Bounds on strong unicity for Chebyshev approximation with bounded coefficients

Andrei Sipoș

Technische Universität Darmstadt
Institute of Mathematics of the Romanian Academy

August 15, 2019
Logic Colloquium 2019
Praha, Česko

Proof mining:

- an applied subfield of mathematical logic

Proof mining:

- an applied subfield of mathematical logic
- goals: to find explicit and uniform witnesses or bounds and to remove superfluous premises from concrete mathematical statements by analyzing their proofs

Proof mining

Proof mining:

- an applied subfield of mathematical logic
- goals: to find explicit and uniform witnesses or bounds and to remove superfluous premises from concrete mathematical statements by analyzing their proofs
- tools used: primarily proof interpretations (modified realizability, negative translation, functional interpretation)

A brief history

- Early efforts
  - David Hilbert: "Über das Unendliche" (1926)
  - Grete Hermann: "The Question of Finitely Many Steps in Polynomial Ideal Theory" (1926)

A brief history

- Early efforts
  - David Hilbert: "Über das Unendliche" (1926)
  - Grete Hermann: "The Question of Finitely Many Steps in Polynomial Ideal Theory" (1926)
- Georg Kreisel's program of "unwinding of proofs"
  - the shift of emphasis in the early 1950s
  - Kreisel: Littlewood's theorem, Hilbert's 17th problem (1957)
  - the publication of Gödel's *Dialectica* interpretation (1958)
  - Jean-Yves Girard: bounds on van der Waerden numbers by strategic cut elimination (1987)
  - Horst Luckhardt: growth conditions on Herbrand terms and the number of solutions in Roth's theorem (1989)

A brief history

- Early efforts
  - David Hilbert: "Über das Unendliche" (1926)
  - Grete Hermann: "The Question of Finitely Many Steps in Polynomial Ideal Theory" (1926)
- Georg Kreisel's program of "unwinding of proofs"
  - the shift of emphasis in the early 1950s
  - Kreisel: Littlewood's theorem, Hilbert's 17th problem (1957)
  - the publication of Gödel's *Dialectica* interpretation (1958)
  - Jean-Yves Girard: bounds on van der Waerden numbers by strategic cut elimination (1987)
  - Horst Luckhardt: growth conditions on Herbrand terms and the number of solutions in Roth's theorem (1989)
- Ulrich Kohlenbach: contemporary proof mining
  - uniqueness in approximation theory (since 1990)
  - nonlinear analysis, convex optimization et al. (since 2001)
  - ergodic theory, commutative algebra, differential algebra: work by Avigad, Towsner, Simmons (since 2007)

A brief history

- Early efforts
  - David Hilbert: "Über das Unendliche" (1926)
  - Grete Hermann: "The Question of Finitely Many Steps in Polynomial Ideal Theory" (1926)
- Georg Kreisel's program of "unwinding of proofs"
  - the shift of emphasis in the early 1950s
  - Kreisel: Littlewood's theorem, Hilbert's 17th problem (1957)
  - the publication of Gödel's *Dialectica* interpretation (1958)
  - Jean-Yves Girard: bounds on van der Waerden numbers by strategic cut elimination (1987)
  - Horst Luckhardt: growth conditions on Herbrand terms and the number of solutions in Roth's theorem (1989)
- Ulrich Kohlenbach: contemporary proof mining
  - uniqueness in approximation theory (since 1990)
  - nonlinear analysis, convex optimization et al. (since 2001)
  - ergodic theory, commutative algebra, differential algebra: work by Avigad, Towsner, Simmons (since 2007)

Chebyshev approximation

We have the following classical Chebyshev approximation result.

Theorem (de la Vallée Poussin, Young – 1900s)

For every $n \in \mathbb{N}$ and every continuous $f : [0,1] \to \mathbb{R}$ there is an unique $p \in P_n$ (the set of real polynomials of degree at most n) such that

$$\|f - p\| = \min_{q \in P_n} \|f - q\|$$

(where $\|\cdot\|$ denotes the supremum norm).

Chebyshev approximation

We have the following classical Chebyshev approximation result.

Theorem (de la Vallée Poussin, Young – 1900s)

For every $n \in \mathbb{N}$ and every continuous $f : [0,1] \to \mathbb{R}$ there is an unique $p \in P_n$ (the set of real polynomials of degree at most n) such that

$$\|f - p\| = \min_{q \in P_n} \|f - q\|$$

(where $\|\cdot\|$ denotes the supremum norm).

Kohlenbach extracted in 1990 a *modulus of uniqueness* – a function Ψ with the property that if p_1 and p_2 are such that $\|f - p_1\|$, $\|f - p_2\| \leq \min + \Psi(\delta)$, then $\|p_1 - p_2\| \leq \delta$.

Chebyshev approximation

We have the following classical Chebyshev approximation result.

Theorem (de la Vallée Poussin, Young – 1900s)

For every $n \in \mathbb{N}$ and every continuous $f : [0,1] \to \mathbb{R}$ there is an unique $p \in P_n$ (the set of real polynomials of degree at most n) such that

$$\|f - p\| = \min_{q \in P_n} \|f - q\|$$

(where $\|\cdot\|$ denotes the supremum norm).

Kohlenbach extracted in 1990 a *modulus of uniqueness* – a function Ψ with the property that if p_1 and p_2 are such that $\|f - p_1\|$, $\|f - p_2\| \leq \min + \Psi(\delta)$, then $\|p_1 - p_2\| \leq \delta$.

He did this by analyzing the uniqueness proof and obtaining an approximate version of it. Let us see how the original proof flows.

A sketch of de la Vallée Poussin's proof

Take p_1 and p_2 that attain the minimum distance E. Then also $\frac{p_1+p_2}{2}$ attains the minimum and we denote it by p.

A sketch of de la Vallée Poussin's proof

Take p_1 and p_2 that attain the minimum distance E. Then also $\frac{p_1+p_2}{2}$ attains the minimum and we denote it by p. By a result called the *alternation theorem*, we have that there is a $j \in \{0,1\}$ and $x_1 < \ldots < x_{n+1}$ in $[0,1]$ such that for every $i \in \{1,\ldots,n+1\}$,

$$(p-f)(x_i) = (-1)^{i+j}E.$$

A sketch of de la Vallée Poussin's proof

Take p_1 and p_2 that attain the minimum distance E. Then also $\frac{p_1+p_2}{2}$ attains the minimum and we denote it by p. By a result called the *alternation theorem*, we have that there is a $j \in \{0,1\}$ and $x_1 < \ldots < x_{n+1}$ in $[0,1]$ such that for every $i \in \{1,\ldots,n+1\}$,

$$(p-f)(x_i) = (-1)^{i+j}E.$$

Let $i \in \{1,\ldots,n+1\}$ and assume wlog that $i+j$ is even. Then $(p-f)(x_i) = E$, so

$$\frac{p_1(x_i)-f(x_i)}{2} + \frac{p_2(x_i)-f(x_i)}{2} = E.$$

A sketch of de la Vallée Poussin's proof

Take p_1 and p_2 that attain the minimum distance E. Then also $\frac{p_1+p_2}{2}$ attains the minimum and we denote it by p. By a result called the *alternation theorem*, we have that there is a $j \in \{0,1\}$ and $x_1 < \ldots < x_{n+1}$ in $[0,1]$ such that for every $i \in \{1,\ldots,n+1\}$,

$$(p-f)(x_i) = (-1)^{i+j}E.$$

Let $i \in \{1,\ldots,n+1\}$ and assume wlog that $i+j$ is even. Then $(p-f)(x_i) = E$, so

$$\frac{p_1(x_i)-f(x_i)}{2} + \frac{p_2(x_i)-f(x_i)}{2} = E.$$

Since $\|p_1 - f\| = E$, $p_1(x_i) - f(x_i) \leq E$. Similarly, $p_2(x_i) - f(x_i) \leq E$. By the above, we have that both are actually equal to E and so $p_1(x_i) = p_2(x_i)$.

A sketch of de la Vallée Poussin's proof

Take p_1 and p_2 that attain the minimum distance E. Then also $\frac{p_1+p_2}{2}$ attains the minimum and we denote it by p. By a result called the *alternation theorem*, we have that there is a $j \in \{0,1\}$ and $x_1 < \ldots < x_{n+1}$ in $[0,1]$ such that for every $i \in \{1,\ldots,n+1\}$,

$$(p-f)(x_i) = (-1)^{i+j}E.$$

Let $i \in \{1,\ldots,n+1\}$ and assume wlog that $i+j$ is even. Then $(p-f)(x_i) = E$, so

$$\frac{p_1(x_i)-f(x_i)}{2} + \frac{p_2(x_i)-f(x_i)}{2} = E.$$

Since $\|p_1 - f\| = E$, $p_1(x_i) - f(x_i) \leq E$. Similarly, $p_2(x_i) - f(x_i) \leq E$. By the above, we have that both are actually equal to E and so $p_1(x_i) = p_2(x_i)$. Since p_1 and p_2 coincide on at least $n+1$ points, they must be equal.

Approximating the proof

Let us now see how one approximates the proof on the previous slide. First, for trivial reasons, the polynomials can be assumed to be in the closed ball Z of radius $\frac{5}{2}\|f\|$ (which is compact, as it lies inside the finite dimensional space P_n).

Approximating the proof

Let us now see how one approximates the proof on the previous slide. First, for trivial reasons, the polynomials can be assumed to be in the closed ball Z of radius $\frac{5}{2}\|f\|$ (which is compact, as it lies inside the finite dimensional space P_n).

1. for all p_1, $p_2 \in Z$ and all $\varepsilon > 0$, if $\|f - p_1\|$, $\|f - p_2\| \leq E + \Phi_1(\varepsilon)$, then $\left\|f - \frac{p_1+p_2}{2}\right\| \leq E + \varepsilon$.

Approximating the proof

Let us now see how one approximates the proof on the previous slide. First, for trivial reasons, the polynomials can be assumed to be in the closed ball Z of radius $\frac{5}{2}\|f\|$ (which is compact, as it lies inside the finite dimensional space P_n).

1. for all p_1, $p_2 \in Z$ and all $\varepsilon > 0$, if $\|f - p_1\|$, $\|f - p_2\| \leq E + \Phi_1(\varepsilon)$, then $\left\|f - \frac{p_1+p_2}{2}\right\| \leq E + \varepsilon$.
2. (the "ε-alternation theorem") for all $p \in Z$ and all $\varepsilon > 0$ with $\|f - p\| \leq E + \Phi_2(\varepsilon)$ there is a $j \in \{0, 1\}$ and $x_1 < \ldots < x_{n+1}$ in $[0, 1]$ such that for every $i \in \{1, \ldots, n+1\}$,

$$|(p - f)(x_i) - (-1)^{i+j} E| \leq \varepsilon.$$

Last steps

I shall omit steps 3 and 4, as I am not going to focus on them.

Last steps

I shall omit steps 3 and 4, as I am not going to focus on them.

5. for all p_1, $p_2 \in Z$ and all δ, $\beta > 0$, $x_1 < \ldots < x_{n+1}$ in $[0, 1]$ such that for all $i \in \{1, \ldots, n\}$, $x_{i+1} - x_i \geq \beta$ and for all $i \in \{1, \ldots, n+1\}$, $|(p_1 - p_2)(x_i)| \leq \Phi_5(\beta, \delta)$, we have that $\|p_1 - p_2\| \leq \delta$.

Last steps

I shall omit steps 3 and 4, as I am not going to focus on them.

5. for all p_1, $p_2 \in Z$ and all δ, $\beta > 0$, $x_1 < \ldots < x_{n+1}$ in $[0,1]$ such that for all $i \in \{1, \ldots, n\}$, $x_{i+1} - x_i \geq \beta$ and for all $i \in \{1, \ldots, n+1\}$, $|(p_1 - p_2)(x_i)| \leq \Phi_5(\beta, \delta)$, we have that $\|p_1 - p_2\| \leq \delta$.

Kohlenbach has extracted moduli Φ_1-Φ_5 and by putting them together he obtained the modulus of uniqueness. This was possible, by the metatheorems of proof mining, because the uniqueness proof could be formalized in WE-PA$^\omega$+WKL+QF-AC^{0,0}.

Directions to follow

Kohlenbach also suggested in his 1990 thesis to extend the techniques to the following results:

Directions to follow

Kohlenbach also suggested in his 1990 thesis to extend the techniques to the following results:

- L_1-best approximation: analyzed by K. and Paulo Oliva in the early 2000s

Directions to follow

Kohlenbach also suggested in his 1990 thesis to extend the techniques to the following results:

- L_1-best approximation: analyzed by K. and Paulo Oliva in the early 2000s
- Chebyshev approximation with bounded coefficients
    - a 1971 result of Roulier and Taylor

Directions to follow

Kohlenbach also suggested in his 1990 thesis to extend the techniques to the following results:

- L_1-best approximation: analyzed by K. and Paulo Oliva in the early 2000s
- Chebyshev approximation with bounded coefficients
  - a 1971 result of Roulier and Taylor
  - its analysis stood for 30 years as an open problem in proof mining

Directions to follow

Kohlenbach also suggested in his 1990 thesis to extend the techniques to the following results:

- L_1-best approximation: analyzed by K. and Paulo Oliva in the early 2000s
- Chebyshev approximation with bounded coefficients
  - a 1971 result of Roulier and Taylor
  - its analysis stood for 30 years as an open problem in proof mining

The last one is what we are going to focus on.

The result

Theorem (Roulier and Taylor, 1971)

Let $n, m \in \mathbb{N}$ be such that $m \leq n$ and $(k_i)_{i=1}^m \subseteq \mathbb{N}$ be such that $0 < k_1 < \ldots < k_m \leq n$. In addition, let $(a_i)_{i=1}^m$ and $(b_i)_{i=1}^m$ be finite sequences in $\mathbb{R} \cup \{\pm\infty\}$ be such that for all $i \in \{1, \ldots, m\}$, $a_i \leq b_i$, $a_i \neq \infty$ and $b_i \neq -\infty$. If one sets

$$K := \left\{ \sum_{i=0}^{n} c_i X^i \in P_n \;\middle|\; \text{ for all } i \in \{1, \ldots, m\},\; a_i \leq c_{k_i} \leq b_i \right\},$$

then for any continuous $f : [0,1] \to \mathbb{R}$ there is a unique $p \in K$ such that

$$\|f - p\| = \min_{q \in K} \|f - q\|.$$

The result

Theorem (Roulier and Taylor, 1971)

Let n, $m \in \mathbb{N}$ be such that $m \leq n$ and $(k_i)_{i=1}^m \subseteq \mathbb{N}$ be such that $0 < k_1 < \ldots < k_m \leq n$. In addition, let $(a_i)_{i=1}^m$ and $(b_i)_{i=1}^m$ be finite sequences in $\mathbb{R} \cup \{\pm\infty\}$ be such that for all $i \in \{1, \ldots, m\}$, $a_i \leq b_i$, $a_i \neq \infty$ and $b_i \neq -\infty$. If one sets

$$K := \left\{ \sum_{i=0}^{n} c_i X^i \in P_n \mid \text{ for all } i \in \{1, \ldots, m\},\ a_i \leq c_{k_i} \leq b_i \right\},$$

then for any continuous $f : [0,1] \to \mathbb{R}$ there is a unique $p \in K$ such that

$$\|f - p\| = \min_{q \in K} \|f - q\|.$$

The proof resembles the one from before, so we shall focus on the part which is **fundamentally** different.

The approximate form of the new proof

In the ε-alternation step one obtains (among others) an $r \leq n$, a sequence of degrees $n \geq d_1 > d_2 > \ldots > d_{r+1} = 0$ and $x_1 < \ldots < x_{r+1}$ in $[0, 1]$.

The approximate form of the new proof

In the ε-alternation step one obtains (among others) an $r \leq n$, a sequence of degrees $n \geq d_1 > d_2 > \ldots > d_{r+1} = 0$ and $x_1 < \ldots < x_{r+1}$ in $[0, 1]$.

In the last step we deal with the difference $p_1 - p_2$ as before, only we split it as $p_1 - p_2 = Q_1 + Q_2$ where Q_2 has only terms of degrees $d_1, \ldots, d_{r+1}$.

The approximate form of the new proof

In the ε-alternation step one obtains (among others) an $r \leq n$, a sequence of degrees $n \geq d_1 > d_2 > \ldots > d_{r+1} = 0$ and $x_1 < \ldots < x_{r+1}$ in $[0, 1]$.

In the last step we deal with the difference $p_1 - p_2$ as before, only we split it as $p_1 - p_2 = Q_1 + Q_2$ where Q_2 has only terms of degrees $d_1, \ldots, d_{r+1}$.

It is thus enough to show that for each i, $\|Q_i\| \leq \frac{\delta}{2}$.

The approximate form of the new proof

In the ε-alternation step one obtains (among others) an $r \leq n$, a sequence of degrees $n \geq d_1 > d_2 > \ldots > d_{r+1} = 0$ and $x_1 < \ldots < x_{r+1}$ in $[0, 1]$.

In the last step we deal with the difference $p_1 - p_2$ as before, only we split it as $p_1 - p_2 = Q_1 + Q_2$ where Q_2 has only terms of degrees $d_1, \ldots, d_{r+1}$.

It is thus enough to show that for each i, $\|Q_i\| \leq \frac{\delta}{2}$.

Q_1 is easily bounded by classical methods (using the way the d_i's were chosen).

The approximate form of the new proof

In the ε-alternation step one obtains (among others) an $r \leq n$, a sequence of degrees $n \geq d_1 > d_2 > \ldots > d_{r+1} = 0$ and $x_1 < \ldots < x_{r+1}$ in $[0, 1]$.

In the last step we deal with the difference $p_1 - p_2$ as before, only we split it as $p_1 - p_2 = Q_1 + Q_2$ where Q_2 has only terms of degrees $d_1, \ldots, d_{r+1}$.

It is thus enough to show that for each i, $\|Q_i\| \leq \frac{\delta}{2}$.

Q_1 is easily bounded by classical methods (using the way the d_i's were chosen).

For Q_2, one must generalize the proof of the original step 5.

Proof of the original step 5

Set $p := p_1 - p_2$. By the classical Lagrangian interpolation formula, we have that:

$$p = \sum_{j=1}^{n+1} \left(\prod_{i \neq j} \frac{X - x_i}{x_j - x_i} \right) \cdot p(x_j).$$

Proof of the original step 5

Set $p := p_1 - p_2$. By the classical Lagrangian interpolation formula, we have that:

$$p = \sum_{j=1}^{n+1} \left(\prod_{i \neq j} \frac{X - x_i}{x_j - x_i} \right) \cdot p(x_j).$$

Since we have, for all $x \in [0, 1]$,

$$\left| \prod_{i \neq j} \frac{X - x_i}{x_j - x_i} \right| \leq \frac{1}{\prod_{i \neq j} \beta |i - j|} \leq \frac{1}{\beta^n},$$

Proof of the original step 5

Set $p := p_1 - p_2$. By the classical Lagrangian interpolation formula, we have that:

$$p = \sum_{j=1}^{n+1} \left(\prod_{i \neq j} \frac{X - x_i}{x_j - x_i} \right) \cdot p(x_j).$$

Since we have, for all $x \in [0, 1]$,

$$\left| \prod_{i \neq j} \frac{X - x_i}{x_j - x_i} \right| \leq \frac{1}{\prod_{i \neq j} \beta |i - j|} \leq \frac{1}{\beta^n},$$

we get, for all $x \in [0, 1]$,

$$|p(x)| \leq \sum_{j=1}^{n+1} \left| \prod_{i \neq j} \frac{X - x_i}{x_j - x_i} \right| \cdot |p(x_j)| \leq (n+1) \cdot \frac{1}{\beta^n} \cdot \Phi_5(\beta, \delta).$$

Proof of the original step 5

Set $p := p_1 - p_2$. By the classical Lagrangian interpolation formula, we have that:

$$p = \sum_{j=1}^{n+1} \left(\prod_{i \neq j} \frac{X - x_i}{x_j - x_i} \right) \cdot p(x_j).$$

Since we have, for all $x \in [0,1]$,

$$\left| \prod_{i \neq j} \frac{X - x_i}{x_j - x_i} \right| \leq \frac{1}{\prod_{i \neq j} \beta |i - j|} \leq \frac{1}{\beta^n},$$

we get, for all $x \in [0,1]$,

$$|p(x)| \leq \sum_{j=1}^{n+1} \left| \prod_{i \neq j} \frac{X - x_i}{x_j - x_i} \right| \cdot |p(x_j)| \leq (n+1) \cdot \frac{1}{\beta^n} \cdot \Phi_5(\beta, \delta).$$

Since we want the right hand side to be smaller or equal to δ, one may take $\Phi_5(\beta, \delta) := \frac{\beta^n}{n+1} \cdot \delta$.

The lemma

Our new step 5 takes the form of the following lemma.

Lemma

Let n, $r \in \mathbb{N}$ with $r \leq n$ and $(d_i)_{i=1}^{r+1} \subseteq \mathbb{N}$ with $n \geq d_1 > d_2 > \ldots > d_{r+1} = 0$. Let β, $\delta > 0$ and $(x_j)_{j=1}^{r+1} \subseteq [0,1]$ such that for all $j \in \{1, \ldots, r\}$, $x_{j+1} - x_j \geq \beta$. Suppose that we have a polynomial

$$p = \sum_{i=1}^{r+1} \eta_i X^{d_i}$$

such that for all $j \in \{1, \ldots, r+1\}$,

$$|p(x_j)| \leq \widetilde{\Phi}_5(\beta, \delta).$$

Then $\|p\| \leq \delta$.

The lemma

Our new step 5 takes the form of the following lemma.

Lemma

Let n, $r \in \mathbb{N}$ with $r \leq n$ and $(d_i)_{i=1}^{r+1} \subseteq \mathbb{N}$ with $n \geq d_1 > d_2 > \ldots > d_{r+1} = 0$. Let β, $\delta > 0$ and $(x_j)_{j=1}^{r+1} \subseteq [0,1]$ such that for all $j \in \{1, \ldots, r\}$, $x_{j+1} - x_j \geq \beta$. Suppose that we have a polynomial

$$p = \sum_{i=1}^{r+1} \eta_i X^{d_i}$$

such that for all $j \in \{1, \ldots, r+1\}$,

$$|p(x_j)| \leq \widetilde{\Phi}_5(\beta, \delta).$$

Then $\|p\| \leq \delta$.

To obtain $\widetilde{\Phi}_5$, we need to generalize the Lagrangian formula.

Towards the new formula

Using the form of p in the lemma, we get that for all $j \in \{1, \dots, r+1\}$,

$$p(x_j) = \sum_{i=1}^{r+1} \eta_i x_j^{d_i}.$$

Towards the new formula

Using the form of p in the lemma, we get that for all $j \in \{1, \ldots, r+1\}$,

$$p(x_j) = \sum_{i=1}^{r+1} \eta_i x_j^{d_i}.$$

Therefore, we have

$$\begin{pmatrix} p \\ p(x_1) \\ \vdots \\ p(x_{r+1}) \end{pmatrix} = \sum_{i=1}^{r+1} \eta_i \begin{pmatrix} X^{d_i} \\ x_1^{d_i} \\ \vdots \\ x_{r+1}^{d_i} \end{pmatrix},$$

Towards the new formula

Using the form of p in the lemma, we get that for all $j \in \{1, \dots, r+1\}$,

$$p(x_j) = \sum_{i=1}^{r+1} \eta_i x_j^{d_i}.$$

Therefore, we have

$$\begin{pmatrix} p \\ p(x_1) \\ \vdots \\ p(x_{r+1}) \end{pmatrix} = \sum_{i=1}^{r+1} \eta_i \begin{pmatrix} X^{d_i} \\ x_1^{d_i} \\ \vdots \\ x_{r+1}^{d_i} \end{pmatrix},$$

so

$$\begin{vmatrix} p & X^{d_1} & \cdots & X^{d_{r+1}} \\ p(x_1) & x_1^{d_1} & \cdots & x_1^{d_{r+1}} \\ \vdots & \vdots & \ddots & \vdots \\ p(x_{r+1}) & x_{r+1}^{d_1} & \cdots & x_{r+1}^{d_{r+1}} \end{vmatrix} = 0.$$

Towards the new formula

Using the form of p in the lemma, we get that for all $j \in \{1, \dots, r+1\}$,

$$p(x_j) = \sum_{i=1}^{r+1} \eta_i x_j^{d_i}.$$

Therefore, we have

$$\begin{pmatrix} p \\ p(x_1) \\ \vdots \\ p(x_{r+1}) \end{pmatrix} = \sum_{i=1}^{r+1} \eta_i \begin{pmatrix} X^{d_i} \\ x_1^{d_i} \\ \vdots \\ x_{r+1}^{d_i} \end{pmatrix},$$

so

$$\begin{vmatrix} p & X^{d_1} & \cdots & X^{d_{r+1}} \\ p(x_1) & x_1^{d_1} & \cdots & x_1^{d_{r+1}} \\ \vdots & \vdots & \ddots & \vdots \\ p(x_{r+1}) & x_{r+1}^{d_1} & \cdots & x_{r+1}^{d_{r+1}} \end{vmatrix} = 0.$$

We are thus led to use Vandermonde determinants.

Generalizing Vandermonde

Remember the ordinary Vandermonde determinant:

$$V(y_1,\ldots,y_{r+1}) := \begin{vmatrix} y_1^r & y_1^{r-1} & \cdots & 1 \\ y_2^r & y_2^{r-1} & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ y_{r+1}^r & y_{r+1}^{r-1} & \cdots & 1 \end{vmatrix} = \prod_{1\le i<j\le r+1} (y_i - y_j).$$

Generalizing Vandermonde

Remember the ordinary Vandermonde determinant:

$$V(y_1, \ldots, y_{r+1}) := \begin{vmatrix} y_1^r & y_1^{r-1} & \cdots & 1 \\ y_2^r & y_2^{r-1} & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ y_{r+1}^r & y_{r+1}^{r-1} & \cdots & 1 \end{vmatrix} = \prod_{1 \le i < j \le r+1} (y_i - y_j).$$

Now define the following generalization (where $h_1 > \ldots > h_{r+1}$):

$$V(h_1, \ldots, h_{r+1}; y_1, \ldots, y_{r+1}) := \begin{vmatrix} y_1^{h_1} & y_1^{h_2} & \cdots & y_1^{h_{r+1}} \\ y_2^{h_1} & y_2^{h_2} & \cdots & y_2^{h_{r+1}} \\ \vdots & \vdots & \ddots & \vdots \\ y_{r+1}^{h_1} & y_{r+1}^{h_2} & \cdots & y_{r+1}^{h_{r+1}} \end{vmatrix}.$$

Generalizing Vandermonde

Remember the ordinary Vandermonde determinant:

$$V(y_1, \ldots, y_{r+1}) := \begin{vmatrix} y_1^r & y_1^{r-1} & \cdots & 1 \\ y_2^r & y_2^{r-1} & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ y_{r+1}^r & y_{r+1}^{r-1} & \cdots & 1 \end{vmatrix} = \prod_{1 \leq i < j \leq r+1} (y_i - y_j).$$

Now define the following generalization (where $h_1 > \ldots > h_{r+1}$):

$$V(h_1, \ldots, h_{r+1}; y_1, \ldots, y_{r+1}) := \begin{vmatrix} y_1^{h_1} & y_1^{h_2} & \cdots & y_1^{h_{r+1}} \\ y_2^{h_1} & y_2^{h_2} & \cdots & y_2^{h_{r+1}} \\ \vdots & \vdots & \ddots & \vdots \\ y_{r+1}^{h_1} & y_{r+1}^{h_2} & \cdots & y_{r+1}^{h_{r+1}} \end{vmatrix}.$$

Armed with these notations, by expanding the determinant on the previous slide along its first column, we get that

$$p = \sum_{j=1}^{r+1} (-1)^{j-1} \frac{V(d_1, \ldots, d_{r+1}; X, x_1, \ldots, \widehat{x_j}, \ldots, x_{r+1})}{V(d_1, \ldots, d_{r+1}; x_1, \ldots, x_{r+1})} \cdot p(x_j).$$

Young tableaux

We shall need some definitions from algebraic combinatorics to help us in dealing with those determinants.

- **partition**: a finite sequence $(\lambda_i)_{i=1}^{r+1} \subseteq \mathbb{N}$ with $\lambda_1 \geq \ldots \geq \lambda_{r+1}$

Young tableaux

We shall need some definitions from algebraic combinatorics to help us in dealing with those determinants.

- **partition**: a finite sequence $(\lambda_i)_{i=1}^{r+1} \subseteq \mathbb{N}$ with $\lambda_1 \geq \ldots \geq \lambda_{r+1}$
- we can move bijectively between strictly decreasing sequences h and partitions λ by the formula $\lambda_i^h := h_i + i - r - 1$

Young tableaux

We shall need some definitions from algebraic combinatorics to help us in dealing with those determinants.

- **partition**: a finite sequence $(\lambda_i)_{i=1}^{r+1} \subseteq \mathbb{N}$ with $\lambda_1 \geq \ldots \geq \lambda_{r+1}$
- we can move bijectively between strictly decreasing sequences h and partitions λ by the formula $\lambda_i^h := h_i + i - r - 1$
- if $r \in \mathbb{N}$ and λ is a partition of length $r+1$, then a **semistandard Young tableau** of weight λ is a jagged array with $r+1$ rows where for any $i \in \{1, \ldots, r+1\}$, the i'th line has λ_i entries which are elements of the set $\{1, \ldots, r+1\}$, such that the entries on each row are (weakly) increasing and the entries on each column are strictly increasing

1	1	2	7	8
2	3	3		
4	4			
5	6			
6				

Schur functions

Now, if T is such a semistandard Young tableau in which for each $i \in \{1, \dots, r+1\}$, i appears t_i times in T, one denotes by y^T the monomial $y_1^{t_1} \dots y_{r+1}^{t_{r+1}}$. Then the **Schur function** associated to a partition λ is defined by

$$s_\lambda := \sum_T y^T,$$

where T ranges over all semistandard Young tableaux of weight λ.

Schur functions

Now, if T is such a semistandard Young tableau in which for each $i \in \{1, \ldots, r+1\}$, i appears t_i times in T, one denotes by y^T the monomial $y_1^{t_1} \ldots y_{r+1}^{t_{r+1}}$. Then the **Schur function** associated to a partition λ is defined by

$$s_\lambda := \sum_T y^T,$$

where T ranges over all semistandard Young tableaux of weight λ.

The result which is relevant to our ends states that for any r and any strictly decreasing h of length $r+1$,

$$V(h_1, \ldots, h_{r+1}; y_1, \ldots, y_{r+1}) = V(y_1, \ldots, y_{r+1}) \cdot s_{\lambda^h}(y_1, \ldots, y_{r+1}).$$

Schur functions

Now, if T is such a semistandard Young tableau in which for each $i \in \{1, \dots, r+1\}$, i appears t_i times in T, one denotes by y^T the monomial $y_1^{t_1} \dots y_{r+1}^{t_{r+1}}$. Then the **Schur function** associated to a partition λ is defined by

$$s_\lambda := \sum_T y^T,$$

where T ranges over all semistandard Young tableaux of weight λ.

The result which is relevant to our ends states that for any r and any strictly decreasing h of length $r+1$,

$$V(h_1, \dots, h_{r+1}; y_1, \dots, y_{r+1}) = V(y_1, \dots, y_{r+1}) \cdot s_{\lambda^h}(y_1, \dots, y_{r+1}).$$

A simple proof may be found in:
R. A. Proctor, Equivalence of the combinatorial and the classical definitions of Schur functions. *J. Combin. Theory Ser. A* 51, no. 1, 135–137, 1989.

Back to the formula

The formula for p now becomes

$$p = \sum_{j=1}^{n+1} \left(\prod_{i \neq j} \frac{X - x_i}{x_j - x_i} \right) \cdot p(x_j) \cdot \frac{s_{\lambda^d}(X, x_1, \dots, \widehat{x_j}, \dots, x_{r+1})}{s_{\lambda^d}(x_1, \dots, x_{r+1})}.$$

This formula differs from the Lagrangian one only by the additional Schur factors, so we only need to bound those in order to get $\widetilde{\Phi}_5$.

The upper bound

For any partition λ of length $r+1$, the number of semistandard Young tableaux of weight λ can be shown to be

$$N_\lambda := \prod_{1 \le i < j \le r+1} \frac{\lambda_i - \lambda_j + j - i}{j - i}.$$

The upper bound

For any partition λ of length $r+1$, the number of semistandard Young tableaux of weight λ can be shown to be

$$N_\lambda := \prod_{1\le i<j\le r+1} \frac{\lambda_i - \lambda_j + j - i}{j-i}.$$

Moreover, for any n there is a finite number of strictly decreasing h's with length smaller or equal to $n+1$ and with $h_1 \le n$. If we set, for any n, N_n to be the maximum of all the N_{λ^h}'s for all these h's, this number is easily seen to be computable.

The upper bound

For any partition λ of length $r+1$, the number of semistandard Young tableaux of weight λ can be shown to be

$$N_\lambda := \prod_{1 \leq i < j \leq r+1} \frac{\lambda_i - \lambda_j + j - i}{j - i}.$$

Moreover, for any n there is a finite number of strictly decreasing h's with length smaller or equal to $n+1$ and with $h_1 \leq n$. If we set, for any n, N_n to be the maximum of all the N_{λ^h}'s for all these h's, this number is easily seen to be computable.

Proposition

For all n, $r \in \mathbb{N}$ with $r \leq n$, any strictly decreasing h of length $r+1$ and with $h_1 \leq n$, and any $y_1, \ldots, y_{r+1} \in [0,1]$,

$$0 \leq s_{\lambda^h}(y_1, \ldots, y_{r+1}) \leq N_n.$$

The lower bound

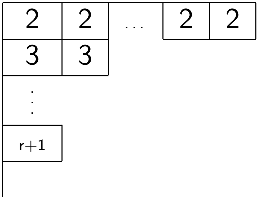
First, for all $j \in \{2, \ldots, n\}$, we have that

$$1 \geq x_k \geq x_2 \geq x_2 - x_1 \geq \beta.$$

First, for all $j \in \{2, \ldots, n\}$, we have that

$$1 \geq x_k \geq x_2 \geq x_2 - x_1 \geq \beta.$$

Since $d_{r+1} = 0$, $\lambda^d_{r+1} = 0$, so using the following semistandard Young tableau of weight λ^d:



The lower bound

First, for all $j \in \{2, \ldots, n\}$, we have that

$$1 \geq x_k \geq x_2 \geq x_2 - x_1 \geq \beta.$$

Since $d_{r+1} = 0$, $\lambda^d_{r+1} = 0$, so using the following semistandard Young tableau of weight λ^d:

$$\begin{array}{|c|c|c|c|c|}
\hline
2 & 2 & \ldots & 2 & 2 \\
\hline
3 & 3 & & & \\
\cline{1-2}
\vdots & & & & \\
\cline{1-1}
\text{r+1} & & & & \\
\cline{1-1}
\end{array}$$

we get that

$$\begin{aligned} s_{\lambda^d}(x_1, \ldots, x_{r+1}) &\geq x_2^{\lambda^d_1} \ldots x_{r+1}^{\lambda^d_r} \geq \beta^{\sum_{i=1}^{r} \lambda^d_i} \\ &\geq \beta^{r \cdot \lambda^d_1} = \beta^{r(d_1 - r)} \geq \beta^{r(n-r)} \geq \beta^{\frac{n^2}{4}}. \end{aligned}$$

Wrapping up

We may take then

$$\widetilde{\Phi}_5(\beta, \delta) := \frac{\beta^{n+\frac{n^2}{4}}}{N_n(n+1)} \cdot \delta.$$

Wrapping up

We may take then

$$\widetilde{\Phi}_5(\beta, \delta) := \frac{\beta^{n+\frac{n^2}{4}}}{N_n(n+1)} \cdot \delta.$$

The lower bound also shows that the "Lagrange-Schur" formula for p is well-defined, i.e. that the denominator is nonzero.

Wrapping up

We may take then

$$\widetilde{\Phi}_5(\beta,\delta) := \frac{\beta^{n+\frac{n^2}{4}}}{N_n(n+1)} \cdot \delta.$$

The lower bound also shows that the "Lagrange-Schur" formula for p is well-defined, i.e. that the denominator is nonzero.

In addition, like with the original Lagrange formula, we may also show the existence of an interpolation polynomial with prescribed degrees, by reversing the above argument (there is a catch, but it is easily taken care of).

The final modulus

Of course, there is much more to the extraction of the modulus. For example, the Schur formula also plays a role in the corresponding ε-alternation result.

The final modulus

Of course, there is much more to the extraction of the modulus. For example, the Schur formula also plays a role in the corresponding ε-alternation result. In the end, we get the modulus of uniqueness

$$\Psi(\delta) := \frac{\left(\frac{\chi_{\omega,n,M}\left(\frac{L}{2}\right)}{2}\right)^{\frac{n^2}{2}+2n}}{10 \cdot N_n^2 (n+1)(nF_n+1)} \cdot \delta,$$

which depends (in addition to δ) on

- the norm of a polynomial p_0 in K;
- the degree n;
- a lower bound L on E;
- a modulus of uniform continuity ω for f;
- the norm of f.

Some remarks

- The modulus does not depend on the bounds on the coefficients, except via p_0, which is in line with what Kohlenbach's metatheorems predict.

Some remarks

- The modulus does not depend on the bounds on the coefficients, except via p_0, which is in line with what Kohlenbach's metatheorems predict.
- The dependence on the norm of f may be removed at virtually no cost, by a shifting trick.

Some remarks

- The modulus does not depend on the bounds on the coefficients, except via p_0, which is in line with what Kohlenbach's metatheorems predict.
- The dependence on the norm of f may be removed at virtually no cost, by a shifting trick.
- The fact that the modulus is linear in δ corresponds to its coefficient being what approximation theorists call a **constant of strong unicity**, the existence of which having been shown before in this setting only nonconstructively.

Some remarks

- The modulus does not depend on the bounds on the coefficients, except via p_0, which is in line with what Kohlenbach's metatheorems predict.
- The dependence on the norm of f may be removed at virtually no cost, by a shifting trick.
- The fact that the modulus is linear in δ corresponds to its coefficient being what approximation theorists call a **constant of strong unicity**, the existence of which having been shown before in this setting only nonconstructively.
- One may even remove the dependence on L, though at the expense of linearity.

All this can be found in:

A. Sipoș, Bounds on strong unicity for Chebyshev approximation with bounded coefficients. arXiv:1904.10284 [math.CA], 2019.

Thank you for your attention.