Real Solutions to Polynomial Equations

1 01/06 – Basic results

We will be interested in estimating the number of real solutions to systems of polynomial equations. We will start off by studying well-known elementary methods for the simplest case: one polynomial with real coefficients in a single variable. We know that such a polynomial may have few real roots even if the degree is high. For example, the polynomial $x^d - 1$ has at most two real roots regardless of the value of d. Let us start by studying the most basic upper bound:

Theorem 1.1 (Descartes' rule of signs – 1637)**.** *Let $f = \sum_{i=0}^{d} c_i x^i \in \mathbb{R}[x]$ and denote the number of positive real roots of f counted with multiplicity as $n(f)$. Then*

$$n(f) \leq \operatorname{signvar}(c_0, c_1, \ldots, c_d),$$

where $\operatorname{signvar}(c_0, c_1, \ldots, c_d)$ *stands for the number of sign changes between consecutive non-zero coefficients. Moreover, the difference between $n(f)$ and* $\operatorname{signvar}(c_0, c_1, \ldots, c_d)$ *is even.*

Notice that it suffices to estimate the number of positive roots of f only, since the number of negative roots of f is the number of positive roots of g, where $g(x) = f(-x)$. Therefore, the number of real roots of f, which we will denote as $N(f)$, is bounded by

$$1 + \operatorname{signvar}(c_0, \ldots, c_d) + \operatorname{signvar}(c_0, -c_1, \ldots, (-1)^d c_d).$$

One thus obtains the following bound:

$$N(f) \leq 1 + 2d,$$

which is obviously very rough. A finer bound is given by

$$N(f) \leq 1 + 2s,$$

where $s + 1$ is the number of non-zero coefficients of f. A nice corollary of this bound is that sparse polynomials have few real roots – in fact, if $s < (d-1)/2$ then there is at least one non-real root. We will see that this bound is optimal in a particular sense that will be explained later.

In order to prove the theorem, we will start with a general lemma first:

Lemma 1.2. *For any differentiable function $f : I \subseteq \mathbb{R} \to \mathbb{R}$ with a finite number of roots in I, we have*

$$n(f) \leq n(f') + 1,$$

where $n(f)$ is the number of roots of f in I counted with multiplicity.

Proof. Let $r_1, \ldots, r_n$ be the different roots of f, with multiplicities $m_1, \ldots, m_n$. Then f' has r_i as a root of multiplicity $m_i - 1$. Moreover, f' has at least one root between any two roots of f. Thus, we obtain

$$n(f') \geq \sum_{i=1}^{n}(m_i - 1) + n - 1 = \sum_{i=1}^{n} m_i - n + n - 1 = n(f) - 1,$$

as wanted. $\square$

We now proceed with the proof of Descartes' rule of signs:

Proof. Let $f(x) = \sum_{i=0}^{d} c_i x^i$. We will suppose without loss of generality that $c_0 c_d \neq 0$, since otherwise we can divide by a factor of the form x^k without altering the number of positive roots.

Suppose then that $c_0 c_d > 0$. Then, $\text{signvar}(c_0, \ldots, c_d)$ is even, since both of its endpoints have the same sign. Moreover, $f(0) \cdot \lim_{x\to\infty} f(x) > 0$, which means that f has the same sign at zero and at infinity. Now, f does not change sign near a root of even multiplicity, so it follows that f has an even number of roots with odd multiplicity. This in turn implies that $n(f)$ is even, and so the difference between $n(f)$ and $\text{signvar}(c_0, \ldots, c_d)$ is even too. If $c_0 c_d < 0$ reasoning analogously we conclude that both numbers are odd, so their difference is even, which proves the last statement of the theorem.

We will now prove the first statement by induction on d. By our previous lemma, we know that $n(f) \leq n(f') + 1$, and by inductive hypothesis we have that $n(f') \leq \text{signvar}(c_1, \ldots, c_d)$.

We now consider two different cases. If $\text{signvar}(c_0, \ldots, c_d) = \text{signvar}(c_1, \ldots, c_d) + 1$, we are done since then $n(f) \leq n(f') + 1 \leq \text{signvar}(c_0, \ldots, c_d)$ as wanted. Otherwise, $\text{signvar}(c_0, \ldots, c_d) = \text{signvar}(c_1, \ldots, c_d)$. This tells us that if $f(x) = c_0 + c_i x^i + \cdots + c_d x^d$, where we only write non-zero coefficients, then $c_0 c_i \geq 0$. Assume without loss of generality that both c_0 and c_i are non-negative (the other case is analogous). In this case we have that $f(0) \geq 0$ and that f is increasing in some interval $(0, \varepsilon)$ not containing any root of f. This implies that f' has a positive root smaller than the smallest positive root of f. Let $r_1 < \cdots < r_n$ be the positive roots of f. Then, by our previous lemma we know that f' has at least $n(f) - 1$ roots (counted with multiplicity) in $(r_1 - \delta, \ldots, r_n + \delta)$ for some $\delta > 0$ such that $\varepsilon < r_1 - \delta$, and at least one in $(0, \varepsilon)$. Therefore

$$n(f) \leq n(f') \leq \text{signvar}(c_1, \ldots, c_d) = \text{signvar}(c_0, \ldots, c_d)$$

as wanted. $\square$

Corollary 1.3. If $\operatorname{signvar}(c_0, \ldots, c_d)$ is odd, then f has at least one real root.

Proof. Since the difference between $n(f)$ and $\operatorname{signvar}(c_0, \ldots, c_d)$ is even, it follows that $n(f)$ is odd, and in particular non-zero. $\square$

We will now provide some generalizations of this theorem. Given a sequence $F = (f_0, \ldots, f_k)$ of real polynomials and some scalar $a \in \overline{\mathbb{R}}$, we define

$$\operatorname{signvar}(F, a) = \operatorname{signvar}(f_0(a), \ldots, f_k(a)),$$

where $f(\pm\infty)$ stands for $\lim_{x\to\pm\infty} f(x)$. For $f \in \mathbb{R}[x]$ of degree d, we will denote

$$\delta f = (f, f', \ldots, f^{(d)})$$

the sequence of consecutive derivatives of f.

Theorem 1.4 (Budan–Fourier). *Let $f = \sum_{i=0}^{d} c_i x^i \in \mathbb{R}[x]$, $a, b \in \overline{\mathbb{R}}$ and denote the number of positive real roots of f counted with multiplicity in $(a, b]$ as $n_{(a,b]}(f)$. Then*

$$n_{(a,b]}(f) \leq \operatorname{signvar}(\delta f, a) - \operatorname{signvar}(\delta f, b).$$

Moreover, the difference between both sides of the inequality is even.

Remark 1.5. When $a = 0$,

$$\operatorname{signvar}(\delta f, a) = \operatorname{signvar}(f(0), f'(0), \ldots, f^{(d)}(0)) = \operatorname{signvar}(c_0, \ldots, c_d).$$

Moreover, it is easy to see that if $b = +\infty$ then $\operatorname{signvar}(\delta f, b) = 0$, and so we conclude that Budan-Fourier contains Descartes' rule of signs as a particular case.

We now give a sketch of the proof:

Proof. The quantity $\operatorname{signvar}(\delta f, t)$ can change only at a root t of some of the polynomials in δf. Let r be a root of f of multiplicity m; at $t = r$, $n_{(a,t]}(f)$ increases by m. Let $\varepsilon > 0$ be such that $(r - \varepsilon, r + \varepsilon)$ does not contain roots different than r of any polynomial in δf. Then, we have that

$$\operatorname{signvar}(\delta f, r) = \operatorname{signvar}(\delta f, r + \varepsilon)$$

and

$$\operatorname{signvar}(\delta f, r - \varepsilon) \geq \operatorname{signvar}(\delta f, r) + m,$$

and by a similar reasoning as in the proof of Descartes' rule of signs, it follows that the difference between both sides is even. We arrive at the result by summing up over all of the roots of f. $\square$

Let us now state (without proof) one more basic result. Given $f \in \mathbb{R}[x]$ of degree d, consider the sequence $F = (f_0, f_1, \ldots, f_k)$ where $f_0 = f$, $f_1 = f'$ and $f_i = -r(f_{i-2}, f_{i-1})$ where r stands for the remainder of the Euclidean algorithm. We stop the sequence at the first k such that $f_{k+1} = 0$.

Theorem 1.6 (Sturm). *Let $f = \sum_{i=0}^{d} c_i x^i \in \mathbb{R}[x]$, $a, b \in \overline{\mathbb{R}}$. Then*

$$n_{(a,b)}(f) = \operatorname{signvar}(F, a) - \operatorname{signvar}(F, b).$$

We now turn to the simplest case of combinatorial patchworking, a technique developed in the early 1980s to produce real plane algebraic curves with prescribed topologies, in order to prove that the Descartes bound is sharp.

Let us write $f(x) = \sum_{i=0}^{d} c_i x^i = \sum_{j=1}^{s} d_j x^{a_j}$, where all $d_j \neq 0$ and $a_1 < \cdots < a_s$. Descartes' rule of signs implies that $n(f) \leq \operatorname{signvar}(d_1, \ldots, d_s)$. Consider the parameterized family of polynomials

$$f_t(x) = \sum_{j=1}^{s} d_j t^{h_j} x^{a_j},$$

where $h_1, \ldots, h_s \in \mathbb{N}$. We set $t > 0$, so that the sign variations of f and f_t match.

Theorem 1.7 (Viro). *Suppose we pick $h_1, \ldots, h_n$ so that the points $(a_1, h_1), \ldots (a_n, h_n) \in \mathbb{N}^2$ are vertices of their convex hull, and moreover, are found in its* lower hull *(that is, the edges of the convex hull such that their normal vectors have negative vertical component). Then, there exists $t_0 > 0$ such that for any fixed $t \in (0, t_0)$,*

$$n(f_t) = \operatorname{signvar}(d_1 t^{h_1}, \ldots, d_s t^{h_s}) = \operatorname{signvar}(d_1, \ldots, d_s).$$

In particular, for any sequence of coefficients, there is a polynomial with the same sign sequence such that the Descartes bound is sharp.

2 03/06

Let us start by proving Viro's theorem:

Proof. Let E_i be the line segment $[(a_i, h_i), (a_{i+1}, h_{i+1})]$. The segment E_i is the graph of an affine function $y \mapsto \alpha_i y + \beta_i$ on $[a_i, a_{i+1}]$. Then

$$\frac{f_t(xt^{-\alpha_i})}{t^{\beta_i}} = \sum_{j=1}^{s} d_j t^{h_j - \alpha_i a_j - \beta_i} x_j^a = d_i x^{a_i} + d_{i+1} x^{a_{i+1}} + p(x, t)$$

for some rational function p in which t appears with positive exponent in all of its terms. This happens since points of the form (a_j, h_j) are vertices of the lower hull, then, $h_j - \alpha_i a_j - \beta_i > 0$ for all $j \neq i, i+1$, and so a positive power of t appears in all these terms.

Let us call $g_i(x) = d_i x^{a_i} + d_{i+1} x^{a_{i+1}}$. By Descartes' rule of signs, we know that g_i has at most one positive, simple root ρ_i, and this happens iff $x^{a_{i+1} - a_i} = -d_i / d_{i+1}$ (recall that no d_j vanishes). This obviously happens iff d_i and d_{i+1} have different signs. If

d_i and d_{i+1} have different signs, the existence of the positive root ρ_i of g_i proves the existence of a real, positive root $t^{-\alpha_i}\rho_{i,t}$ of f_t for small enough t, since f_t is g_i modulo a small perturbation p and simple roots are stable. Repeating this argument for each g_i we obtain a total of $\operatorname{signvar}(d_1, \dots, d_s)$ simple roots.

Consider now a compact set K in $(0, +\infty)$ such that K contains all of the roots $\rho_{i,t}$ and ρ_i: for some suitably small t, the compact sets $t^{-\alpha_1}K, \dots, t^{-\alpha_s}K$ are disjoint, and since $t^{-\alpha_i}\rho_{i,t} \in t^{-\alpha_i}K$, we know that all of our roots $t^{-\alpha_i}\rho_{i,t}$ of f_t are different. We have thus obtained $\operatorname{signvar}(d_1, \dots, d_s)$ roots of f_t and by Descartes' rule of signs we know that there are no more than this. □

Remark 2.1. If $f(x) = \sum_{i=0}^{d} c_i x^i$ with $c_d \neq 0$, then if f has d positive roots then all of the coefficients c_i are non-zero, by Descartes' bound.

Exercise 2.2. What can one say if f has d real (not necessarily positive) roots?

Consider now sufficiently differentiable functions $h_1, \dots, h_k : I \subseteq \mathbb{R} \to \mathbb{R}$. Their *Wronskian* is defined as

$$W(f_1, \dots, f_k) = \det \begin{bmatrix} f_1 & f_2 & \cdots & f_k \\ f_1' & f_2' & \cdots & f_k' \\ \vdots & \vdots & \ddots & \vdots \\ f_1^{(k-1)} & f_2^{(k-1)} & \cdots & f_k^{(k-1)} \end{bmatrix}$$

Obviously, the Wronskian of a set of linearly dependent functions vanishes identically on I, and the converse holds provided the functions are analytic. This may not be true without analitycity: for instance, consider the functions x^2 and $x|x|$ on $(-1, 1)$.

Proposition 2.3. *Let $f_1, \dots, f_k, g : I \subseteq \mathbb{R} \to \mathbb{R}$ be sufficiently differentiable. Then:*

1. *$W(gh_1, \dots, gh_k) = g^k W(h_1, \dots, h_k)$.*
2. *Assuming h_1 does not vanish on I, then*

$$W(h_1, \dots, h_k) = h_1^k W\left(\left(\frac{h_2}{h_1}\right)', \dots, \left(\frac{h_k}{h_1}\right)'\right).$$

Proof. Statement 1. is basic linear algebra. Statement 2. follows from 1. observing that

$$W(h_1, \dots, h_k) = h_1^k W\left(1, \frac{h_2}{h_1}, \dots, \frac{h_k}{h_1}\right)$$

and expanding the determinant by the first column of the matrix. □

Theorem 2.4. *Let $h_1, \dots, h_k : I \subseteq \mathbb{R} \to \mathbb{R}$ be sufficiently differentiable functions such that $W(h_1, \dots, h_i)$ does not vanish on I for any $i = 1, \dots, k$. Then, for any real numbers $a_1, \dots, a_k$ not all zero, the function $f = \sum_{i=1}^{k} a_i h_i$ has at most $k - 1$ roots on I, counted with multiplicity.*

Proof. By induction on k. For $k = 1$ the result is obvious since $W(h_1) = h_1$ does not vanish on I. Let $k > 1$ and consider the result proved for $k' < k$. Call N_f the number of roots of f counted with multiplicity. If $a_2 = \cdots = a_k = 0$, then $f = a_1 h_1$ and the result is once again obvious. Suppose then that not all $a_2, \ldots a_k$ are zero. As h_1 does not vanish on I, we have that $N_f = N_g$, where

$$g = a_1 + \sum_{i=2}^{k} a_i \frac{h_i}{h_1}.$$

As we proved the last class, we may bound $N_g \leq 1 + N_{g'}$. Now,

$$g' = \sum_{i=2}^{k} a_i \left(\frac{h_i}{h_1}\right),$$

and

$$W\left(\left(\frac{h_2}{h_1}\right)', \ldots, \left(\frac{h_i}{h_1}\right)'\right) = \frac{W(h_1, \ldots, h_i)}{h_1^i}$$

and moreover not all of $a_2, \ldots, a_n$ are zero, so the claim follows by induction. $\square$

Exercise 2.5. Compute the Wronskians for the families $\{x^{b_1}, \ldots, x^{b_k}\}$ where $b_i \in \mathbb{N}$ and $\{p_1(x)^{-1}, \ldots, p_k(x)^{-1}\}$ where $p_i(x) \in \mathbb{R}[x]$ are of degree 1. Notice that the second Wronskian is highly dependent on the domain and where the roots of the p_i lie.

We say that $h_1, \ldots, h_k : I \subseteq \mathbb{R} \to \mathbb{R}$ with sufficient derivatives satisfy Descartes' rule of signs (or DRS for short) if for any real numbers $a_1, \ldots, a_k$ not all zero, then the number of real roots of $\sum_{i=1}^{k} a_i h_i$ in I counted with multiplicity does not exceed $\text{signvar}(a_1, \ldots, a_k)$. For example, any set of monomials $x^{b_1}, \ldots, x^{b_k}$ with $b_1 < b_2 < \cdots < b_k \in \mathbb{N}$ satisfies the DRS.

Theorem 2.6. *A set of functions $h_1, \ldots, h_k : I \subseteq \mathbb{R} \to \mathbb{R}$ satisfy the DRS if:*

1. *for any set of indices $j_1, \ldots, j_l \in \{1, \ldots, k\}$, the Wronskian $W(h_{j_1}, \ldots, h_{j_l})$ does not vanish on I, and*
2. *for any pair of sets of indices $j_1, \ldots, j_l$ and $j'_1, \ldots, j'_l \in \{1, \ldots, k\}$ with $j_1 < \cdots < j_l$ and $j'_1 < \cdots < j'_l$, then*
$$W(h_{j_1}, \ldots, h_{j_l}) W(h_{j'_1}, \ldots, h_{j'_l}) > 0$$
on I, that is, they have the same sign.

Exercise 2.7. Check that both conditions are satisfied for ordered monomials. What can one say about the family $\{p_1(x)^{-1}, \ldots, p_k(x)^{-1}\}$ with p_is of degree 1? What about a family of functions of the form $x^{b_i}(1 + x)^{c_i}$?

We will now start studying *systems* of polynomial equations of the form

$$f_1(x_1, \ldots, x_n) = \cdots = f_n(x_1, \ldots, x_n) = 0,$$

where $f_i \in \mathbb{C}[x_1, \ldots, x_n]$. We will write such a polynomial as $\sum_{a \in \mathbb{N}_0^n} c_a x^a$ and denote Q_f the *Newton polytope* of f, which is defined as the convex hull of the set of points in $\mathbb{N}_0^n$ in the support of f. As usual, we are interested in the number N of solutions to the system, but this time we will start by estimating the number of solutions in $\mathbb{C}^n$. The most basic bound follows from Bézout's theorem: assuming N is finite, then

$$N \leq \prod_{i=1}^{n} \deg(f_i).$$

This bound, which we will call the *Bézout bound*, is generically sharp if we fix the degrees of the polynomials involved. We will now state a much sharper bound for a more particular situation:

Theorem 2.8 (Kouchnirenko)**.** *Assume that $f_1, \ldots, f_n$ have the same Newton polytope Q. The number of solutions counted with multiplicity in $(\mathbb{C}^\times)^n$ to the system*

$$f_1(x) = \cdots = f_n(x) = 0,$$

if finite, is bounded by $\mathrm{Vol}_{\mathbb{Z}^n}(Q)$, where $\mathrm{Vol}_{\mathbb{Z}^n}(Q) = n!\,\mathrm{Vol}(Q)$.

Remark 2.9. The symbol $\mathrm{Vol}_{\mathbb{Z}^n}$ is just a renormalization of the usual euclidean volume, so that the unit n-simplex has volume 1. In this way, one easily bounds the number of solutions by cutting up the Newton polytope into unit n-simplices and counting the pieces.