Sturm's theorem

In mathematics, the **Sturm sequence** of a univariate polynomial p is a sequence of polynomials associated with p and its derivative by a variant of Euclid's algorithm for polynomials. **Sturm's theorem** expresses the number of distinct real roots of p located in an interval in terms of the number of changes of signs of the values of the Sturm sequence at the bounds of the interval. Applied to the interval of all the real numbers, it gives the total number of real roots of p.[1]

Whereas the fundamental theorem of algebra readily yields the overall number of complex roots, counted with multiplicity, it does not provide a procedure for calculating them. Sturm's theorem counts the number of distinct real roots and locates them in intervals. By subdividing the intervals containing some roots, it can isolate the roots into arbitrarily small intervals, each containing exactly one root. This yields the oldest real-root isolation algorithm, and arbitrary-precision root-finding algorithm for univariate polynomials.

For computing over the reals, Sturm's theorem is less efficient than other methods based on Descartes' rule of signs. However, it works on every real closed field, and, therefore, remains fundamental for the theoretical study of the computational complexity of decidability and quantifier elimination in the first order theory of real numbers.

The Sturm sequence and Sturm's theorem are named after Jacques Charles François Sturm, who discovered the theorem in 1829.[2]

Contents

The theorem

The **Sturm chain** or **Sturm sequence** of a univariate polynomial $P(x)$ with real coefficients is the sequence of polynomials $P_0, P_1, \ldots,$ such that

$$\begin{aligned} P_0 &= P, \\ P_1 &= P', \\ P_{i+1} &= -\operatorname{rem}(P_{i-1}, P_i), \end{aligned}$$

for $i \geq 1$, where P' is the derivative of P, and $\mathbf{rem}(P_{i-1}, P_i)$ is the remainder of the Euclidean division of P_{i-1} by P_i. The length of the Sturm sequence is at most the degree of P.

The number of sign variations at ξ of the Sturm sequence of P is the number of sign changes–ignoring zeros—in the sequence of real numbers

$$P_0(\xi), P_1(\xi), P_2(\xi), \ldots.$$

This number of sign variations is denoted here $V(\xi)$.

Sturm's theorem states that, if P is a square-free polynomial, the number of distinct real roots of P in the half-open interval $(a, b]$ is $V(a) - V(b)$ (here, a and b are real numbers such that $a < b$).[1]

The theorem extends to unbounded intervals by defining the sign at $+\infty$ of a polynomial as the sign of its leading coefficient (that is, the coefficient of the term of highest degree). At $-\infty$ the sign of a polynomial is the sign of its leading coefficient for a polynomial of even degree, and the opposite sign for a polynomial of odd degree.

In the case of a non-square-free polynomial, if neither a nor b is a multiple root of p, then $V(a) - V(b)$ is the number of *distinct* real roots of P.

The proof of the theorem is as follows: when the value of x increases from a to b, it may pass through a zero of some P_i ($i > 0$); when this occurs, the number of sign variations of (P_{i-1}, P_i, P_{i+1}) does not change. When x passes through a root of $P_0 = P$, the number of sign variations of (P_0, P_1) decreases from 1 to 0. These are the only values of x where some sign may change.

Example

Suppose we wish to find the number of roots in some range for the polynomial $p(x) = x^4 + x^3 - x - 1$. So

$$p_0(x) = p(x) = x^4 + x^3 - x - 1$$
$$p_1(x) = p'(x) = 4x^3 + 3x^2 - 1$$

The remainder of the Euclidean division of p_0 by p_1 is $-\frac{3}{16}x^2 - \frac{3}{4}x - \frac{15}{16}$; multiplying it by -1 we obtain

$$p_2(x) = \tfrac{3}{16}x^2 + \tfrac{3}{4}x + \tfrac{15}{16}.$$

Next dividing p_1 by p_2 and multiplying the remainder by -1, we obtain

$$p_3(x) = -32x - 64.$$

Now dividing p_2 by p_3 and multiplying the remainder by -1, we obtain

$$p_4(x) = -\tfrac{3}{16}.$$

As this is a constant, this finishes the computation of the Sturm sequence.

To find the number of real roots of p_0 one has to evaluate the sequences of the signs of these polynomials at $-\infty$ and ∞, which are respectively $(+, -, +, +, -)$ and $(+, +, +, -, -)$. Thus

$$V(-\infty) - V(+\infty) = 3 - 1 = 2,$$

which shows that p has two real roots.

This can be verified by noting that $p(x)$ can be factored as $(x^2 - 1)(x^2 + x + 1)$, where the first factor has the roots -1 and 1, and second factor has no real roots. This last assertion results from the quadratic formula, and also from Sturm's theorem, which gives the sign sequences $(+, -, -)$ at $-\infty$ and $(+, +, -)$ at $+\infty$.

Generalization

Sturm sequences have been generalized in two directions. To define each polynomial in the sequence, Sturm used the negative of the remainder of the Euclidean division of the two preceding ones. The theorem remains true if one replaces the negative of the remainder by its product or quotient by a positive constant or the square of a polynomial. It is also useful (see below) to consider sequences where the second polynomial is not the derivative of the first one.

A *generalized Sturm sequence* is a finite sequence of polynomials with real coefficients

$$P_0, P_1, \ldots, P_m$$

such that

- the degrees are decreasing after the first one: $\deg P_i < \deg P_{i-1}$ for $i = 2, \ldots, m$;
- P_m does not have any real root or has no sign changes near its real roots.
- if $P_i(\xi) = 0$ for $0 < i < m$ and ξ a real number, then $P_{i-1}(\xi)\, P_{i+1}(\xi) < 0$.

The last condition implies that two consecutive polynomials do not have any common real root. In particular the original Sturm sequence is a generalized Sturm sequence, if (and only if) the polynomial has no multiple real root (otherwise the first two polynomials of its Sturm sequence have a common root).

When computing the original Sturm sequence by Euclidean division, it may happen that one encounters a polynomial that has a factor that is never negative, such a x^2 or $x^2 + 1$. In this case, if one continues the computation with the polynomial replaced by its quotient by the nonnegative factor, one gets a generalized Sturm sequence, which may also be used for computing the number of real roots, since the proof of Sturm's theorem still applies (because of the third condition). This may sometimes simplify the computation, although it is generally difficult to find such nonnegative factors, except for even powers of x.

Use of pseudo-remainder sequences

In computer algebra, the polynomials that are considered have integer coefficients or may be transformed to have integer coefficients. The Sturm sequence of a polynomial with integer coefficients generally contains polynomials whose coefficients are not integers (see above example).

To avoid computation with rational numbers, a common method is to replace Euclidean division by pseudo-division for computing polynomial greatest common divisors. This amounts to replacing the remainder sequence of the Euclidean algorithm by a pseudo-remainder sequence, a pseudo remainder sequence being a sequence $p_0, \ldots, p_k$ of polynomials such that there are constants a_i and b_i such that $b_i p_{i+1}$ is the remainder of the Euclidean division of $a_i p_{i-1}$ by p_i. (The different kinds of pseudo-remainder sequences are defined by the choice of a_i and b_i; typically, a_i is chosen for not introducing denominators during Euclidean division, and b_i is a common divisor of the coefficients of the resulting

remainder; see Pseudo-remainder sequence for details.) For example, the remainder sequence of the Euclidean algorithm is a pseudo-remainder sequence with $a_i = b_i = 1$ for every i, and the Sturm sequence of a polynomial is a pseudo-remainder sequence with $a_i = 1$ and $b_i = -1$ for every i.

Various pseudo-remainder sequences have been designed for computing greatest common divisors of polynomials with integer coefficients without introducing denominators (see Pseudo-remainder sequence). They can all be made generalized Sturm sequences by choosing the sign of the b_i to be the opposite of the sign of the a_i. This allows the use of Sturm's theorem with pseudo-remainder sequences.

Root isolation

For a polynomial with real coefficients, *root isolation* consists of finding, for each real root, an interval that contains this root, and no other roots.

This is useful for root finding, allowing the selection of the root to be found and providing a good starting point for fast numerical algorithms such as Newton's method; it is also useful for certifying the result, as if Newton's method converge outside the interval one may immediately deduce that it converges to the wrong root.

Root isolation is also useful for computing with algebraic numbers. For computing with algebraic numbers, a common method is to represent them as a pair of a polynomial to which the algebraic number is a root, and an isolation interval. For example $\sqrt{2}$ may be unambiguously represented by $(x^2 - 2, [0, 2])$.

Sturm's theorem provides a way for isolating real roots that is less efficient (for polynomials with integer coefficients) than other methods involving Descartes' rule of signs. However, it remains useful in some circumstances, mainly for theoretical purposes, for example for algorithms of real algebraic geometry that involve infinitesimals.

For isolating the real roots, one starts from an interval $(a, b]$ containing all the real roots, or the roots of interest (often, typically in physical problems, only positive roots are interesting), and one computes $V(a)$ and $V(b)$. For defining this starting interval, one may use bounds on the size of the roots (see Properties of polynomial roots § Bounds on (complex) polynomial roots). Then, one divides this interval in two, by choosing c in the middle of $(a, b]$. The computation of $V(c)$ provides the number of real roots in $(a, c]$ and $(c, b]$, and one may repeat the same operation on each subinterval. When one encounters, during this process an interval that does not contain any root, it may be suppressed from the list of intervals to consider. When one encounters an interval containing exactly one root, one may stop dividing it, as it is an isolation interval. The process stops eventually, when only isolating intervals remain.

This isolating process may be used with any method for computing the number of real roots in an interval. Theoretical complexity analysis and practical experiences show that methods based on Descartes' rule of signs are more efficient. It follows that, nowadays, Sturm sequences are rarely used for root isolation.

Application

Generalized Sturm sequences allow counting the roots of a polynomial where another polynomial is positive (or negative), without computing these root explicitly. If one knows an isolating interval for a root of the first polynomial, this allows also finding the sign of the second polynomial at this particular root of the first polynomial, without computing a better approximation of the root.

Let $P(x)$ and $Q(x)$ be two polynomials with real coefficients such that P and Q have no common root and P has no multiple roots. In other words, P and $P'Q$ are coprime polynomials. This restriction does not really affect the generality of what follows as GCD computations allows reducing the general case to this case, and the cost of the computation of a Sturm sequence is the same as that of a GCD.

Let $W(a)$ denote the number of sign variations at a of a generalized Sturm sequence starting from P and $P'Q$. If $a < b$ are two real numbers, then $W(a) - W(b)$ is the number of roots of P in the interval $(a, b]$ such that $Q(a) > 0$ minus the number of roots in the same interval such that $Q(a) < 0$. Combined with the total number of roots of P in the same interval given by Sturm's theorem, this gives the number of roots of P such that $Q(a) > 0$ and the number of roots of P such that $Q(a) < 0$.[1]

See also

- Routh–Hurwitz theorem
- Hurwitz's theorem (complex analysis)
- Descartes' rule of signs
- Rouché's theorem
- Properties of polynomial roots
- Gauss–Lucas theorem
- Turán's inequalities

References

1. (Basu, Pollack & Roy 2006)
2. O'Connor, John J.; Robertson, Edmund F. "Sturm's theorem" (https://mathshistory.st-andrews.ac.uk/Biographies/Sturm.html). *MacTutor History of Mathematics archive*. University of St Andrews.

- Basu, Saugata; Pollack, Richard; Roy, Marie-Françoise (2006). "Section 2.2.2". *Algorithms in real algebraic geometry* (2nd ed.). Springer. pp. 52–57. ISBN 978-3-540-33098-1.
- Sturm, Jacques Charles François (1829). "Mémoire sur la résolution des équations numériques". *Bulletin des Sciences de Férussac*. **11**: 419–425.
- Sylvester, J. J. (1853). "On a theory of the syzygetic relations of two rational integral functions, comprising an application to the theory of Sturm's functions, and that of the greatest algebraical common measure" (https://zenodo.org/record/1432412). *Phil. Trans. R. Soc. Lond*. **143**: 407–548. doi:10.1098/rstl.1853.0018 (https://doi.org/10.1098%2Frstl.1853.0018). JSTOR 108572 (https://www.jstor.org/stable/108572).
- Thomas, Joseph Miller (1941). "Sturm's theorem for multiple roots". *National Mathematics Magazine*. **15** (8): 391–394. doi:10.2307/3028551 (https://doi.org/10.2307%2F3028551). JSTOR 3028551 (https://www.jstor.org/stable/3028551). MR 0005945 (https://www.ams.org/mathscinet-getitem?mr=0005945).
- Heindel, Lee E. (1971). *Integer arithmetic algorithms for polynomial real zero determination*. *Proc. SYMSAC '71*. p. 415. doi:10.1145/800204.806312 (https://doi.org/10.1145%2F800204.806312). MR 0300434 (https://www.ams.org/mathscinet-getitem?mr=0300434). S2CID 9971778 (https://api.semanticscholar.org/CorpusID:9971778).
- Panton, Don B.; Verdini, William A. (1981). "A fortran program for applying Sturm's theorem in counting internal rates of return". *J. Financ. Quant. Anal*. **16** (3): 381–388. doi:10.2307/2330245 (https://doi.org/10.2307%2F2330245). JSTOR 2330245 (https://www.jstor.org/stable/2330245).

- Akritas, Alkiviadis G. (1982). "Reflections on a pair of theorems by Budan and Fourier". *Math. Mag*. **55** (5): 292–298. doi:10.2307/2690097 (https://doi.org/10.2307%2F2690097). JSTOR 2690097 (https://www.jstor.org/stable/2690097). MR 0678195 (https://www.ams.org/mathscinet-getitem?mr=0678195).
- Pedersen, Paul (1991). "Multivariate Sturm theory". In Mattson, Harold F.; Mora, Teo; Rao, T. R. N. (eds.). *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes, 9th International Symposium, AAECC-9, New Orleans, LA, USA, October 7–11, 1991, Proceedings*. Lecture Notes in Computer Science. Vol. 539. Berlin: Springer. pp. 318–332. doi:10.1007/3-540-54522-0_120 (https://doi.org/10.1007%2F3-540-54522-0_120). ISBN 978-3-540-54522-4. MR 1229329 (https://www.ams.org/mathscinet-getitem?mr=1229329).
- Yap, Chee (2000). *Fundamental Problems in Algorithmic Algebra* (http://www.cs.nyu.edu/yap/book/berlin/). Oxford University Press. ISBN 0-19-512516-9.
- Rahman, Q. I.; Schmeisser, G. (2002). *Analytic theory of polynomials*. London Mathematical Society Monographs. New Series. Vol. 26. Oxford: Oxford University Press. ISBN 0-19-853493-0. Zbl 1072.30006 (https://zbmath.org/?format=complete&q=an:1072.30006).
- Baumol, William. *Economic Dynamics*, chapter 12, Section 3, "Qualitative information on real roots"
- D.G. Hook and P. R. McAree, "Using Sturm Sequences To Bracket Real Roots of Polynomial Equations" in Graphic Gems I (A. Glassner ed.), Academic Press, pp. 416–422, 1990.
