Formulations and Consequences of Nakayama's Lemma

Andrew Tindall
Final Project
Algebra II

December 13, 2019

In this paper, we will formulate and prove an important lemma in commutative algebra called Nakayama's lemma, named after Tadashi Nakayama, who introduced it in the main form we will see here in 1951. [1]

1 Nakayama's Lemma

We first state the central lemma. In the following, R is an arbitrary ring with identity. The definitions and proofs in the following section are all from Eisenbud, [3], expanded and clarified where it seemed appropriate.

Definition 1.1. The **Jacobson radical** of R is the intersection of all the maximal ideals of R.

Theorem 1.2. Nakayama Let I be an ideal contained in the Jacobson radical of a ring R, and let M be a finitely generated R-module.

a. If $IM = M$, then $M = 0$.

b. If $m_1, \ldots, m_n \in M$ have images in M/IM that generate it as an R-module, then $m_1, \ldots m_n$ generate M as an R-module.

To prove this, we first prove a theorem which is commonly known as the *Cayley-Hamilton Theorem*, although it is stated more generally than the theorem of linear algebra which usually goes by this name.

Theorem 1.3. Let R be a ring, $I \subset R$ an ideal, and M an R-module that can be generated by n elements. Let φ be an endomorphism of M. If

$$\varphi(M) \subset IM,$$

then there is a monic polynomial

$$p(x) = x^n + p_1 x^{n-1} + \cdots + p_n$$

with $p_j \in I$ for each j, such that $p(\varphi) = 0$ as an endomorphism of M.

Proof. Let $m_1, \ldots, m_n$ be a set of generators of M. Any element of IM can be written as a finite sum $\sum_i a_i n_i$, where $a_i \in I$ and $n_i \in M$. Expanding n_i as a sum of the generators m_j of M, we see that any element x of IM can be written in terms of m_j, with coefficients in I:

$$x = \sum a_i n_i = \sum_i a_i (\sum_j b_j m_j) = \sum_j \left(\sum_i (a_i b_j) \right) m_j$$

Specifically, because $\varphi(M) \subset IM$, the image of each generator, $\varphi(m_i)$, can be written in terms of the m_j, using coefficients in I:

$$\varphi(m_i) = \sum_j a_{ij} m_j \tag{1}$$

The coefficients a_{ij} form a matrix A whose entries are all elements of I.

As shown in Dummit & Foote, We can take any R-module to be an $R[x]$-module, by having x act as a particular endomorphism. In this way, we have an $R[x]$-action on our module M, where x acts as φ. Equation 1 then says

$$x \cdot m_i - \sum_j a_{ij} \cdot m_j = 0 \tag{2}$$

Now, let A be the matrix whose entries are the elements a_{ij} of I, viewed as a matrix over $R[x]$, and let $x\mathbf{1}$ be the $n \times n$ diagonal matrix whose diagonal entries are all x. If m is the vector in $M^{\oplus n}$ whose entries are m_j, then equation 2, for $1 \leq i \leq n$, gives the relation

$$(x\mathbf{1} - A) \cdot m = 0.$$

Where $(x\mathbf{1} - A)$ is a matrix of elements of $R[x]$, acting as an endomorphism of the $R[x]$-module $M^{\oplus n}$. We can do linear algebra on this module: in particular, by multiplying $x\mathbf{1} - A$ by its matrix of cofactors, we obtain

$$[\det(x\mathbf{1} - A)]\mathbf{1} \cdot m = 0, \tag{3}$$

where $\det(x\mathbf{1} - A)$ is the formal determinant of the matrix $x\mathbf{1} - A$. As shown in §VI.3 of [1], this determinant is a polynomial in the elements of $(x\mathbf{1} - A)$; in particular, it is a monic element $p(x)$ of $R[x]$. For each basis element m_i, equation 3 gives us

$$p(x) \cdot m_i = 0$$

Since $p(x) \cdot m_i = p(\varphi(m_i))$, and the elements m_i generate M, we see that $p(\varphi)$ takes any element b of M to 0:

$$\begin{aligned} p(\varphi(b)) &= p\left(\sum_i b_i m_i \right) \\ &= \sum_i b_i p(m_i) \\ &= \sum_i 0 = 0 \end{aligned}$$

□

Therefore, the monic polynomial $p(x)$ takes φ to 0, as an endomorphism of M. Also, because the elements of the matrix $(x\mathbf{1} - A)$ are all elements of I, or are $(x - a_{ii})$, where $a_{ii} \in I$, and $p(x)$ is polynomial in the elements of $(x\mathbf{1} - A)$, we can see that the non-leading coefficients of $p(x)$ are all elements of I.

We now use the Cayley-Hamilton theorem to prove the following corollary:

Corollary 1.4. If M is a finitely generated R-module and I is an ideal of R such that $IM = M$, then there is an element $r \in I$ that acts as the identity on M; that is, such that $(1 - r)M = 0$.

Proof. Let φ be the identity map $\mathrm{Id}_M : M \to M$; then $\varphi(M) = IM$. The Cayley-Hamilton theorem above gives us a monic polynomial $p(x)$ such that $p(\mathrm{Id}_M) = 0$. Let the non-leading coefficients of this polynomial be $p_1, \dots p_n$. Since $(\mathrm{Id}_M)^n = \mathrm{Id}_M$, the function $p(\mathrm{Id}_M)$ takes an element $y \in M$ to

$$\begin{aligned} p(\mathrm{Id}_M)(y) &= y + p_1 \cdot y + p_2 \cdot y + \cdots + p_n \cdot y \\ &= (1 + p_1 + \cdots + p_n) \cdot y \end{aligned}$$

Since $p(\mathrm{Id}_M)(y) = 0$ for all $y \in M$, this shows that the element $-(p_1 + \cdots + p_n) \in I$ acts as the identity on M. □

We are now ready to prove Nakayama's lemma. Let I be an ideal contained in the Jacobson radical of R, and let M be a finitely generated R-module.

a. If $IM = M$, then $M = 0$.

   Proof. Apply Corollary 1.4: we have an $r \in I$ such that $(1 - r)M = 0$. Since r is in the Jacobson radical of R, it is in every maximal ideal; therefore, $1 - r$ is not in any maximal ideal. However, every non-unit of a ring is contained in some maximal ideal (by Zorn's lemma). So $1 - r$ must be a unit. So,

$$\begin{aligned} M &= ((1 - r)^{-1}(1 - r))M \\ &= (1 - r)^{-1}((1 - r)M) \\ &= (1 - r)^{-1}0 \\ &= 0 \end{aligned}$$

   We see that M must be 0. □

b. If $m_1, \dots, m_n \in M$ have images in M/IM that generate it as an R-module, then $m_1, \dots, m_n$ generate M as an R-module.

   Proof. Let $N = M/(\sum_i Rm_i)$. We want to show that $\sum_i Rm_i = M$, which is equivalent to saying $N = 0$. □

2 Consequences of the Lemma

Nakayama's Lemma is useful in many contexts. Many of the applications of the lemma involve the special case of a local ring, which gives a particularly nice form of the lemma (the statement, but not proof, of the following lemma is from Wikipedia)

Lemma 2.1. If M is a finitely generated module over a local ring R with maximal ideal m, the quotient M/mM is a vector space over the field R/m, and a basis of M/mM lifts to a minimal set of generators of M. Conversely, every minimal set of generators of M is obtained in this way, and any two sets of generators are related by an invertible matrix with elements in the ring.

Proof. The forward direction of this proof is a simple application of the lemma. Let $m_1, \ldots, m_n$ be elements of M whose images form a basis for M/mM. In the case of a local ring R, the Jacobson radical is just the unique maximal ideal m, so Theorem 1.2 a.) implies that $m_1, \ldots, m_n$ generate M.

We see that they must be a minimal set of generators, because if any proper subset $m_1, \ldots, m_{n-k}$ generated M, then the images of $m_1, \ldots, m_{n-k}$ would generate M/mM. But the images of $m_1, \ldots m_n$ are a basis of the vector space M/mM, so no proper subset of them can span the whole space.

Conversely, let $m_1, \ldots, m_n$ be a minimal set of generators of M. Their images must span the vector space M/mM, and we see that they must be a basis: if the images of any smaller subset $m_1, \ldots, m_{n-k}$ spanned M/mM, then this would lift to a generating set $m_1, \ldots, m_{n-k}$ of M, contradicting the minimality of $m_1, \ldots, m_n$.

Finally, let $a_1, \ldots, a_n$ and $b_1, \ldots, b_m$ be two minimal generating sets of M. First, because they both descend to bases of the vector space M/mM, which has well-defined dimension, the number of generators in each set must be equal: $n = m$. And, each a_i is defined as a sum of b_js:

$$a_i = \sum_j c_{ij} b_j$$

The coefficients c_{ij} form a matrix C, so that $\langle a_1, \ldots, a_n \rangle = C\langle b_1, \ldots, b_n \rangle$. Similarly,

$$b_i = \sum_j d_{ij} a_j$$

with the entries d_{ij} forming a matrix D, so that $\langle b_1, \ldots, b_n \rangle = D\langle a_1, \ldots, a_n \rangle$. It is immediate that $CDa_i = a_i$ and $DCb_i = b_i$ for any of the generators from either set. Then, for any element $x \in M$, we see that $DCm = CDm = m$, so that $DC = CD = \mathrm{Id}$, and the two

matrices are inverses:

$$\begin{aligned} CDx &= CD\left(\sum_i x_i a_i\right) \\ &= C\left(\sum_i x_i Da_i\right) \\ &= \sum_i x_i (CDa_i) \\ &= \sum_i x_i a_i \\ &= x \end{aligned}$$

The proof in the opposite direction is symmetric. Thus, any two minimal generating sets of M are lifts of bases of M/mM, and are related by an invertible matrix. □

This local case of Nakayama's lemma is useful in algebraic geometry, since it essentially reduces many questions about generators of local rings to questions about bases of vector spaces.

Another useful application of the lemma is in proving the following criterion for morphisms of finitely generated modules, which is similar to the case of finite dimensional vector spaces: On a finite dimensional vector space, the fact that a surjective endomorphism is injective follows from some considerations about bases and dimension; Nakayama's lemma allows us to generalize this to the case of a commutative ring.

Lemma 2.2. Let R be a commutative ring, M a finitely generated R module, and let φ be a surjective endormorphism of M. Then φ is also injective, and is therefore an automorphism.

Proof. The following proof is from [5]: we wish to show that $f(u) = 0$ implies $u = 0$. As in the proof of the Cayley-Hamilton Theorem, we can view M as an $R[x]$ module by letting x act on M by φ. Then $xM = \varphi(M) = M$, and by Corrolary 1.4, there is some element $r \in (x)$ such that $(1 - r)M = 0$. Writing r as $p(x)x$ for some polynomial $p \in R[x]$, we have $(1 - p(x)x)M = 0$.

Let $u \in \mathrm{Ker} f$. Then

$$\begin{aligned} 0 &= (1 - p(x)x) \cdot u \\ &= u - p(x) \cdot \varphi(u) \\ &= u - p(x) \cdot 0 \\ &= u \end{aligned}$$

So, $u = 0$, meaning the kernel of f must be trivial. Therefore, we see that every surjective endomorphism of a finite R-module is an isomorphism. □

Over a local ring, due to the connection with vector spaces via Lemma 2.1, we have the following:

Lemma 2.3. Let R be a local ring with maximal ideal m, and let M, N be finitely generated R-modules. If $\varphi : M \to N$ is an R-linear map such that $\varphi_m : M/mM \to N/mN$ is surjective, then φ is surjective.

Proof. Let $n_1, \ldots, n_k$ be elements of N which descend to a basis of N/mN. By Lemma 2.1, $n_1, \ldots, n_k$ give a minimal generating set of N. Because φ_m is surjective, we have

$$\varphi_m(\overline{m_i}) = \overline{n_i}$$

for some elements $m_1, \ldots, m_k$ of M. Then $\varphi(m_i) = n_i'$, for some element $n_i' \in N$ over $\overline{n_i}$. Because the n_i' lie over a basis of N/nN, they form a generating set of N, and φ is surjective. □

Finally, assuming some more knowledge of local ring theory, we have the following theorem (from [4]:

Lemma 2.4. Let $f : A \to B$ be a local homomorphism of local noetherian rings, such that

a. $A/m_A \to B/m_B$ is an isomorphism

b. $m_a \to m_B/m_B^2$ is surjective, and

c. B is a finitely generated A-module,

then f is surjective.

Here,

- a local homomorphism of local rings is a homomorphism φ from a local ring A to a local ring B, such that the image of the maximal ideal of A lies in the maximal ideal of B: $\varphi(m_A) \subset m_B$
- the map
$$A/m_A \to B/m_B$$
is defined by first taking the composition $A \xrightarrow{f} B \xrightarrow{\pi} B/m_B$, and then descending to a homomorphism on A/m_A, by the fact that the kernel of $\pi \circ f$ contains m_A
- the map $m_A \to m_B/m_B^2$ is defined by restricting f to m_A, which gives a function $m_A \to m_B$, and composing with the quotient $m_B \to m_B/m_B^2$.
- B is considered as an A-module through the action of f.

Proof. Consider the ideal $\mathfrak{a} = m_A B$ of B. Because $m_A B = f(m_A)B \subset m_B B = m_B$, and by a. we have that $\mathfrak{a}$ contains a set of generators for m_B/m_B^2. By Nakayama's lemma for the B-module B, we can lift these generators to a set of generators for m_B, so that $\mathfrak{a} = m_B$.

Now, by c.), we have that B is a finitely generated A-module. By a., A/m_A is isomorphic to $B/m_B = B/m_A B$, so the element 1 generates $B/m_A B$ as an A-module. By Nakayama's lemma again, this time for B as an A-module, we can lift 1 to a generator of B as an A-module. This implies that f is surjective, as if 1 generates B, then any element $b \in B$ is equal to $a \cdot 1 = f(a) \cdot 1 = f(a)$, for some $a \in A$. □

The use of Nakayama's lemma for local rings gives many more useful results in local ring theory and algebraic goemetry.

Notes

[1]Other important figures in the development of this theory have been Wolfgang Krull, who discovered it in the specific case of ideals over a commutative ring; Goro Azumaya, who formulated it in a more general form, possibly before Nakayama, and Nathan Jacobson, who formulated the noncommutative version of the lemma in 1945. This same lemma has been called the "Krull-Azumaya Lemma," even just "Azumaya's Lemma," and the noncommutative form is often called the "Jacobson-Azumaya Lemma."

I found the general history of this lemma, and notes about its various names, in Appendix $A.2$ of Nagata's 1962 book on Local Rings, [6].

References

[1] Aluffi, Paulo. Algebra, Chapter 0. Graduate Studies in Mathematics, Vol 104, AMS, 2009

[2] Dummit, David, & Foote, R. Abstract Algebra, 3rd Edition. Wiley, 2004

[3] Eisenbud, David. Commutative Algebra with a View Toward Algebraic Geometry. Graduate Texts in Mathematics Vol 150, Springer, 1995

[4] Hartshorne, Robin. Algebraic Geometry. Graduate Texts in Mathematics, Vol 52. Springer,1977

[5] Matsumura, Hideyuki. Commutative Ring Theory. Cambridge Studies in Advanced Mathematics, Vol 8. Cambridge, 1989

[6] Nagata, Masayoshi. Local Rings. Wiley, 1962