Making polynomials robust to noise

A. A. Sherstov

Goal. We aim to approximate Boolean functions by real polynomials in a *robust* way. By D we denote the discrete set $\{-1, +1\}$, by I the interval $[-1, +1]$. Let f be a Boolean function $D^n \to D$, p a polynomial $I^n \to \mathbf{R}$ and $\varepsilon \in \mathbf{R}$ a constant, we say that

- p is an ε-approximation of f if $|f(x) - p(x)| \leq \varepsilon$ for every $x \in D^n$,
- p is a robust ε-approximation of f if $|f(x) - p(x + \delta)| \leq \varepsilon$ for every $x \in D^n$ and every $\delta \in (I/3)^n$.

The main result can be stated compactly as follows:

Theorem 1. *Let $p\colon D^n \to I$ be a polynomial. Then for every $\varepsilon > 0$ there is a polynomial p' of degree $O(\deg p + \log \frac{1}{\varepsilon})$ that robustly ε-approximates p.*

Tools. For functions from a set X to $\mathbf{R}$ we use the notation $\|\phi\|_\infty = \sup_{x\in X}|\phi(x)|$ and $\|\phi\|_1 = \sum_{x\in X}|\phi(x)|$. We use the *characteristic sign functions*, that is, for $S \subseteq \{1..n\}$ we have $\chi_S\colon D^n \to D$ defined as

$$\chi_S(x) = \prod_{i\in S} x_i.$$

These functions form an orthogonal basis for the vector space of the functions $D^n \to \mathbf{R}$ (with the scalar product $\langle f, g\rangle = 2^{-n}\sum_{x\in D^n} f(x)g(x)$). Thus every function can be represented as a linear combination of its *Fourier characters*, that is,

$$\phi = \sum_{S\subseteq\{1..n\}} \hat{\phi}(S)\chi_S,$$

where $\hat{\phi}(S) = \langle \phi, \chi_S\rangle$. This provides a multilinear extension to $\mathbf{R}^n \to \mathbf{R}$.

Battle plan. We gradually construct robust approximations for the following classes of polynomials:

1. Parity polynomials, $p(x) = \prod_{i\in\{1..n\}} x_i$.
2. Homogeneous polynomials, $p(x) = \sum_{|S|=d} a_S \prod_{i\in S} x_i$.
3. General polynomials.

Take a deep breath, the next page is tough.

Parity polynomials, $p(x) = \prod_{i\in\{1..n\}} x_i$.

Lemma 1. *For* $x_{1..n} \in (-\sqrt{2}, 0) \cup (0, \sqrt{2})$*,*

$$\operatorname{sgn}(x_1 \cdot x_2 \cdots x_n) = x_1 \cdot x_2 \cdots x_n \sum_{i_1, i_2, \dots, i_n \in \mathbf{N}} \prod_{j=1}^{n} \left(-\frac{1}{4}\right)^{i_j} \binom{2i_j}{i_j} (x_j^2 - 1)^{i_j}.$$

Theorem 2. *Fix* $\varepsilon \in [0, 1)$ *and let* $X = [-\sqrt{1+\varepsilon}, -\sqrt{1-\varepsilon}] \cup [\sqrt{1-\varepsilon}, \sqrt{1+\varepsilon}]$*. Then for every natural number* N *there is an explicitly given polynomial* $p\colon \mathbf{R}^n \to \mathbf{R}$ *of degree at most* $2N + n$ *such that*

$$\max_{X^n} |\operatorname{sgn} x_1 \cdot x_2 \cdots x_n - p(x)| \le \varepsilon^{N+1}(1+\varepsilon)^{n/2} \binom{N+n}{N+1} N.$$

Reduction from general case to homogeneous polynomials.

Lemma 2. *Let* $p(t) = \sum_{i=0}^{d} a_i t^i$ *be a given polynomial. Then for every* $i = 0, 1, \dots, d$ *we have*

$$|a_i| \le (4e)^d \max_{j=0,1,\dots,d} \left| p\left(\frac{j}{d}\right) \right|.$$

Theorem 3. *Let* $\phi\colon D^n \to \mathbf{R}$ *be a function,* $\deg \phi = d$*. Write* $\phi = \phi_0 + \phi_1 + \cdots + \phi_d$*, where* $\phi_i\colon D^n \to \mathbf{R}$ *is given by* $\phi_i = \sum_{|S|=i} \hat{\phi}(S) \chi_S$*. Then for* $i = 0, 1, \dots, d$ *we have*

$$\|\phi_i\|_\infty \le (4e)^d \|\phi\|_\infty.$$

Homogeneous polynomials, $p(x) = \sum_{|S|=d} a_S \prod_{i\in S} x_i$.

Abrahadabra. Let v be a vector in $\{0,1\}^d$, the operator A_v acting on functions from D^n to $\mathbf{R}$ is defined by

$$(A_v f)(x) = \mathbf{E}_{z\in D^d} \left\{ z_1 \cdot z_2 \cdots z_d f\left(\frac{1}{d}\sum_{i=1}^{d} z_i x_1^{v_i}, \frac{1}{d}\sum_{i=1}^{d} z_i x_2^{v_i}, \dots, \frac{1}{d}\sum_{i=1}^{d} z_i x_n^{v_i} \right) \right\}.$$

Theorem 4. *Let* $\phi\colon D^n \to \mathbf{R}$ *be given such that* $\hat{\phi}(S) = 0$ *whenever* $|S| \ne d$*. Fix arbitrary symmetric function* $\delta\colon D^n \to \mathbf{R}$ *and define* $\Delta\colon D^n \to \mathbf{R}$ *by*

$$\Delta(x) = \sum_{|S|=d} \hat{\phi}(S) \delta(x|S).$$

Then

$$\|\Delta\|_\infty \le \frac{d^d}{d!} \|\phi\|_\infty \|\hat{\delta}\|_1.$$

Theorem 5. *Let* $\phi\colon D^n \to \mathbf{R}$ *be given such that* $\hat{\phi}(S) = 0$ *whenever* $|S| \ne d$*. Fix* $\varepsilon \in [0, 1)$ *and let* X *be as in Lemma 1. Then for every natural number* M *there is an explicitly given polynomial* $p\colon \mathbf{R}^d \to \mathbf{R}$ *of degree at most* $2M + d$ *such that*

$$P(x) = \sum_{S \in \binom{n}{d}} \hat{\phi}(S) p(x|S)$$

obeys

$$\max_{X^n} |\phi(\operatorname{sgn} x_1, \operatorname{sgn} x_2, \dots, \operatorname{sgn} x_n) - P(x)| \le (1+\varepsilon)^{d/2} \frac{d^d}{d!} \varepsilon^M \binom{M+d}{M} M \|\phi\|_\infty.$$