Math 346 Lecture #33 12.3 The Resolvent

The Jordan Canonical Form, or spectral decomposition, of a linear operator on a finite dimension vector space has important applications in many areas such as differential equations and dynamical systems. Computing the Jordan Canonical form depends on finding bases for the generalized eigenspaces, and as mentioned before, this is a poorly conditioned numerical algorithm.

A more powerful approach to finding a spectral decomposition of a linear operator on a finite dimensional vector space uses the tools of complex analysis. This theoretical approach is basis-free, meaning we do not have to find bases of the generalized eigenspaces to get a spectral decomposition.

12.3.1 Properties of the Resolvent

Definition 12.3.1. The resolvent set of $A \in M_n(\mathbb{C})$, denoted by $\rho(A)$, is the set of points $z \in \mathbb{C}$ for which zI - A is invertible.

Note that the complement $\sigma(A) = \mathbb{C} \setminus \rho(A)$ is the spectrum of A.

The resolvent of A is the map $R(A, \cdot) : \rho(A) \to M_n(\mathbb{C})$ defined by

$$R(A, z) = (zI - A)^{-1}.$$

We will sometimes make use of the notation $R_A(z)$ for R(A, z) to make clear the dependence of the resolvent on A.

When there is no ambiguity, we denote R(A, z) simply as R(z).

Example (in lieu of 12.3.2). The resolvent R_A for

$$A = \begin{bmatrix} 1 & 1\\ 4 & 1 \end{bmatrix} \in M_2(\mathbb{C})$$

is

$$R_A(z) = \begin{bmatrix} z - 1 & -1 \\ -4 & z - 1 \end{bmatrix}^{-1}$$
$$= \frac{1}{(z - 1)^2 - 4} \begin{bmatrix} z - 1 & 1 \\ 4 & z - 1 \end{bmatrix}$$
$$= \frac{1}{z^2 - 2z - 3} \begin{bmatrix} z - 1 & 1 \\ 4 & z - 1 \end{bmatrix}$$
$$= \frac{1}{(z - 3)(z + 1)} \begin{bmatrix} z - 1 & 1 \\ 4 & z - 1 \end{bmatrix}$$

The resolvent has simple poles precisely on $\sigma(A) = \{3, -1\}$. The domain of the resolvent R_A is the resolvent set $\rho(A) = \mathbb{C} \setminus \sigma(A)$.

The domain of the resolvent n_A is the resolvent set $p(A) = C \setminus O(A)$.

Remark 12.3.3. Each entry of the resolvent is a rational function by Cramer's Rule:

$$R_A(z) = (zI - A)^{-1} = \frac{1}{\det(zI - A)} \operatorname{adj}(zI - A)$$

where $\det(zI - A)$ is the characteristic polynomial of A, and $\operatorname{adj}(zI - A)$ is the adjugate matrix, i.e., the transpose of the matrix of signed minors of zI - A, which satisfies $(zI - A)\operatorname{adj}(zI - A) = \det(zI - A)I$ (see Definition 2.9.19 and Theorem 2.9.22).

The rational function nature of the entries of the resolvent shows that the resolvent has poles, some possibly not simple, precisely on $\sigma(A)$.

Example (in lieu of 12.3.4). (i) Find the resolvent for

$$A = \begin{bmatrix} 6 & 1 & 0 \\ 0 & 6 & 7 \\ 0 & 0 & 4 \end{bmatrix}.$$

Because A is upper triangular, the characteristic polynomial of A is

$$\det(zI - A) = (z - 6)^2(z - 4).$$

The adjugate of A is

adj
$$(zI - A) = \begin{bmatrix} (z - 6)(z - 4) & z - 4 & 7\\ 0 & (z - 6)(z - 4) & 7(z - 6)\\ 0 & 0 & (z - 6)^2 \end{bmatrix}$$
.

We verify this by computing (zI - A)adj(zI - A) = det(zI - A)I. \checkmark The resolvent of A is

$$R(z) = \begin{bmatrix} (z-6)^{-1} & (z-6)^{-2} & 7(z-6)^{-2}(z-4)^{-1} \\ 0 & (z-6)^{-1} & 7(z-6)^{-1}(z-4)^{-1} \\ 0 & 0 & (z-4)^{-1} \end{bmatrix}.$$

(ii) Find the resolvent of

$$A = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 5 \end{bmatrix}$$

Because A is upper triangular, the characteristic polynomial of A is

$$\det(zI - A) = (z - 2)^3(z - 5).$$

The required adjugate of A is

$$\operatorname{adj}(zI - A) = \begin{bmatrix} (z - 2)^2(z - 5) & (z - 2)(z - 5) & z - 5 & 3\\ 0 & (z - 2)^2(z - 5) & (z - 2)(z - 5) & 3(z - 2)\\ 0 & 0 & (z - 2)^2(z - 5) & 3(z - 2)^2\\ 0 & 0 & 0 & (z - 2)^3 \end{bmatrix}.$$

The resolvent has several basic properties as expressed next. Lemma 12.3.5. Let $A, A_1, A_2 \in M_n(\mathbb{C})$. (i) If $z_1, z_2 \in \rho(A)$, then

$$R(z_2) - R(z_1) = (z_1 - z_2)R(z_2)R(z_1).$$

[This is known as Hilbert's Identity.]

(ii) If $z \in \rho(A_1) \cap \rho(A_2)$, then

$$R(A_2, z) - R(A_1, z) = R(A_1, z)(A_2 - A_1)R(A_1, z).$$

(iii) If $z \in \rho(A)$, then

$$R(z)A = AR(z)$$

(iv) If $z_1, z_2 \in \rho(A)$, then

$$R(z_1)R(z_2) = R(z_2)R(z_1).$$

Proof. (i) Multiplying the identity

$$(z_1 - z_2)I = (z_1I - A) - (z_2I - A)$$

on the left by $R(z_2)$ and on the right by $R(z_1)$ gives

$$R(z_2)(z_1 - z_2)R(z_1) = R(z_2)(z_1I - A)R(z_1) - R(z_2)(z_2I - A)R(z_1).$$

Since $R(z_1) = (z_1I - A)^{-1}$ and $R(z_2) = (z_2I - A)^{-1}$ the equation simplifies to

$$R(z_1)(z_1 - z_2)R(z_2) = R(z_2) - R(z_1).$$

The term $z_1 - z_2$ is a complex scalar and so commutes with the matrix $R(z_2)$ to give

$$(z_1 - z_2)R(z_2)R(z_1) = R(z_2) - R(z_1).$$

(ii) Multiply the identity

$$A_2 - A_1 = (zI - A_1) - (zI - A_2)$$

on the left by $R(A_1, z)$ and on the right by $R(A_2, z)$ to obtain

$$R(A_1, z)(A_2 - A_1)R(A_2, z) = R(A_1, z)(zI - A_1)R(A_2, z) - R(A_1, z)(zI - A_2)R(A_2, z).$$

Since $R(A_1, z) = (zI - A_1)^{-1}$ and $R(A_2, z) = (zI - A_2)^{-1}$ the equation simplifies to

$$R(A_1, z)(A_2 - A_1)R(A_2, z) = R(A_2, z) - R(A_1, z).$$

(iii) Since R(z)(zI - A) = I and (zI - A)R(z) = I, we have

$$R(z)(zI - A) = (zI - A)R(z).$$

Multiplying this out gives

$$zR(z) - R(z)A = zR(z) - AR(z).$$

Cancelling the common term zR(z) gives

$$R(z)A = AR(z)$$

(iv) When $z_1 = z_2$ it follows trivially that $R(z_1)R(z_2) = R(z_2)R(z_1)$.

When $z_1 \neq z_2$, then Hilbert's identity gives

$$R(z_2)R(z_1) = \frac{R(z_2) - R(z_1)}{z_1 - z_2} = \frac{R(z_1) - R(z_2)}{z_2 - z_1} = R(z_1)R(z_2).$$

This gives the desired result.

12.3.2 Local Properties

We show that the resolvent R_A is a matrix-valued holomorphic function on $\rho(A)$ by finding power series expansions of R_A at all points $z \in \rho(A)$. The original lower bounds on the radii of convergence for these power series are enlarged by means of a limit quantity called the spectral radius of A.

Throughout this subsection we assume that $\|\cdot\|$ is a matrix norm on $M_n(\mathbb{C})$, i.e., a norm on $M_n(\mathbb{C})$ that for all $A, B \in M_n(\mathbb{C})$ satisfies

$$||AB|| \le ||A|| \, ||B||.$$

Each induced p-norm

$$||A||_p = \sup\left\{\frac{||A\mathbf{x}||_p}{||\mathbf{x}||_p} : \mathbf{x} \in \mathbb{C}^n \setminus \{0\}\right\}$$

is a matrix norm, as is the Frobenius norm $\|\cdot\|_F$.

Theorem 12.3.6. For $A \in M_n(\mathbb{C})$, the resolvent set $\rho(A)$ is open, and R is holomorphic on $\rho(A)$ where for each $z_0 \in \rho(A)$, the resolvent is given by the power series

$$R(z) = \sum_{k=0}^{\infty} (-1)^k (z - z_0)^k R^{k+1}(z_0),$$

whose the radius of convergence is at least as large as $||R(z_0)||^{-1}$.

Proof. For $B \in M_n(\mathbb{C})$, if ||B|| < 1, then by Proposition 5.7.4, I - B is invertible and there holds the Neumann series

$$\sum_{k=0}^{\infty} B^k = (I - B)^{-1}.$$

For $A \in M_n(\mathbb{C})$, using Hilbert's Identity for points $z_0, z \in \rho(A)$, the resolvent R satisfies

$$R(z_0) = R(z) + (z - z_0)R(z_0)R(z) = [I + (z - z_0)R(z_0)]R(z).$$

Setting $B = -(z - z_0)R(z_0)$ the condition ||B|| < 1 implies that

$$|z - z_0| ||R(z_0)|| = ||B|| < 1.$$

Since $z_0 \in \rho(A)$, the matrix $R(z_0)$ is invertible so that $||R(z_0)|| > 0$. Thus for z satisfying $|z - z_0| \le ||R(z_0)||^{-1}$, the matrix

$$I - B = I + (z - z_0)R(z_0)$$

is invertible and

$$[I + (z - z_0)R(z_0)]^{-1} = \sum_{k=0}^{\infty} (-1)^k (z - z_0)^k R^k(z_0)$$

Since $R(z_0) = R(z) + (z - z_0)R(z_0)R(z) = [I + (z - z_0)R(z_0)]R(z)$ we obtain

$$R(z) = [I - (z - z_0)R(z_0)]^{-1}R(z_0) = \sum_{k=0}^{\infty} (-1)^k (z - z_0)^k R^{k+1}(z_0),$$

which power series converges on the open disk $D(z_0) = |z - z_0| \le ||R(z_0)||^{-1}$. The union $\bigcup_{z_0 \in \rho(A)} D(z_0)$ is the same as $\rho(A)$, implying that $\rho(A)$ is open. Therefore, by Remark 11.2.10, the resolvent R is holomorphic on the open $\rho(A)$. Remark 12.3.7. Comparing the power series

$$R(z) = \sum_{k=0}^{\infty} (-1)^k (z - z_0)^k R^{k+1}(z_0)$$

in Theorem 12.3.6 and the Taylor series

$$R(z) = \sum_{k=0}^{\infty} \frac{R^{(k)}(z_0)}{k!} (z - z_0)^k$$

reveals, by the uniqueness of the Taylor series, a relationship between the powers of R and its derivatives, namely that

$$(-1)^k R^{k+1}(z_0) = \frac{R^{(k)}(z_0)}{k!}.$$

As this holds for every $z_0 \in \rho(A)$, we may replace z_0 by z to get

$$R^{(k)}(z) = k!(-1)^k R^{k+1}(z) = k!(-1)^k (zI - A)^{-(k+1)}.$$

Theorem 12.3.8. For $A \in M_n(\mathbb{C})$, the Laurent series of R(z) on the open annulus |z| > ||A|| exists and is given by

$$R(z) = \sum_{k=0}^{\infty} \frac{A^k}{z^{k+1}}.$$

Proof. The resolvent is

$$R(z) = (zI - A)^{-1} = (z(I - z^{-1}A))^{-1} = z^{-1}(I - z^{-1}A)^{-1}.$$

To express $(I - z^{-1}A)^{-1}$ as a Neumann series requires that $||z^{-1}A|| < 1$. This condition gives the annulus |z| > ||A||, on which we have

$$R(z) = z^{-1}(I - z^{-1}A)^{-1} = z^{-1}\sum_{k=0}^{\infty} \frac{A^k}{z^k} = \sum_{k=0}^{\infty} \frac{A^k}{z^{k+1}} = \frac{I}{z} + \frac{A}{z^2} + \frac{A^2}{z^3} + \cdots$$

This is the Laurent series of R(z) on the open annulus |z| > ||A||.

Remark. One might be tempted to say that the resolvent R(z) has an essential singularity at z = 0, but the open annulus |z| > ||A|| is not a punctured disk unless A = 0 in which case the resolvent has a simple pole at 0. But we can use the Laurent series of R(z) on |z| > ||A|| to say something about the behavior of R(z) as $|z| \to \infty$. For this we speak of R(z) being holomorphic in a neighbourhood of ∞ , which means that the function $(R \circ g)(w)$, for the change of variables z = g(w) = 1/w, is holomorphic on an open disk centered at w = 0. Here |z| = 1/|w| so that $|w| \to 0$ if and only if $|z| \to \infty$.

Corollary 12.3.9. For any $A \in M_n(\mathbb{C})$, the resolvent R is holomorphic in a neighbourhood of ∞ , and moreover there holds

$$\lim_{|z|\to\infty} \|R(z)\| = 0.$$

Proof. From Theorem 12.3.8, the Laurent series

$$R(z) = \sum_{k=0}^{\infty} \frac{A^k}{z^{k+1}} = \sum_{k=0}^{\infty} A^k \left(\frac{1}{z}\right)^{(k+1)}$$

converges on |z| > ||A||.

If A = 0, then $R(z) = z^{-1}I$ and under the change of variables z = g(w) = 1/w we obtain $(R \circ g)(w) = wI$ which is entire.

Now suppose $A \neq 0$. Then ||A|| > 0.

Under the change of variables z = g(w) = 1/w, the annulus |z| > ||A|| becomes 1/|w| > ||A|| which is the disk |w| < 1/||A||.

With the change of variables z = g(w) = 1/w we obtain the power series

$$(R \circ g)(w) = \sum_{k=0}^{\infty} A^k w^{k+1}$$

that converges on the disk |w| < 1/||A||.

Now from the Laurent series for the resolvent and $|z| > \|A\|$ we have that $\|A\|/|z| < 1$ and that

$$\|R(z)\| \le \sum_{k=0}^{\infty} \frac{\|A\|^k}{|z|^{k+1}} = \frac{1}{|z|} \sum_{k=0}^{\infty} \left(\frac{\|A\|}{|z|}\right)^k = \frac{1}{|z|} \left(1 - \frac{\|A\|}{|z|}\right)^{-1} = \frac{1}{|z| - \|A\|}.$$

This shows that $\lim_{|z|\to\infty} ||R(z)|| = 0.$

Remark 12.3.11. We saw in Remark 12.3.3 that the values of z for which (zI - A) is not invertible are precisely the eigenvalues of A, i.e., the roots of the characteristic polynomial det(zI - A). The Fundamental Theorem of Algebra implies that the characteristic polynomial has n roots (counting repeated eigenvalues if any). This means that the resolvent set $\rho(A)$ is not all of \mathbb{C} , that $\sigma(A) = \mathbb{C} \setminus \rho(A) \neq \emptyset$. Next is another proof that the spectrum of A is not empty.

Corollary 12.3.12. For any $A \in M_n(\mathbb{C})$, the spectrum $\sigma(A)$ is not empty.

Proof. Suppose to the contrary that $\sigma(A) = \emptyset$.

Then $\rho(A) = \mathbb{C}$, and hence that the resolvent R is entire.

Corollary 12.3.9 implies that ||R|| is a bounded entire function.

Liouville's Theorem implies that R is a constant function.

Since $\lim_{|z|\to\infty} ||R(z)|| = 0$, the resolvent R is the zero function.

But this is a contradiction because the resolvent satisfies I = (zI - A)R(z).

Note. The radius of convergence of the power series of R(z) about $z_0 \in \rho(A)$ is at least as large as $||R(z_0)||^{-1}$ by Theorem 12.3.6. The inner radius of the annulus for the Laurent series of R(z) about 0 is at least as small as ||A||. We show that these radii can be improved through a limit quantity known as the spectral radius.

Lemma 12.3.13. For any $A \in M_n(\mathbb{C})$, the limit

$$r(A) = \lim_{k \to \infty} \|A^k\|^{1/k}$$

exists and is bounded above by ||A||.

Proof. If there exists $L \in \mathbb{N}$ such that $||A^L|| = 0$, then $A^L = 0$ and $||A^l|| = 0$ for all $l \ge L$, and hence $||A^l|| = 0$ for all $l \ge L$, implying that r(A) = 0.

Suppose that $||A^l|| > 0$ for all $l \in \mathbb{N}$.

Using the submultiplicative property of the given norm, we have for integers $1 \le s < t$ that

$$||A^{s}|| = ||A^{t-s}A^{s}|| \le ||A^{t-s}|| \, ||A^{s}||$$

Since none of the norms are zero, we can apply the logarithm function to the inequality to get

$$\log \|A^t\| \le \log \|A^{t-s}\| + \log \|A^s\|.$$

Setting $a_l = \log ||A^l||$ for each $l \in \mathbb{N}$ gives whenever $1 \leq s < t$ that

$$a_t \le a_{t-s} + a_s.$$

Now fix integers k and m satisfying $1 \le m < k$.

Since $k > m \ge 1$, there exists by the Division Algorithm unique integers q and p where $q \ge 1$ and $0 \le p < m$ such that

$$k = mq + p.$$

We can express $q = \lfloor k/m \rfloor$, the greatest integer equal to or less then k/m = q + p/m.

Setting t = k and s = p in $a_t \le a_{t-s} + a_s$ gives $a_k \le a_{k-p} + a_p = a_{mq} + a_p$. Setting t = mq and s = m(q-1) in $a_t \le a_{t-s} + a_s$ gives $a_{mq} \le a_m + a_{m(q-1)}$. Setting t = m(q-1) and s = m(q-2) in $a_t \le a_{t-s} + a_s$ gives $a_{m(q-1)} \le a_m + a_{m(q-2)}$. Continuing this gives $a_{mq} \le qa_m$. Thus for $1 \le m < k$ we obtain

$$a_k = a_{mq+p} \le qa_m + a_p$$

This implies that

$$\frac{a_k}{k} \le \frac{q}{k}a_m + \frac{1}{k}a_p$$

Note that

$$\exp\left(\frac{a_k}{k}\right) = \exp\left(\frac{\log\|A^k\|}{k}\right) = \exp\left(\|A^k\|^{1/k}\right) = \|A^k\|^{1/k}$$

Leaving *m* fixed (for now) and letting k > m be arbitrary we have that $\lfloor k/m \rfloor = q$ for all $k = mq, mq + 1, \ldots, mq + (m-1)$, i.e., $\lfloor (mq+p)/m \rfloor = q$ for all $p = 0, 1, \ldots, m-1$. Hence

$$\frac{q}{k} = \frac{q}{mq+p} \le \frac{q}{mq} = \frac{1}{m}$$

with equality realized when p = 0.

This implies that

$$\limsup_{k \to \infty} \frac{q}{k} = \frac{1}{m}$$

We then have

$$\limsup_{k \to \infty} \frac{a_k}{k} \le \limsup_{k \to \infty} \left(\frac{q}{k} a_m + \frac{1}{k} a_p \right) = \frac{a_m}{m}.$$

This holds for each m, so that

$$\limsup_{k \to \infty} \frac{a_k}{k} \le \liminf_{m \to \infty} \frac{a_m}{m}.$$

This implies that the limit of $(1/k)a_k$ and hence that the limit of $\exp((1/k)a_k) = ||A^k||^{1/k}$ exists by the continuity of the exponential function.

To get the upper bound of ||A|| on $\rho(A)$, we fix m = 1 to get

$$\log r(A) = \lim_{k \to \infty} \frac{\log \|A^k\|}{k} = \limsup_{k \to \infty} \frac{a_k}{k} \le \frac{a_1}{1} = \log \|A\|.$$

Exponentiating the inequality gives $r(A) \leq ||A||$.

Remark. We will see in the next section that the quantity r(A) is independent of the matrix norm used in the limit.

Definition. The spectral radius of $A \in M_n(\mathbb{C})$ is defined to be the quantity r(A).

Theorem 12.3.14. For $A \in M_n(\mathbb{C})$, the power series

$$R(z) = \sum_{k=0}^{\infty} (-1)^k (z - z_0)^k R^{k+1}(z_0)$$

converges on $|z - z_0| < [r(R(z_0))]^{-1}$, and the Laurent series

$$R(z) = \sum_{k=0}^{\infty} \frac{A^k}{z^{k+1}}$$

converges on the annulus |z| > r(A).

Proof. The power series

$$R(z) = \sum_{k=0}^{\infty} (-1)^k (z - z_0)^k R^{k+1}(z_0) = R(z_0) \sum_{k=0}^{\infty} (-1)^k (z - z_0)^k R^k(z_0).$$

converges if there exists $N \in \mathbb{N}$ such that for all $k \geq N$ there holds

$$\left\|\left((z-z_0)R(z_0)\right)^k\right\| < 1.$$

To simplify notation set $r = r(R(z_0))$.

Let $z \in \mathbb{C}$ satisfy $|z - z_0| < r^{-1}$.

Then there exists $\epsilon > 0$ such that $|z - z_0| < (r + 2\epsilon)^{-1}$.

For this same ϵ , since $||R(z_0)^k||^{1/k} \to r$, there exists $N \in \mathbb{N}$ such that $||R(z_0)^k||^{1/k} < r + \epsilon$, or $||R(z_0)^k|| \le (r+\epsilon)^k$, for all $k \ge N$.

Thus $|z - z_0|^k < (r + 2\epsilon)^{-k}$ for all $k \ge N$, so that

$$\left\| \left((z-z_0)R(z_0) \right)^k \right\| = |z-z_0|^k \|R(z_0)^k\| < \left(\frac{r+\epsilon}{r+2\epsilon}\right)^k < 1.$$

Thus the power series converges on $|z - z_0| < r^{-1}$. The Laurent series

$$R(z) = \sum_{k=0}^{\infty} \frac{A^k}{z^{k+1}} = \frac{1}{z} \sum_{k=0}^{\infty} \left(\frac{A}{z}\right)^k$$

converges if there exists $N \in \mathbb{N}$ such that for all $k \ge N$ there holds

$$\left\| \left(\frac{A}{z}\right)^k \right\| < 1.$$

To simplify notation set r = r(A), and let $z \in \mathbb{C}$ satisfy |z| > r. There exists $\epsilon > 0$ such that $|z| > r + 2\epsilon$. Hence $|z|^k > (r+2\epsilon)^k$ for all k which implies that

$$\frac{1}{|z|^k} < \frac{1}{(r+2\epsilon)^k} \text{ for all } k.$$

For this same ϵ , since $||A^k||^{1/k} \to r$, there exists $N \in \mathbb{N}$ such that for all $k \ge N$ there holds $||A^k||^{1/k} < r + \epsilon$, or $||A^k|| < (r + \epsilon)^k$.

Thus for all $k \ge N$ there holds

$$\left\| \left(\frac{A}{z}\right)^k \right\| < \frac{\|A^k\|}{|z|^k} < \left(\frac{r+\epsilon}{r+2\epsilon}\right)^k < 1.$$

Therefore the Laurent series converges on |z| > r.

Remark. The radii of convergence given in Theorem 12.3.14 do improve the radii $||R(z_0)||^{-1}$ and ||A|| given in Theorem 12.3.6 and Theorem 12.3.8. For the first radii this is because $r(R(z_0)) \leq ||R(z_0)||$ implies

$$||R(z_0)||^{-1} \le [r(R(z_0))]^{-1},$$

and for the second radii this is because $r(A) \leq ||A||$, i.e., a smaller inner radius of the open annulus is possible.

Remark 12.3.15. A lower bound on the spectral radius r(A) is given by quantity

$$\sigma_M = \sup\{|\lambda| : \lambda \in \sigma(A)\}.$$

Justification of this lower bound follows from Theorem 12.3.14 because the open annulus on which the Laurent series converges cannot include any point in $\sigma(A)$ where the resolvent has poles. In the next section we will show that $r(A) = \sigma_M$ which implies that the spectral radius is independent of the matrix norm used to compute r(A).