IDP Talk

A Formalisation of Sturm's Theorem in Isabelle/HOL

Manuel Eberl
eberlm@in.tum.de

Motivation

We have: a polynomial with real coefficients

Motivation

We have: a polynomial with real coefficients

We want: the number of real roots in a specific interval

Motivation

We have: a polynomial with $\overbrace{\text{real coefficients}}^{\text{not really}}$

We want: the number of real roots in a specific interval

Motivation

We have: a polynomial with $\overbrace{\text{real coefficients}}^{\text{not really}}$

We want: the number of real roots in a specific interval

For "real" computations: restricted to appropriate subset of $\mathbb{R}$, such as $\mathbb{Q}$.

Motivation

The solution: Sturm's Theorem

Provides a method for counting real roots algorithmically.

$\Longrightarrow$ Let's formalise it in Isabelle/HOL

Notation

Sign changes: $\sigma(\mathrm{P}_0, \ldots, \mathrm{P}_{n-1}; x)$ denotes denotes the number of sign changes in the sequence $\mathrm{P}_0(x), \ldots, \mathrm{P}_{n-1}(x)$

For the functionally inclined:

$$\sigma(\mathrm{ps}; x) = (\mathsf{length} \circ \mathsf{remdups\_adj} \circ \mathsf{filter}\ (\neq 0) \circ \mathsf{map}\ (\lambda p.\ p(x)))\ \mathrm{ps}\ -\ 1$$

Sturm's Theorem

Sturm's Theorem: Let P be a real polynomial and $\mathrm{P}_0,\ldots,\mathrm{P}_{n-1}$ a Sturm sequence of P. Then

$$\sigma(\mathrm{P}_0,\ldots,\mathrm{P}_{n-1};a)-\sigma(\mathrm{P}_0,\ldots,\mathrm{P}_{n-1};b)$$

is the number of real roots of P in the interval $(a;b]$.

Sturm sequence

But what is a Sturm sequence of P?
A (reasonably) general definition:

- $P_0 = P$

Sturm sequence

But what is a Sturm sequence of P?
A (reasonably) general definition:

- $P_0 = P$
- P_0 and P_1 have no common roots

Sturm sequence

But what is a Sturm sequence of P?
A (reasonably) general definition:

- $P_0 = P$
- P_0 and P_1 have no common roots
- P_{n-1} (last element) does not change its sign

Sturm sequence

But what is a Sturm sequence of P?
A (reasonably) general definition:

- $P_0 = P$
- P_0 and P_1 have no common roots
- P_{n-1} (last element) does not change its sign
- if x_0 is root of P_0: $P_0P_1(x) < 0$ in some left-NH and $P_0P_1(x) > 0$ in some right-NH of x_0

Sturm sequence

But what is a Sturm sequence of P?
A (reasonably) general definition:

- $\mathrm{P}_0 = \mathrm{P}$
- P_0 and P_1 have no common roots
- $\mathrm{P}_{\mathrm{n}-1}$ (last element) does not change its sign
- if x_0 is root of P_0: $\mathrm{P}_0\mathrm{P}_1(\mathrm{x}) < 0$ in some left-NH and $\mathrm{P}_0\mathrm{P}_1(\mathrm{x}) > 0$ in some right-NH of x_0
- if x_0 is root of another P_i: $\mathrm{P}_{\mathrm{i}-1}\mathrm{P}_{\mathrm{i}+1}(\mathrm{x}_0) < 0$

Assessment

The good news:

formalisation of real analysis, polynomials, algebra already exists

Assessment

The good news:

formalisation of real analysis, polynomials, algebra already exists

The bad news:

no formalisation of limits of polynomials, very little on divisibility

Assessment

The good news:
: formalisation of real analysis, polynomials, algebra already exists

The bad news:
: no formalisation of limits of polynomials, very little on divisibility

The ugly news:
: textbook proofs of Sturm's theorem are extremely informal proof sketchs at best

Proving Sturm's Theorem

Assume we already have a Sturm chain. Why does it count roots? Follow $x \mapsto \sigma(P_0,\ldots,P_{n-1};x)$ passing over $\mathbb{R}$. Obviously, it can only change at x_0 if one of the P_i has a root at x_0

Proving Sturm's Theorem

Assume we already have a Sturm chain. Why does it count roots? Follow $x \mapsto \boldsymbol{\sigma}(P_0,\ldots,P_{n-1};x)$ passing over $\mathbb{R}$. Obviously, it can only change at x_0 if one of the P_i has a root at x_0

- if $P_i \neq P_0$ has x_0 as a root, $P_{i-1}P_{i+1}(x_0) < 0$

Proving Sturm's Theorem

Assume we already have a Sturm chain. Why does it count roots? Follow $x \mapsto \boldsymbol{\sigma}(P_0, \ldots, P_{n-1}; x)$ passing over $\mathbb{R}$. Obviously, it can only change at x_0 if one of the P_i has a root at x_0

- if $P_i \neq P_0$ has x_0 as a root, $P_{i-1}P_{i+1}(x_0) < 0$
  $\Rightarrow$ signs of P_{i-1}, P_{i+1} are $\neq 0$, opposite, constant in NH of x_0

Proving Sturm's Theorem

Assume we already have a Sturm chain. Why does it count roots? Follow $x \mapsto \boldsymbol{\sigma}(P_0,\ldots,P_{n-1};x)$ passing over $\mathbb{R}$. Obviously, it can only change at x_0 if one of the P_i has a root at x_0

- if $P_i \neq P_0$ has x_0 as a root, $P_{i-1}P_{i+1}(x_0) < 0$
  $\Rightarrow$ signs of P_{i-1}, P_{i+1} are $\neq 0$, opposite, constant in NH of x_0
  $\Rightarrow$ signs $[1, \_\,, -1]$ or $[-1, \_\,, 1]$ in the entire NH, i.e. one sign change

Proving Sturm's Theorem

Assume we already have a Sturm chain. Why does it count roots?
Follow $x \mapsto \sigma(P_0, \ldots, P_{n-1}; x)$ passing over $\mathbb{R}$. Obviously, it can only change at x_0 if one of the P_i has a root at x_0

- if $P_i \neq P_0$ has x_0 as a root, $P_{i-1}P_{i+1}(x_0) < 0$
  $\Rightarrow$ signs of P_{i-1}, P_{i+1} are $\neq 0$, opposite, constant in NH of x_0
  $\Rightarrow$ signs $[1, \_\,, -1]$ or $[-1, \_\,, 1]$ in the entire NH, i.e. one sign change
  $\Rightarrow$ total number of sign changes not influenced

Proving Sturm's Theorem

Assume we already have a Sturm chain. Why does it count roots? Follow $x \mapsto \sigma(P_0, \ldots, P_{n-1}; x)$ passing over $\mathbb{R}$. Obviously, it can only change at x_0 if one of the P_i has a root at x_0

- if $P_i \neq P_0$ has x_0 as a root, $P_{i-1}P_{i+1}(x_0) < 0$
  $\Rightarrow$ signs of P_{i-1}, P_{i+1} are $\neq 0$, opposite, constant in NH of x_0
  $\Rightarrow$ signs $[1, \_, -1]$ or $[-1, \_, 1]$ in the entire NH, i.e. one sign change
  $\Rightarrow$ total number of sign changes not influenced
- if P_0 has x_0 as root, $P_0P_1(x_0) < 0$ in left-NH of x_0, > 0 in right-NH

Proving Sturm's Theorem

Assume we already have a Sturm chain. Why does it count roots?
Follow $x \mapsto \sigma(P_0, \ldots, P_{n-1}; x)$ passing over $\mathbb{R}$. Obviously, it can only change at x_0 if one of the P_i has a root at x_0

- if $P_i \neq P_0$ has x_0 as a root, $P_{i-1}P_{i+1}(x_0) < 0$
  $\Rightarrow$ signs of P_{i-1}, P_{i+1} are $\neq 0$, opposite, constant in NH of x_0
  $\Rightarrow$ signs $[1, \_\,, -1]$ or $[-1, \_\,, 1]$ in the entire NH, i.e. one sign change
  $\Rightarrow$ total number of sign changes not influenced
- if P_0 has x_0 as root, $P_0P_1(x_0) < 0$ in left-NH of x_0, > 0 in right-NH
  $\Rightarrow$ signs are different left of x_0 and the same right of x_0

Proving Sturm's Theorem

Assume we already have a Sturm chain. Why does it count roots?
Follow $x \mapsto \sigma(P_0, \dots, P_{n-1}; x)$ passing over $\mathbb{R}$. Obviously, it can only change at x_0 if one of the P_i has a root at x_0

- if $P_i \neq P_0$ has x_0 as a root, $P_{i-1}P_{i+1}(x_0) < 0$
  $\Rightarrow$ signs of P_{i-1}, P_{i+1} are $\neq 0$, opposite, constant in NH of x_0
  $\Rightarrow$ signs $[1, \_, -1]$ or $[-1, \_, 1]$ in the entire NH, i.e. one sign change
  $\Rightarrow$ total number of sign changes not influenced
- if P_0 has x_0 as root, $P_0P_1(x_0) < 0$ in left-NH of x_0, > 0 in right-NH
  $\Rightarrow$ signs are different left of x_0 and the same right of x_0
  $\Rightarrow$ total number of sign changes decreases by one

Proving Sturm's Theorem

Formal proof: a lot of induction on the sequences and number of roots

$\Longrightarrow$ messy and not terribly interesting, I'll spare you the details

Proving Sturm's Theorem

We now know that Sturm sequences can count roots.
But how do we construct one?

Construction Sturm sequences

Canonical construction for P with no multiple roots (i.e. $\gcd(P, P') = 1$):

$$P_i = \begin{cases} P & \text{for } i = 0 \\ P' & \text{for } i = 1 \\ -(P_{i-2} \bmod P_{i-1}) & \text{otherwise} \end{cases}$$

Construction Sturm sequences

Why does it work? Nonobvious parts:

Construction Sturm sequences

Why does it work? Nonobvious parts:

If x_0 is root of $P_0 = P$: $PP'(x_0) < 0$ in left-NH and $PP'(x_0) > 0$ in right-NH

– pick neighbourhood without roots of P_0 and P_1 (except for x_0), apply mean value theorem

Construction Sturm sequences

Why does it work? Nonobvious parts:

If x_0 is root of $P_0 = P$: $PP'(x_0) < 0$ in left-NH and $PP'(x_0) > 0$ in right-NH
– pick neighbourhood without roots of P_0 and P_1 (except for x_0), apply mean value theorem

If x_0 is root of another P_i: $P_{i-1}P_{i+1}(x_0) < 0$ in some NH of x_0
– by construction, $P_{i-1} = Q \cdot P_i - P_{i+1}$ for some $Q \in \mathbb{R}[X]$
$\Longrightarrow P_{i-1}(x_0) = -P_{i+1}(x_0)$
also: $P_{i-1}(x_0) \neq 0$ since $\gcd(P_{i-1}, P_i) = \gcd(P_0, P_1) = 1$

Construction Sturm sequences

This construction assumed no multiple roots.
What do we do if there are multiple roots?

Construction Sturm sequences

In case of multiple roots: Let $D := \mathsf{gcd}(P, P')$. Then:

Construction Sturm sequences

In case of multiple roots: Let $D := \mathbf{gcd}(P, P')$. Then:

The obvious way:

- compute canonical Sturm chain of P/D
  (“divide out” multiple roots)

Construction Sturm sequences

In case of multiple roots: Let $D := \mathbf{gcd}(P, P')$. Then:

The obvious way:

- compute canonical Sturm chain of P/D ("divide out" multiple roots)

The clever way:

- we can compute the canonical Sturm chain of P and divide by D afterwards

Construction Sturm sequences

In case of multiple roots: Let $D := \mathsf{gcd}(P, P')$. Then:

The obvious way:

- compute canonical Sturm chain of P/D ("divide out" multiple roots)

The clever way:

- we can compute the canonical Sturm chain of P and divide by D afterwards
- but: if $D(x) \neq 0$, dividing by D does not change the number of sign changes at x

Construction Sturm sequences

In case of multiple roots: Let $D := \mathsf{gcd}(P, P')$. Then:

The obvious way:

- compute canonical Sturm chain of P/D ("divide out" multiple roots)

The clever way:

- we can compute the canonical Sturm chain of P and divide by D afterwards
- but: if $D(x) \neq 0$, dividing by D does not change the number of sign changes at x
- $\Longrightarrow$ unless the interval bounds are multiple roots, we can use the canonical construction without changes

Making a decision procedure

count_roots_between p a b: picks the most efficient Sturm chain construction and:

$$\text{count\_roots\_between p a b} \;=\; |\{\text{x. a} < \text{x} \wedge \text{x} \leq \text{b} \wedge \text{p(x)} = 0\}|$$

Making a decision procedure

count_roots_between p a b: picks the most efficient Sturm chain construction and:

$$\text{count\_roots\_between p a b} \;\; = \;\; |\{\text{x. a} < \text{x} \;\wedge\; \text{x} \leq \text{b} \;\wedge\; \text{p}(\text{x}) = 0\}|$$

Some fluff:

- case distinctions allow arbitrary combination og $\leq$ and $<$ in bounds
- “limit signs” allow infinite bounds

In summary: we can count roots in any open/halfopen/closed, bounded/unbounded real interval

Making a decision procedure

Some more fluff:

- and/or: count x with

$$P(x) = 0 \ \wedge \ Q(x) = 0 \quad \text{or} \quad P(x) = 0 \ \vee \ Q(x) = 0$$

Making a decision procedure

Some more fluff:

- and/or: count x with

$$P(x) = 0 \ \wedge \ Q(x) = 0 \qquad \text{or} \qquad P(x) = 0 \ \vee \ Q(x) = 0$$

- $\forall$-inequalities:

$$\forall x.\ P(x) \neq Q(x) \ \wedge \ R(x) \neq S(x) \ \vee \ T(x) \neq U(x)$$

Making a decision procedure

Some more fluff:

- and/or: count x with

$$P(x) = 0 \wedge Q(x) = 0 \quad \text{or} \quad P(x) = 0 \vee Q(x) = 0$$

- $\forall$-inequalities:

$$\forall x.\ P(x) \neq Q(x) \wedge R(x) \neq S(x) \vee T(x) \neq U(x)$$

- $\forall$ with $<$ and $>$:

$$\forall x.\ P(x) < Q(x) \wedge R(x) > S(x) \vee T(x) \neq U(x)$$

Making a decision procedure

Examples:

lemma "card $\{x::real.\ (x-1)^2 * (x+1) = 0\} = 2$" **by** sturm

lemma "card $\{x::real.\ -0.010831 < x \wedge x < 0.010831 \wedge$

poly $[:0,\ -17/2097152,\ -49/16777216,\ 1/6,\ 1/24,\ 1/120:]\ x = 0\} = 3$" **by** sturm

lemma "card $\{x::real.\ x^3 + x = 2*x^2 \wedge x^3 - 6*x^2 + 11*x = 6\} = 1$" **by** sturm

lemma "$(x::real)^2 + 1 > 0$" **by** sturm

Size of the formalisation

3725 LOC in total, 185 of that ML, the rest Isabelle