Bounds

In this lecture, we study various bounds on algebraic numbers. Broadly speaking, we will put them into the following categories: bounds on absolute value, root separation bounds, perturbation bounds, and evaluation bounds.

Throughout this lecture, $A(x) = \sum_{i=0}^{n} a_i x^i$ is a degree n polynomial with $a_i \in \mathbb{C}$, $a_n \neq 0$, unless mentioned otherwise. Furthermore, let $\alpha_1, \dots, \alpha_n$ be the n roots of $A(x)$ counted with multiplicities; we use α to denote an arbitrary root of $A(x)$. If $A(x) \in \mathbb{Z}[x]$ then we assume that $\|A\|_\infty \leq 2^L$, i.e., the coefficients are L-bit integers.

1 Bounds on Absolute Value

In this section, we derive upper and lower bounds on $|\alpha|$. In fact, we will only focus on upper bounds since if B is an upper bound on the roots of $x^n A(1/x)$, then we have $B \leq 1/|\alpha|$ which implies that $|\alpha| \geq 1/B$. Thus any function that computes an upper bound can be applied to the reciprocal polynomial $x^n A(1/x)$ to get a lower bound on the absolute value; of course, we have to assume that $a_0 \neq 0$ to be able to do this trick.

Theorem 1 (Upper Bounds [1]). *For any root α of $A(x)$*

1. $|\alpha| \leq \max\left\{1, \sum_{i=0}^{n-1} |a_i|/|a_n|\right\}$.
2. $|\alpha| \leq 1 + (\|A\|_\infty/|a_n|)$. *In particular, if $A \in \mathbb{Z}[x]$ then $|\alpha| \leq 2\|A\|_\infty$.*
3. *If $\lambda_1, \dots, \lambda_n$ are positive real numbers satisfying the condition $\sum_{i=1}^{n} 1/\lambda_i = 1$ then we have*

$$|\alpha| \leq \max_{i=1,\dots,n} \left\{(\lambda_i |a_i|/|a_n|)^{1/(n-i)}\right\}.$$

Proof. Since α is a root, we get the fundamental equation is that

$$|a_n||\alpha|^n = |\sum_{i=0}^{n-1} a_i \alpha^i|. \tag{1}$$

1. If $|\alpha| \leq 1$ then the bound trivially holds, so suppose $|\alpha| > 1$. Then from (1) and triangular inequality it follows that

$$|\alpha|^n \leq \frac{1}{|a_n|} \sum_{i=0}^{n-1} |a_i||\alpha|^i \leq \frac{|\alpha|^{n-1}}{|a_n|} \sum_{i=0}^{n-1} |a_i|$$

which implies that

$$|\alpha| \leq \frac{1}{|a_n|} \sum_{i=0}^{n-1} |a_i|$$

2. We can improve the bound in (i) by instead taking the geometric summation. Again suppose $|\alpha| > 1$.

Then we have from (1) and triangula inequality that

$$
\begin{aligned}
|\alpha|^n &\le \frac{1}{|a_n|}\sum_{i=0}^{n-1}|a_i||\alpha|^i \\
&\le \frac{\max\{|a_0|,|a_1|,\ldots,|a_{n-1}|\}}{|a_n|}\sum_{i=0}^{n-1}|\alpha|^i \\
&\le \frac{\max\{|a_0|,|a_1|,\ldots,|a_{n-1}|\}}{|a_n|}\frac{|\alpha|^n-1}{|\alpha|-1} \\
&\le \frac{\max\{|a_0|,|a_1|,\ldots,|a_{n-1}|\}}{|a_n|}\frac{|\alpha|^n}{|\alpha|-1}.
\end{aligned}
$$

Canceling $|\alpha|^n$ on both sides, we obtain

$$
|\alpha| \le 1+\frac{\max\{|a_0|,|a_1|,\ldots,|a_{n-1}|\}}{|a_n|} \le 1+\frac{\|A\|_\infty}{|a_n|}.
$$

3. To show this general bound, we first start with the following observation: the positive real root of the polynomial

$$
|a_n|x^n-\sum_{i=0}^{n-1}|a_i|x^i=0
$$

is an upper bound on the absolute value of the roots of the polynomial. The root is unique because the polynomial is monotone, and it exists because there is a change of sign from 0 to ∞. Thus any ρ such that

$$
|a_n|\rho^n \ge \sum_{i=0}^{n-1}|a_i|\rho^i
$$

is an upper bound on the absolute value of the roots of the polynomial. In particular, let $\rho := \max_{i=1,\ldots,n}\left\{(\lambda_i|a_{n-i}|/|a_n|)\right.$ Then we have

$$
|a_i| \le |a_n|\rho^{n-i}/\lambda_i.
$$

Thus

$$
\begin{aligned}
\sum_{i=0}^{n-1}|a_i|\rho^i &\le |a_n|\sum_{i=0}^{n-1}\rho^{n-i}\rho^i/\lambda_i \\
&\le |a_n|\rho^n.
\end{aligned}
$$

Q.E.D.

1.1 Cauchy's Bound – Careful Perspective

A common approach to all the bounds above is the following property: Consider the polynomial

$$
Cf(x) := |a_n|x^n-\sum_{i<n}|a_i|x^i. \tag{2}
$$

This polynomial has exactly one positive root $\rho(f)$. This can be shown, e.g., by showing that the polynomial is monotone after the first positive root; one way to show this is that any root ν of the derivative is smaller than $\rho(f)$: consider

$$
n\,Cf(\nu) = n|a_n|\nu^n - n\sum_{i<n}|a_i|x^i < n|a_n|\nu^n-\sum_{i<n}i|a_i|x^i = (Cf)'(\nu)=0.
$$

Our claim is that $\rho := \rho(f)$ is an upper bound on the absolute value of the roots of $f(x)$. If α is any root of $f(x)$ then we know that

$$a_n \alpha^n = -\sum_{i<n} a_i \alpha^i$$

and hence taking the absolute values on both sides and applying the triangular inequality we obtain that

$$|a_n||\alpha|^n - \sum_{i<n} |a_i||\alpha^i| \leq 0,$$

i.e., $Cf(|\alpha|) \leq 0$. Therefore, $|\alpha| \leq \rho$, for all roots α of f, and hence to obtain an upper bound on the absolute value of the roots, we can derive a tight estimate on ρ, e.g., applying Newton-iteration to the polynomial $Cf(x)$ from some suitable starting point. But how good is ρ an estimate to the largest absolute value amongst the roots of f? The next result shows that it is quite good. Let $r := \max_{\alpha : f(\alpha)=0} |\alpha|$. Then from Viete's formula we know that

$$\frac{a_i}{a_n} = \sum_{1 \leq i_1 < i_2 < \cdots < i_{n-i} \leq n} \alpha_{i_1} \ldots \alpha_{i_{n-i}}.$$

Applying triangular inequality on both sides and using the fact that the absolute value of all the roots is smaller than r we obtain that

$$\left| \frac{a_i}{a_n} \right| \leq \binom{n}{i} r^{n-i} \tag{3}$$

which implies that $|a_i| r^i \leq |a_n| r^n \binom{n}{i}$. Since ρ is a root of $Cf(x)$ we know that

$$|a_n|\rho^n = \sum_{i<n} |a_i|\rho^i \leq \sum_{i<n} |a_n| \binom{n}{i} r^{n-i} \rho^i.$$

Therefore,

$$2\rho^n \leq (r+\rho)^n,$$

which implies that $r \geq (\sqrt[n]{2} - 1)\rho$. To summarize, we have shown the following:

$$(\sqrt[n]{2} - 1)\rho \leq \max_{\alpha : f(\alpha)=0} |\alpha| \leq \rho. \tag{4}$$

But can we get some a priori bounds on $\rho(f)$? Let

$$B := \max \left(\binom{n}{i}^{-1} \left| \frac{a_i}{a_n} \right| \right)^{1/(n-i)}.$$

Then from (3) it follows that $\geq r \geq B$. Moreover, an argument similar for r shows that $(\sqrt[n]{2} - 1)\rho \leq B$, and since $\rho \geq r$, we have the following a priori estimate:

$$B \leq \max_{\alpha : f(\alpha)=0} |\alpha| \leq \frac{B}{(\sqrt[n]{2} - 1)}.$$

The above bounds are useful when we want to bound the absolute value of one root. What if we are interested in bounding the absolute value of products of roots? A naive bound would be roughly $(2\|A\|_\infty)^n$. But this is too pessimistic, as we see next.

1.2 The Measure of a Polynomial

The **Mahler measure**, $M(A)$ is defined as

$$M(A) := |\texttt{lead}(A)| \prod_{i=1}^{n} \max\{1, |\alpha_i|\}. \tag{5}$$

It is often convenient to write $\max_1 a := \max\{1, a\}$. The key property of the Mahler measure is that is an *amortized bound on the products of any subsets of roots*. This insight is often useful in deriving bounds. From the definition, we also have the following nice multiplicative properties:

$$M(x^n A(1/x)) = M(A) \text{ and } M(AB) = M(A)M(B).$$

It trivially follows from the definition and the first equality above that for any root α of $A(x)$

$$\frac{|\mathtt{lead}(A)|}{M(A)} \leq |\alpha| \leq \frac{M(A)}{|\mathtt{lead}(A)|}. \tag{6}$$

But how large is the measure itself? Our earlier bounds on $|\alpha|$ would imply that $M(A) \lessapprox \|A\|_\infty^{n+1}$. We next show that it is in fact much smaller. More precisely,

$$M(A) \leq \|A\|_2 \quad (\leq \sqrt{n+1}\|A\|_\infty). \tag{7}$$

We will need the following technical result to prove Landau's inequality.

LEMMA 2. *For any $c \in \mathbb{C}$, let $\overline{c}$ be its complex conjugate. Then*

$$\|(x - c) \cdot A\|_2 = \|(\overline{c}x - 1) \cdot A\|_2.$$

Proof. Consider the squared-norm (define $a_{-1} = a_{m+1} = 0$)

$$\begin{aligned}
\|(x-c)\cdot A\|_2^2 &= \sum_{i=0}^{m+1}(a_{i-1} - ca_i)\cdot(\overline{a}_{i-1} - \overline{ca}_i) \\
&= (1+|c|^2)\|A\|_2^2 - \sum_{i=0}^{m+1}(a_{i-1}\overline{ca}_i + ca_i\overline{a}_{i-1}) \\
&= (1+|c|^2)\|A\|_2^2 - \sum_{i=0}^{m+1}((a_{i-1}\overline{c})\overline{a}_i + a_i(c\overline{a}_{i-1})) \\
&= \sum_{i=0}^{m+1}(\overline{c}a_{i-1} - a_i)\cdot(c\overline{a}_{i-1} - \overline{a}_i) \\
&= \|(\overline{c}x - 1)A\|_2^2.
\end{aligned}$$

Q.E.D.

Suppose we order the roots of A s.t.

$$|\alpha_1| \geq |\alpha_2| \geq \cdots \geq |\alpha_k| \geq 1 > |\alpha_{k+1}| \geq \cdots \geq |\alpha_n|.$$

Then repeatedly applying Lemma 2 we get that

$$\begin{aligned}
\|A\|_2 &= \|a\prod_{i=1}^{n}(x-\alpha_i)\|_2 \\
&= \|a(\overline{\alpha_1}x - 1)\prod_{i=2}^{n}(x-\alpha_i)\|_2 \\
&= \ldots \\
&= \|a\prod_{i=1}^{k}(\overline{\alpha_i}x - 1)\prod_{i=k+1}^{n}(x-\alpha_i)\|_2.
\end{aligned}$$

Let $B := a\prod_{i=1}^{k}(\overline{\alpha_1}x - 1)\prod_{i=k+1}^{n}(x-\alpha_i)$. Then

$$|\mathtt{lead}(B)| = a\prod_{i=1}^{k}|\overline{\alpha_i}| = a\prod_{i=1}^{k}|\alpha_i| = M(A).$$

Since $\|A\|_2 = \|B\|_2 \geq |\texttt{lead}(A)|$ we get that

$$\|A\|_2 \geq M(A)$$

as desired.

We can now easily bound the absolute values of the roots of sum of two polynomials $A(x), B(x)$. From (6), we know that the absolute values of the roots of $A(x) \pm B(x)$ is bounded by

$$M(A \pm B) \leq \|A \pm B\|_2 \leq \|A\|_2 + \|B\|_2 \leq \|A\|_1 + \|B\|_1 \leq 2^n(M(A) + M(B))$$

where the last inequality follows from the observation that

$$\begin{aligned} |a_{n-i}| &= |a_n| \left| \sum_{1 \leq j_1 < j_2 < \cdots < j_i \leq n} \alpha_{j_1} \dots \alpha_{j_i} \right| \\ &\leq |a_n| \sum_{1 \leq j_1 < j_2 < \cdots < j_i \leq n} |\alpha_{j_1} \dots \alpha_{j_i}| \\ &\leq \sum_{1 \leq j_1 < j_2 < \cdots < j_i \leq n} M(A) \\ &\leq \binom{n}{i} M(A) \end{aligned}$$

which implies that

$$\|A\|_1 \leq 2^n M(A).$$

A similar argument shows that

$$\|A\|_2 \leq \binom{2n}{n}^{1/2} M(A)$$

which gives a tighter bound than the one described above.

An equivalent analytic definition of $M(A)$ is the geometric mean of the absolute value of the polynomial on the unit circle in $\mathbb{C}$:

$$M(A) = \exp\left[\frac{1}{2\pi}\int_0^{2\pi} \log|A(e^{i\theta})| d\theta\right]. \tag{8}$$

¶1. Factor Bounds The measure clearly gives us a good control on the size of the factors dividing a certain polynomial.

2 Root Separation Bounds

Another fundamental bound in working with polynomials is how close can two distinct roots of $A(x)$ be. More precisely, the **minimum root separation** of a polynomial $A(x)$ is defined as

$$\text{sep}(A) := \min_{i,j \in [n]; \alpha_i \neq \alpha_j} |\alpha_i - \alpha_j|. \tag{9}$$

Lower bounds on $\text{sep}(A)$ play a crucial role in the analysis of root finding algorithms.

To study root separation, we introduce another classic tool called the **discriminant** of the polynomial:

$$\text{disc}(A) := a_n^{2n-2} \prod_{1 \leq i < j \leq n} (\alpha_i - \alpha_j)^2. \tag{10}$$

The discriminant tells us whether or not the polynomial has a multiple/repeated root, thus generalizing the concept in the quadratic case. A polynomial with no multiple roots, or only simple roots, is called **square-free** polynomial. It is easy to see that a root of multiplicity m of $A(x)$ occurs as a root of multiplicity $(m-1)$ in $A'(x)$. Based upon this relation, we see the following surprising correlation between $\text{disc}(A)$ and $\texttt{res}(A, A')$, which also implies that $\text{disc}(A)$ is an integer.

Lemma 3.

$$a \cdot \mathrm{disc}(A) = (-1)^{\binom{m}{2}} \mathtt{res}(A, A').$$

Proof. We first show that $A'(\alpha_i) = a \prod_{j=1:j\neq i}^{n} (\alpha_i - \alpha_j)$. To see this observe that

$$A'(x) = a \sum_{k=1}^{n} \prod_{j=1:j\neq k}^{n} (x - \alpha_j).$$

Substituting $x = \alpha_i$ zeros out all products on the RHS except the product that does not contain the term $(x - \alpha_i)$. Thus

$$A'(\alpha_i) = a \prod_{j=1:j\neq i}^{n} (\alpha_i - \alpha_j).$$

Now, we know that

$$\mathtt{res}(A, A') = a^{n-1} \prod_{i=1}^{n} A'(\alpha_i).$$

Substituting the expression for $A'(\alpha_i)$ we immediately get

$$\begin{aligned} \mathtt{res}(A, A') &= a^{n-1} \prod_{i=1}^{n} \left(a \prod_{j=1:j\neq i}^{n} (\alpha_i - \alpha_j) \right) \\ &= a^{2n-1} \prod_{1\le i<j\le n} (-1)(\alpha_i - \alpha_j)^2 \\ &= (-1)^{\binom{n}{2}} a \cdot \mathrm{disc}(A). \end{aligned}$$

Q.E.D.

To show a lower bound on $\mathrm{sep}(A)$, we first derive an upper bound on $\prod_{1\le i<j\le n} |\alpha_i - \alpha_j|$. It is easy to see that $|\alpha_i - \alpha_j| \le 2 \max_1(|\alpha_i|) \max_1(|\alpha_j|)$. Thus

$$\prod_{1\le i<j\le n} |\alpha_i - \alpha_j| \le 2^{\binom{n}{2}} |\alpha_1|^{n-1} |\alpha_2|^{n-1} \dots |\alpha_n|^{n-1} \le 2^{\binom{n}{2}} \left(\frac{M(A)}{a_n} \right)^{n-1}.$$

For any choice of (i, j), let Ω be the set consisting of all pairs of indices (i', j') except (i, j). Then

$$|\alpha_i - \alpha_j| \ge \mathrm{disc}(A)^{1/2} a_n^{1-n} \prod_{(i',j')\in\Omega} |\alpha_{i'} - \alpha_{j'}|^{-1}.$$

Substituting the upper bound derived on the product of roots we obtain

$$|\alpha_i - \alpha_j| \ge \mathrm{disc}(A)^{1/2} 2^{-\binom{n}{2}} M(A)^{1-n}. \tag{11}$$

For the special case, when A is square-free and integer polynomial, we know that $|\mathrm{disc}(A)| \ge 1$. Thus

$$\mathrm{sep}(A) \ge 2^{-\binom{n}{2}} M(A)^{1-n} \ge 2^{-\binom{n}{2}} \|A\|_2^{1-n}. \tag{12}$$

The following theorem generalizes this result to a product of distances between pairs of roots.

Theorem 4 (The Davenport-Mahler-Mignotte Bound [2]). *Let Ω be any set of k pairs of indices (i, j), $1 \le i < j \le n$. Let $\alpha_1, \alpha_2, \dots, \alpha_n$ be the n roots of $A(x)$. Then*

$$2^k M(A)^{2k} \ge \prod_{(i,j)\in\Omega} |\alpha_i - \alpha_j| \ge 2^{k-\binom{n}{2}} M(A)^{2(1-n)} \sqrt{\mathrm{disc}(A)}. \tag{13}$$

Proof. We know from the definition of discriminant that

$$(\mathrm{disc}(A))^{1/2} = a^{n-1} \prod_{1 \le i < j \le n} |\alpha_i - \alpha_j| = a^{n-1} \prod_{(i,j) \in \Omega} |\alpha_i - \alpha_j| \prod_{(i,j) \notin \Omega} |\alpha_i - \alpha_j|. \tag{14}$$

We first derive an upper bound on the second product in the RHS. Applying the inequality

$$|a - b| \le 2 \max_1 |a| \max_1 |b|$$

we get

$$\prod_{(i,j) \notin \Omega} |\alpha_i - \alpha_j| \le 2^{\binom{n}{2} - k} \prod_{(i,j) \notin \Omega} \max_1 |\alpha_i| \max_1 |\alpha_j|.$$

How many times can the term $\max_1 |\alpha_i|$ occur? Given the constraints, it can appear at most $n-1$ times. Thus

$$\prod_{(i,j) \notin \Omega} |\alpha_i - \alpha_j| \le 2^{\binom{n}{2} - k} \prod_{i=1}^{n} \left(\max_1 |\alpha_i|^{n-1} \right) \left(\prod_{j=1}^{n} \max_1 |\alpha_j|^{n-1} \right).$$

Since $\prod_{i=1}^{n} \max_1 |\alpha_i|^{n-1} \le (M(A)/a)^{n-1}$, it follows that

$$\prod_{(i,j) \notin \Omega} |\alpha_i - \alpha_j| \le 2^{\binom{n}{2} - k} (M(A)/a)^{2(n-1)}.$$

Applying this upper bound to (14), we obtain

$$\prod_{(i,j) \in \Omega} |\alpha_i - \alpha_j| \ge a^{n-1} 2^{k - \binom{n}{2}} M(A)^{2(1-n)} \sqrt{\mathrm{disc}(A)}.$$

Q.E.D.

3 Inclusion Bounds – Critical Points and Roots

We have derived in Theorem 1, and equation (6) upper bounds on the absolute value of the roots of $A(x)$. But what about the roots of its derivative $A'(x)$, also called the **critical points** of $A(x)$? As a consequence of the following result it follows that any upper bound on the absolute value of the roots of $A(x)$ is also an upper bound on the absolute value of its critical points, and hence recursively, an upper bound on the roots of all its derivatives. The statement is evident in the case of a quadratic polynomial.

THEOREM 5 (Gauß-Lucas Theorem). *The critical points of a polynomial are contained in the convex hull of its roots.*

Proof. The **logarithmic derivative** of $A(z)$ is given as

$$\frac{A'(z)}{A(z)} = \sum_{i=1}^{n} \frac{1}{z - \alpha_i}. \tag{15}$$

Suppose $\gamma \in \mathbb{C}$ is a critical point of $A(z)$. If γ is a root of $A(x)$ then the claim follows trivially, so assume $\gamma \ne \alpha_i$, $i = 1, \ldots, n$. Then it follows that

$$0 = \sum_{i=1}^{n} \frac{1}{\gamma - \alpha_i} = \sum_{i=1}^{n} \frac{\overline{\gamma} - \overline{\alpha_i}}{|\gamma - \alpha_i|^2}. \tag{16}$$

This implies that

$$\overline{\gamma} \left(\sum_{i=1}^{n} \frac{1}{|\gamma - \alpha_i|^2} \right) = \sum_{i=1}^{n} \frac{\overline{\alpha_i}}{|\gamma - \alpha_i|^2}.$$

Taking conjugates on both sides, we see that γ can be expressed as a convex combination of α_i's. **Q.E.D.**

There is a physical interpretation of the result above. The quantity $1/(z-\alpha)$, for any $z \in \mathbb{C}$, can be interpreted as a vector with direction along the unit vector in the direction of the conjugate $\overline{z-\alpha}$ and magnitude $1/|z-\alpha|$. This associates a force field for any point $\alpha \in \mathbb{C}$. Then (16) can be interpreted as follows: the critical points are the points where the force fields of the n roots sum up to zero, i.e., cancel each other out; in other words, critical points are the equilbrium points in the force field. Now consider any point z outside the convex hull of the roots. As there is a line through z such that all the roots are to one side of it, we can see that the vectors $1/(z-\alpha_i)$ are within a cone centered at origin of angle $< \pi$ and hence they cannot sum up to zero.

The result above thus generalizes Rolle's theorem which states that between any two real roots of $A(x)$ there is at least one critical point of A (or the more general statement that between any two points having the same value of A, there is a critical point of A, which is a special case of the mean value theorem). For the special case when the coefficients of $A(x)$ are real numbers, we can say something more. Rolle's theorem gives us a localization of the real critical points, but what can we say about the non-real critical points? Since $A(x) \in \mathbb{R}[x]$, we know that the non-real roots come as complex-conjugate pairs. For such a pair $\alpha, \overline{\alpha}$, let $J_\alpha \subset \mathbb{C}$ be the disc with diameter the line segment $[\alpha, \overline{\alpha}]$. Such a disc is called Jensen's disc. The following theore gives us the location of the non-real critical points:

Theorem 6 (Jensen's Disc Theorem). *The non-real critical points of a real polynomial are contained in the union of Jensen's discs of its roots.*

Proof. The proof again utilizes the physical interpretation mentioned above. **Q.E.D.**

4 Bounds on Eigenvalues of Matrices

We have derived upper and lower bounds on the roots of a polynomial given the coefficients of the polynomial. In many cases, however, we do not have the coefficients of the polynomial explicitly. One such case, which occurs quite frequently, is when the polynomial is the characteristic polynomial $\chi_A(x)$ of a matrix A and we are interested in bounding the absolute values of the eigenvalues. We can, in principle, compute bounds on the sizes of the coefficients of $\chi_A(x)$ and apply the bounds derived earlier; however, this can be quite bad because the bounds on the coefficients can be quite large (e.g., those given by Hadamard's bound). Ideally, we would like to derive bounds on the absolute values of the eigenvalues directly in terms of the entries of the matrix. The next result gives us more than we desire.

Theorem 7 (Gerschgorin's Disc Theorem). *Let $A = [a_{ij}]$ be a square matrix. Then the eigenvalues $\lambda_1, \dots, \lambda_n$ are contained in the union of the n discs D_i, $i = 1, \dots, n$, where D_i is centered at a_{ii} and has radius $\sum_{j \neq i} |a_{ij}|$. Moreover, if k of these discs form a connected component then the union of these discs contains exactly k eigenvalues counted according to their multiplicities.*

A set in $\mathbb{R}^d$ is called **connected** if there is a path between any two points in the set that is totally contained in the set. A **connected component** of a set is a maximally connected subset of the set, i.e., there is no superset that contains the component and is also connected.

Proof. Consider an eigenvalue λ, and let $\mathbf{x}$ be the corresponding eigenvector. We know that $(A-\lambda I)\mathbf{x} = 0$, or in other words for all $i = 1, \dots, n$, $(a_{ii} - \lambda)x_i = -\sum_{j \neq i} a_{ij} x_j$. Let i be the index for which $|x_i|$ is the maximum, i.e., $|x_i| \geq |x_j|$ for $j \neq i$. Then we have

$$|a_{ii} - \lambda||x_i| \leq \sum_{j \neq i} |a_{ij}||x_j| \leq \sum_{j \neq i} |a_{ij}||x_i|,$$

which implies $\lambda \in D_i$.

The proof of the next statement is by a continuity argument. Let $\mathcal{C}$ be a connected component of k discs; for the sake of convenience, suppose these discs are $D_1, \dots, D_k$; therefore, $\cup_i D_i = \mathcal{C} \cup \mathcal{D}$, where $\mathcal{C}$ and $\mathcal{D}$ are disjoint. Let B be the diagonal matrix with entries a_{ii} on the diagonal. Consider the following family of matrices parametrized by t:

$$A(t) := B + t(A - B).$$

Thus, $A(0) = B$ and at $A(1) = A$; thus $A(t)$ gives us a continuous family of matrices from B to A. By the first result, we know that the eigenvalues of $A(t)$ are contained in $\mathcal{C}(t) := \cup_i D_i(t)$, where

$$D_i(t) := \left\{ z \in \mathbb{C} : |z - a_{ii}| \leq t \sum_{j \neq i} |a_{i,j}| \right\}.$$

Note that $D_i(t) \subset D_i$, hence the eigenvalues of $A(t)$ are contained in $\mathcal{C} \cup \mathcal{D}$. Moreover, as the eigenvalues vary continuously with t and $\mathcal{C} \cap \mathcal{D} = \emptyset$, the number of eigenvalues in $\mathcal{C}$ and $\mathcal{D}$ remain the same for all $t \in [0, 1]$. For $t = 0$, the eigevnvalues are precisely a_{ii}, and $\mathcal{C}$ contains exactly k of them corresponding to the centers of $D_1, \ldots, D_k$. Therefore, at $t = 1$, there are k eigenvalues of A in $\mathcal{C}$.

Q.E.D.

Note that since taking transpose of a matrix does not change the set of eigenvalues, we can take the column sum, instead of the row sum in Theorem 7. What are the other operations that do not affect the eigenvalues of a matrix? There are similarity transformations of a matrix, i.e., the matrix $T^{-1}AT$ for a non-singular matrix T has the same set of eigenvalues. Certain choices of T, for instace $\texttt{diag}(t_1, \ldots, t_n)$, gives us a better handle on the radii of the discs. The (i, j)th entry of this matrix is $t_j a_{i,j}/t_i$, and hence the radii of D_i is

$$\sum_{i \neq j} \left| \frac{t_j a_{i,j}}{t_i} \right|.$$

Using tools from matrix analysis gives us a new approach to handle roots of arbitrary polynomials, since any polynomial can be expressed as the characteristic polynomial of some matrix. Such a matrix is called a **companion matrix** of the polynomial (it's unique only upto similarity transformations). For instance, given $f(x) = \sum_{i=0}^{n} a_i x^i$, one such companion matrix called the **Frobenius matrix of** f is the following:

$$\begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & -\frac{a_0}{a_n} \\ 1 & 0 & 0 & \cdots & 0 & -\frac{a_1}{a_n} \\ 0 & 1 & 0 & \cdots & 0 & -\frac{a_2}{a_n} \\ & \ddots & \ddots & \ddots & \ddots & \vdots \\ & \ddots & \ddots & \cdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -\frac{a_{n-1}}{a_n} \end{bmatrix} \tag{17}$$

A straightforward application of Theorem 7 gives us the following inclusion of all zeros:

$$\max_{\alpha : f(\alpha) = 0} |\alpha| \leq \max \left(1 + \left| \frac{a_i}{a_n} \right| \right).$$

5 Perturbation Bounds

In the previous section, we used the fact that the zeros of the characteristic polynomial depend continuously on the entries of the matrix. Why is that the case? Since the coefficients of the characteristic polynomial are polynomial functions of the entries of the matrix, and as composition of continuos functions is continuous, we can reformulate the question – Why do the roots of a polynomial depend continuously on its coefficients? What does the statement mean qunatitatively? We know that the roots of a polynomial are not polynomial functions of its coefficients (this is evident from the quadratic case) and hence this is not a self-evident statement. Let us assume for the moment that the polynomial is monic, i.e., $A(x) = x^n + \sum_{i=0}^{n-1} a_i x^i$. The claim is that there is a continuous map from $(a_0, \ldots, a_{n-1})$ to the roots $(\alpha_1, \ldots, \alpha_n)$. Using the standard ϵ-δ definition of continuity, this means that for all $\epsilon > 0$ there exists a $\delta > 0$ such that the roots $\beta_1, \ldots, \beta_n$ of all monic polynomials $B(x) = x^n + \sum_{i=0}^{n-1} b_i x^i$, where $|a_i - b_i| \leq \delta$ for $i = 0, \ldots, n-1$, satisfy $|\alpha_i - \beta_i| \leq \epsilon$. A slightly weaker statement is the following: given an ϵ-perturbation of the roots, there is a δ such that the roots of the polynomial B obtained by any perturbation of the coefficients of A smaller than δ is within an ϵ distance of some root of A. Let's see a proof of this. Let α be some root of A. Then $B(\alpha) = \prod_{i=1}^{n} (\alpha - \beta_i)$.

But as $A(\alpha) = 0$, we can write the lhs as $B(\alpha) - f(\alpha) = \sum_{i=0}^{n-1}(b_i - a_i)\alpha^i$. Taking absolute values on both sides we have

$$\prod_{i=1}^{n} |\alpha - \beta_i| \le \sum_{i=0}^{n-1} |b_i - a_i| \, |\alpha|^i.$$

In particular,

$$\min_i |\alpha - \beta_i| \le \left(\sum_{i=0}^{n-1} |b_i - a_i| \, |\alpha|^i \right)^{1/n}.$$

Since $|a_i - b_i| \le \delta$, and by Cauchy's bound $|\alpha| \le (1 + \|f\|_\infty) \le 2 \max_1 \|f\|$, the rhs can be further simplified to

$$\min_i |\alpha - \beta_i| \le 2 \max_1 \|f\| \delta^{1/n}.$$

Thus if we want the lhs to be smaller than ϵ, our δ must be of the order of ϵ^n.

The proof above is independent of the multiplicity of α. A stronger statement is that if α is a root of A with multiplicity m then a small enough δ perturbation of A has m roots near to α. This is the purport of the following statement:

THEOREM 8. *Let A be a monic polynomial with roots $\alpha_1, \ldots, \alpha_k$, where the multiplicity of α_i is m_i. Then given an $\epsilon \le \mathrm{sep}(f)/2$, there is a $\delta > 0$ such that any δ-perturnation B of A has exactly m_i roots in the disc $D(\alpha_i, \epsilon)$.*

Proof.

Q.E.D.

References

[1] M. Mignotte. *Mathematics for Computer Algebra.* Springer-Verlag, Berlin, 1992.

[2] E. P. Tsigaridas and I. Z. Emiris. On the complexity of real root isolation using continued fractions. *Theor. Computer Science*, 392:158–173, 2008.