of vectors is linearly independent.

We now prove the above claim. Consider an arbitrary finite collection of closed threads whose tips are linearly independent. Denote such a collection by

$$
\begin{aligned}
v_{d_1}^1 &\mapsto v_{d_1-1}^1 \mapsto \cdots \mapsto v_2^1 \mapsto v_1^1 \mapsto v_0^1 \mapsto \mathbf{0} \\
v_{d_2}^2 &\mapsto v_{d_2-1}^2 \mapsto \cdots \mapsto v_2^2 \mapsto v_1^2 \mapsto v_0^2 \mapsto \mathbf{0} \\
&\vdots \\
v_{d_m}^m &\mapsto v_{d_m-1}^m \mapsto \cdots \mapsto v_2^m \mapsto v_1^m \mapsto v_0^m \mapsto \mathbf{0}
\end{aligned}
$$

with $d = d_1 \geqslant d_2 \geqslant \ldots \geqslant d_m \geqslant 1$. Suppose there is a linear combination

$$
\sum_{j=1}^m \sum_{k=0}^{d_m} \alpha_{j,k} v_k^j = \mathbf{0}.
$$

Applying N^d to this identity gives

$$
\mathbf{0} = \sum_j \alpha_{j,d} N^d v_k^j = \sum_j \alpha_{j,d} v_0^j.
$$

Since $v_0^1, \ldots, v_0^m$ are linearly independent, we have $\alpha_{j,d} = 0$ for all relevant j. Now we apply N^{d-1} to the same identity and use that $\alpha_{j,d} = 0$ to get

$$
\mathbf{0} = \sum_j \alpha_{j,d-1} N^{d-1} v_k^j = \sum_j \alpha_{j,d-1} v_0^j.
$$

Again, by linear independence of $v_0^1, \ldots, v_0^m$, we have $\alpha_{j,d-1} = 0$ for all relevant j. Proceeding this way, we end up concluding that all the coefficients are zero. This concludes the proof of the claim.

We now finish the proof of the proposition. Start with a collection of threads whose vectors span G. For instance, let $v_1, \ldots, v_n$ be a basis for G and consider the collection

$$v_1 \mapsto \ ?, \quad v_2 \mapsto \ ?, \quad \ldots \quad , \quad v_n \mapsto \ ?,$$

of very short open threads. This collection spans G.

Now notice that the operation of reduction only adds multiples of some vectors in the collection to some other vectors, so it does not change the space spanned by the collection as a whole. Moreover, the operation of stretching only incorporates more vectors into the collection.

Therefore, the collection we get in the end is a basis which forms closed threads, which is just the definition of Jordan basis rephrased.

This concludes the proof.

As a remark, if for some reason we start with a collection which is not spanning, but get $\mathsf{dim}\, G$ vectors in the end, then it is anyway a Jordan basis.

2   Example

Let $G = \mathbb{F}^4$. Define $N \in \mathcal{L}(G)$ with respect to the canonical basis by

$$[N] = \begin{bmatrix} 1 & 2 & 0 & 5 \\ 6 & -5 & 17 & -38 \\ -3 & 0 & -6 & 9 \\ -2 & 1 & -5 & 10 \end{bmatrix}.$$

Note that $N^3 = \mathbf{0}$.

Start with short threads $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \mathcal{A}_4$:

$$\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \mapsto \ ?, \quad \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \mapsto \ ?, \quad \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \mapsto \ ?, \quad \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \mapsto \ ?.$$

Note that these four vectors span G.

The tips of the threads are linearly independent, so we cannot apply reduction. We stretch the collection, getting new threads $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \mathcal{A}_4$:

$$\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 3 \\ 1 \\ -3 \\ -1 \end{bmatrix} \mapsto \text{ ?}, \quad \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 2 \\ 12 \\ -6 \\ -4 \end{bmatrix} \mapsto \text{ ?}$$

$$\begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} 5 \\ -21 \\ 3 \\ 5 \end{bmatrix} \mapsto \text{ ?}, \quad \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} 7 \\ -43 \\ 9 \\ 11 \end{bmatrix} \mapsto \text{ ?}$$

We could stretch again, but in order to keep the collection small and avoid unnecessary work, let us try to reduce it instead.

Looking at the tips only, by row reduction:

$$[v_1, v_2, v_3, v_4] = \begin{bmatrix} 3 & 2 & 5 & 7 \\ 1 & 12 & -21 & -43 \\ -3 & -6 & 3 & 9 \\ -1 & -4 & 5 & 11 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 & 5 \\ 0 & 1 & -2 & -4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

so v_1 and v_2 are LI, whereas $v_3 = 3v_1 - 2v_2$ and $v_4 = 5v_1 - 4v_2$.

Now we subtract the corresponding multiples from the whole threads, getting:

$$\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 3 \\ 1 \\ -3 \\ -1 \end{bmatrix} \mapsto \text{ ?}, \quad \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 2 \\ 12 \\ -6 \\ -4 \end{bmatrix} \mapsto \text{ ?}, \quad \begin{bmatrix} -3 \\ -1 \\ 3 \\ 1 \end{bmatrix} \mapsto \mathbf{0}, \quad \begin{bmatrix} -5 \\ 0 \\ 4 \\ 1 \end{bmatrix} \mapsto \mathbf{0}.$$

Let us try to reduce again. Looking at the tips only, by row reduction:

$$[v_1, v_2, v_3, v_4] = \begin{bmatrix} 3 & 2 & -3 & -5 \\ 1 & 12 & -1 & 0 \\ -3 & -6 & 3 & 4 \\ -1 & -4 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

so v_1, v_2, v_4 are LI, whereas $v_3 = -v_1$. We want to subtract $(-1)\mathcal{A}_1$ from $\mathcal{A}_3$. Since $\mathcal{A}_3$ only has one vector, we only use the last vector of $\mathcal{A}_1$.

After sorting by size, we get:

$$\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 3 \\ 1 \\ -3 \\ -1 \end{bmatrix} \mapsto \mathbf{?}, \quad \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 2 \\ 12 \\ -6 \\ -4 \end{bmatrix} \mapsto \mathbf{?}, \quad \begin{bmatrix} -5 \\ 0 \\ 4 \\ 1 \end{bmatrix} \mapsto \mathbf{0}, \quad \mathbf{0} \mapsto \mathbf{0}.$$

The tips of the threads are LI, so we can no longer reduce.

So we stretch the collection and sort by size:

$$\begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 2 \\ 12 \\ -6 \\ -4 \end{bmatrix} \mapsto \begin{bmatrix} 6 \\ 2 \\ -6 \\ -2 \end{bmatrix} \mapsto \mathbf{?}, \quad \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 3 \\ 1 \\ -3 \\ -1 \end{bmatrix} \mapsto \mathbf{0}, \quad \begin{bmatrix} -5 \\ 0 \\ 4 \\ 1 \end{bmatrix} \mapsto \mathbf{0}.$$

Let us try to reduce again. Looking at the tips only, by row reduction:

$$[v_1, v_2, v_3] = \begin{bmatrix} 6 & 3 & -5 \\ 2 & 1 & 0 \\ -6 & -3 & 4 \\ -2 & -1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

so v_1, v_3 are LI and $v_2 = \frac{1}{2}v_1$.

As $\mathcal{A}_2$ has 2 vectors, we subtract $\frac{1}{2}$ times the 2 last vectors of $\mathcal{A}_1$ from the two last vectors $\mathcal{A}_2$.

This gives

$$\begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 2 \\ 12 \\ -6 \\ -4 \end{bmatrix} \mapsto \begin{bmatrix} 6 \\ 2 \\ -6 \\ -2 \end{bmatrix} \mapsto \mathbf{0}, \quad \begin{bmatrix} 0 \\ -5 \\ 3 \\ 2 \end{bmatrix} \mapsto \mathbf{0}, \quad \begin{bmatrix} -5 \\ 0 \\ 4 \\ 1 \end{bmatrix} \mapsto \mathbf{0}.$$

Let us try to reduce again. Looking at the tips only, by row reduction:

$$[v_1, v_2, v_3] = \begin{bmatrix} 6 & 0 & -5 \\ 2 & -5 & 0 \\ -6 & 3 & 4 \\ -2 & 2 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -\frac{5}{6} \\ 0 & 1 & -\frac{1}{3} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Subtracting the corresponding linear combination from the third thread will make it vanish, giving:

$$
\begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 2 \\ 12 \\ -6 \\ -4 \end{bmatrix} \mapsto \begin{bmatrix} 6 \\ 2 \\ -6 \\ -2 \end{bmatrix} \mapsto \mathbf{0}, \qquad \begin{bmatrix} 0 \\ -5 \\ 3 \\ 2 \end{bmatrix} \mapsto \mathbf{0}.
$$

As we saw in the previous proof, since we started with a spanning collection, the above threads also span G.

So we can tell that the above 4 vectors are spanning, hence a basis. Since they form closed threads, they form a Jordan basis.

So we did find a Jordan basis!

We already know the basis $\mathcal{B}$ and Jordan form $[N]_{\mathcal{B}}$. Let's double-check:

$$
Q = \begin{bmatrix} 6 & 2 & 0 & 0 \\ 2 & 12 & 1 & -5 \\ -6 & -6 & 1 & 3 \\ -2 & -4 & 0 & 2 \end{bmatrix} \implies Q^{-1}[N]Q = \left[\begin{array}{ccc|c} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array} \right].
$$

3 Stretching first

In the previous example we reduced the threads whenever possible.

Maybe stretching first results in fewer computations, especially row reductions.

Stretching all the way until the end, and sorting by size:

$$
\begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 2 \\ 12 \\ -6 \\ -4 \end{bmatrix} \mapsto \begin{bmatrix} 6 \\ 2 \\ -6 \\ -2 \end{bmatrix} \mapsto \mathbf{0}, \qquad \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} 5 \\ -21 \\ 3 \\ 5 \end{bmatrix} \mapsto \begin{bmatrix} -12 \\ -4 \\ 12 \\ 4 \end{bmatrix} \mapsto \mathbf{0},
$$

$$
\begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} 7 \\ -43 \\ 9 \\ 11 \end{bmatrix} \mapsto \begin{bmatrix} -24 \\ -8 \\ 24 \\ 8 \end{bmatrix} \mapsto \mathbf{0}, \qquad \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 3 \\ 1 \\ -3 \\ -1 \end{bmatrix} \mapsto \mathbf{0}.
$$

Looking at the tips only, by row reduction:

$$[v_1, v_2, v_3, v_4] = \begin{bmatrix} 6 & -12 & -24 & 3 \\ 2 & -4 & -8 & 1 \\ -6 & 12 & 24 & -3 \\ -2 & 4 & 8 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & -4 & \frac{1}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

so $v_2 = -2v_1$, $v_3 = -4v_1$, and $v_4 = \frac{1}{2}v_1$.

Now we subtract the corresponding multiples from the threads, getting:

$$\begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 2 \\ 12 \\ -6 \\ -4 \end{bmatrix} \mapsto \begin{bmatrix} 6 \\ 2 \\ -6 \\ -2 \end{bmatrix} \mapsto \mathbf{0}, \qquad \begin{bmatrix} 0 \\ 2 \\ 3 \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} 9 \\ 3 \\ -9 \\ -3 \end{bmatrix} \mapsto \mathbf{0},$$

$$\begin{bmatrix} 0 \\ 5 \\ 4 \\ 1 \end{bmatrix} \begin{bmatrix} 15 \\ 5 \\ -15 \\ -5 \end{bmatrix} \mapsto \mathbf{0}, \qquad \begin{bmatrix} 0 \\ -5 \\ 3 \\ 2 \end{bmatrix} \mapsto \mathbf{0}.$$

Looking at the tips only, by row reduction:

$$[v_1, v_2, v_3, v_4] = \begin{bmatrix} 6 & 9 & 15 & 0 \\ 2 & 3 & 5 & -5 \\ -6 & -9 & -15 & 3 \\ -2 & -3 & -5 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & \frac{3}{2} & \frac{5}{2} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

so v_1 and v_4 are LI, whereas $v_2 = \frac{3}{2}v_1$ and $v_3\frac{5}{2}v_1$.

Now we subtract the corresponding multiples from the threads, getting:

$$\begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 2 \\ 12 \\ -6 \\ -4 \end{bmatrix} \mapsto \begin{bmatrix} 6 \\ 2 \\ -6 \\ -2 \end{bmatrix} \mapsto \mathbf{0}, \qquad \begin{bmatrix} -3 \\ -16 \\ 12 \\ 7 \end{bmatrix} \mapsto \mathbf{0}, \qquad \begin{bmatrix} -5 \\ -25 \\ 19 \\ 11 \end{bmatrix} \mapsto \mathbf{0}, \qquad \begin{bmatrix} 0 \\ -5 \\ 3 \\ 2 \end{bmatrix} \mapsto \mathbf{0}.$$

Looking at the tips only, by row reduction:

$$[v_1, v_2, v_3, v_4] = \begin{bmatrix} 6 & -3 & -5 & 0 \\ 2 & -16 & -25 & -5 \\ -6 & 12 & 19 & 3 \\ -2 & 7 & 11 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -\frac{1}{18} & \frac{1}{6} \\ 0 & 1 & \frac{14}{9} & \frac{1}{3} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

so v_1 and v_2 are LI, whereas $v_3 = -\frac{1}{18}v_1 + \frac{14}{9}v_2$ and $v_4 = \frac{1}{6}v_1 + \frac{1}{3}v_2$.

Now we subtract the corresponding multiples from the threads, getting:

$$\begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 2 \\ 12 \\ -6 \\ -4 \end{bmatrix} \mapsto \begin{bmatrix} 6 \\ 2 \\ -6 \\ -2 \end{bmatrix} \mapsto \mathbf{0}, \quad \begin{bmatrix} -3 \\ -16 \\ 12 \\ 7 \end{bmatrix} \mapsto \mathbf{0}, \quad \mathbf{0} \mapsto \mathbf{0}, \quad \mathbf{0} \mapsto \mathbf{0}.$$

We found a Jordan basis!

We already know the basis $\mathcal{B}$ and Jordan form $[N]_{\mathcal{B}}$. Let's double-check:

$$Q = \begin{bmatrix} 6 & 2 & 0 & -3 \\ 2 & 12 & 1 & -16 \\ -6 & -6 & 1 & 12 \\ -2 & -4 & 0 & 7 \end{bmatrix} \implies Q^{-1}[N]Q = \left[\begin{array}{ccc|c} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right].$$

4   Using fewer threads

Pick only one vector from the basis, say $(1, 1, 0, 0)$. Stretch its thread:

$$\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 3 \\ 1 \\ -3 \\ -1 \end{bmatrix} \mapsto \mathbf{0}.$$

This is too short. Let's incorporate another vector, say $(0, 1, 1, 0)$, and stretch:

$$\begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 2 \\ 12 \\ -6 \\ -4 \end{bmatrix} \mapsto \begin{bmatrix} 6 \\ 2 \\ -6 \\ -2 \end{bmatrix} \mapsto \mathbf{0}, \quad \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 3 \\ 1 \\ -3 \\ -1 \end{bmatrix} \mapsto \mathbf{0}.$$

Doing row reduction (yes, that's overkilling) with the tips only:

$$\begin{bmatrix} 6 & 3 \\ 2 & 1 \\ -6 & -3 \\ -2 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Subtracting $\frac{1}{2}$ of the first thread from the second:

$$
\begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 2 \\ 12 \\ -6 \\ -4 \end{bmatrix} \mapsto \begin{bmatrix} 6 \\ 2 \\ -6 \\ -2 \end{bmatrix} \mapsto \mathbf{0}, \qquad \begin{bmatrix} 0 \\ -5 \\ 3 \\ 2 \end{bmatrix} \mapsto \mathbf{0}.
$$

The tips are LI, the threads are closed, and there are four vectors. So the procedure is finished.

We already know the basis $\mathcal{B}$ and Jordan form $[N]_{\mathcal{B}}$. Let's double-check:

$$
Q = \begin{bmatrix} 6 & 2 & 0 & 0 \\ 2 & 12 & 1 & -5 \\ -6 & -6 & 1 & 3 \\ -2 & -4 & 0 & 2 \end{bmatrix} \implies Q^{-1}[N]Q = \left[\begin{array}{ccc|c} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array} \right].
$$