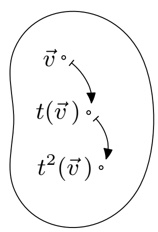
5.III Nilpotence

The goal of this chapter is to show that every square matrix is similar to one that is a sum of two kinds of simple matrices. The prior section focused on the first kind, diagonal matrices. We now consider the other kind.

5.III.1 Self-Composition

This subsection is optional, although it is necessary for later material in this section and in the next one.

A linear transformations $t\colon V \to V$, because it has the same domain and codomain, can be iterated.* That is, compositions of t with itself such as $t^2 = t \circ t$ and $t^3 = t \circ t \circ t$ are defined.



Note that this power notation for the linear transformation functions dovetails with the notation that we've used earlier for their square matrix representations because if $\mathrm{Rep}_{B,B}(t) = T$ then $\mathrm{Rep}_{B,B}(t^j) = T^j$.

1.1 Example For the derivative map $d/dx\colon \mathcal{P}_3 \to \mathcal{P}_3$ given by

$$a + bx + cx^2 + dx^3 \overset{d/dx}{\longmapsto} b + 2cx + 3dx^2$$

the second power is the second derivative

$$a + bx + cx^2 + dx^3 \overset{d^2/dx^2}{\longmapsto} 2c + 6dx$$

the third power is the third derivative

$$a + bx + cx^2 + dx^3 \overset{d^3/dx^3}{\longmapsto} 6d$$

and any higher power is the zero map.

1.2 Example This transformation of the space of 2×2 matrices

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \overset{t}{\longmapsto} \begin{pmatrix} b & a \\ d & 0 \end{pmatrix}$$

* More information on function interation is in the appendix.

has this second power

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \stackrel{t^2}{\longmapsto} \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}$$

and this third power.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \stackrel{t^3}{\longmapsto} \begin{pmatrix} b & a \\ 0 & 0 \end{pmatrix}$$

After that, $t^4 = t^2$ and $t^5 = t^3$, etc.

These examples suggest that on iteration more and more zeros appear until there is a settling down. The next result makes this precise.

1.3 Lemma For any transformation $t\colon V \to V$, the rangespaces of the powers form a descending chain

$$V \supseteq \mathscr{R}(t) \supseteq \mathscr{R}(t^2) \supseteq \cdots$$

and the nullspaces form an ascending chain.

$$\{\vec{0}\,\} \subseteq \mathscr{N}(t) \subseteq \mathscr{N}(t^2) \subseteq \cdots$$

Further, there is a k such that for powers less than k the subsets are proper (if $j < k$ then $\mathscr{R}(t^j) \supset \mathscr{R}(t^{j+1})$ and $\mathscr{N}(t^j) \subset \mathscr{N}(t^{j+1})$), while for powers greater than k the sets are equal (if $j \geq k$ then $\mathscr{R}(t^j) = \mathscr{R}(t^{j+1})$ and $\mathscr{N}(t^j) = \mathscr{N}(t^{j+1})$).

Proof. We will do the rangespace half and leave the rest for Exercise 13. Recall, however, that for any map the dimension of its rangespace plus the dimension of its nullspace equals the dimension of its domain. So if the rangespaces shrink then the nullspaces must grow.

That the rangespaces form chains is clear because if $\vec{w} \in \mathscr{R}(t^{j+1})$, so that $\vec{w} = t^{j+1}(\vec{v})$, then $\vec{w} = t^j(\,t(\vec{v})\,)$ and so $\vec{w} \in \mathscr{R}(t^j)$. To verify the "further" property, first observe that if any pair of rangespaces in the chain are equal $\mathscr{R}(t^k) = \mathscr{R}(t^{k+1})$ then all subsequent ones are also equal $\mathscr{R}(t^{k+1}) = \mathscr{R}(t^{k+2})$, etc. This is because if $t\colon \mathscr{R}(t^{k+1}) \to \mathscr{R}(t^{k+2})$ is the same map, with the same domain, as $t\colon \mathscr{R}(t^k) \to \mathscr{R}(t^{k+1})$ and it therefore has the same range: $\mathscr{R}(t^{k+1}) = \mathscr{R}(t^{k+2})$ (and induction shows that it holds for all higher powers). So if the chain of rangespaces ever stops being strictly decreasing then it is stable from that point onward.

But the chain must stop decreasing. Each rangespace is a subspace of the one before it. For it to be a proper subspace it must be of strictly lower dimension (see Exercise 11). These spaces are finite-dimensional and so the chain can fall for only finitely-many steps, that is, the power k is at most the dimension of V. QED

1.4 Example The derivative map $a + bx + cx^2 + dx^3 \stackrel{d/dx}{\longmapsto} b + 2cx + 3dx^2$ of Example 1.1 has this chain of rangespaces

$$\mathcal{P}_3 \supset \mathcal{P}_2 \supset \mathcal{P}_1 \supset \mathcal{P}_0 \supset \{\vec{0}\,\} = \{\vec{0}\,\} = \cdots$$

and this chain of nullspaces.

$$\{\vec{0}\,\} \subset \mathcal{P}_0 \subset \mathcal{P}_1 \subset \mathcal{P}_2 \subset \mathcal{P}_3 = \mathcal{P}_3 = \cdots$$

1.5 Example The transformation $\pi\colon \mathbb{C}^3 \to \mathbb{C}^3$ projecting onto the first two coordinates

$$\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \stackrel{\pi}{\longmapsto} \begin{pmatrix} c_1 \\ c_2 \\ 0 \end{pmatrix}$$

has $\mathbb{C}^3 \supset \mathscr{R}(\pi) = \mathscr{R}(\pi^2) = \cdots$ and $\{\vec{0}\,\} \subset \mathscr{N}(\pi) = \mathscr{N}(\pi^2) = \cdots$.

1.6 Example Let $t\colon \mathcal{P}_2 \to \mathcal{P}_2$ be the map $c_0 + c_1 x + c_2 x^2 \mapsto 2c_0 + c_2 x$. As the lemma describes, on iteration the rangespace shrinks

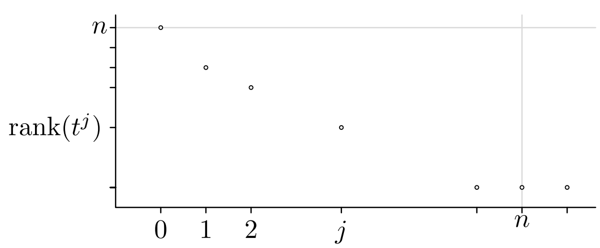
$$\mathscr{R}(t^0) = \mathcal{P}_2 \quad \mathscr{R}(t) = \{a + bx \bigm| a, b \in \mathbb{C}\} \quad \mathscr{R}(t^2) = \{a \bigm| a \in \mathbb{C}\}$$

and then stabilizes $\mathscr{R}(t^2) = \mathscr{R}(t^3) = \cdots$, while the nullspace grows

$$\mathscr{N}(t^0) = \{0\} \quad \mathscr{N}(t) = \{cx \bigm| c \in \mathbb{C}\} \quad \mathscr{N}(t^2) = \{cx + d \bigm| c, d \in \mathbb{C}\}$$

and then stabilizes $\mathscr{N}(t^2) = \mathscr{N}(t^3) = \cdots$.

This graph illustrates Lemma 1.3. The horizontal axis gives the power j of a transformation. The vertical axis gives the dimension of the rangespace of t^j as the distance above zero — and thus also shows the dimension of the nullspace as the distance below the gray horizontal line, because the two add to the dimension n of the domain.



As sketched, on iteration the rank falls and with it the nullity grows until the two reach a steady state. This state must be reached by the n-th iterate. The steady state's distance above zero is the dimension of the generalized rangespace and its distance below n is the dimension of the generalized nullspace.

1.7 Definition Let t be a transformation on an n-dimensional space. The *generalized rangespace* (or the *closure of the rangespace*) is $\mathscr{R}_\infty(t) = \mathscr{R}(t^n)$ The *generalized nullspace* (or the *closure of the nullspace*) is $\mathscr{N}_\infty(t) = \mathscr{N}(t^n)$.