3 Approximating many things by polynomials

In the previous lecture, we showed that we can get block-encodings of $p^{(\mathrm{SV})}(A)$ from block-encodings of A, provided that $p(x)$ is an even or odd polynomial such that $|p(x)| \leq 1$ for $x \in [-1, 1]$. Roughly, this turned a Q-block encoding to a dQ-block encoding. For applications of interest, the main goal is actually to apply a non-polynomial function; to capture these applications, we need tools for approximating the relevant functions with bounded polynomials. In this lecture, we introduce Chebyshev polynomials, our main tool for constructing such approximations. We will see that the class of low-degree bounded polynomials is expressive enough for many applications.

3.1 Chebyshev polynomials and properties

Chebyshev polynomials are a very common tool in applied math; we'll cover a small amount of the theory here.

Definition 3.1 (Chebyshev polynomial)**.** The degree-n *Chebyshev polynomial* (of the first kind), denoted $T_n(x)$, is the function that satisfies, for all $z \in \mathbb{C}$,

$$T_n(\tfrac{1}{2}(z + z^{-1})) = \tfrac{1}{2}(z^n + z^{-n}). \tag{1}$$

We can see this is a polynomial by verifying that T_n satisfies the recurrence

$$T_n = 2x \cdot T_{n-1} - T_{n-2},$$

with $T_0 = 1$ and $T_1 = x$. Plugging in $z = \exp(\mathrm{i}\theta)$ for $\theta \in [-\pi, \pi]$, we get another familiar definition of the Chebyshev polynomials,

$$T_n(\cos(\theta)) = \cos(k\theta).$$

From these definitions we have that $\|T_n(x)\|_{[-1,1]} \leq 1$, and that T_n has the same parity as n, i.e. $T_n(-x) = (-1)^n T_n(x)$.

Under mild "niceness" conditions, any function can be written as a series of Chebyshev polynomials $f(x) = \sum_{k \geq 0} a_k T_k(x)$.

Lemma 3.2 ([Tre19, Theorem 3.1])**.** *Let $f : [-1, 1] \to \mathbb{R}$ be Lipschitz (i.e. $|f(x) - f(y)| \leq C|x - y|$ for finite C). Then f has a unique decomposition into Chebyshev polynomials*

$$f(x) = \sum_{k=0}^{\infty} a_k T_k(x),$$

where the Chebyshev coefficients a_k *absolutely converge.*

This is true for the same reason functions have Fourier series. In fact, the theory of Chebyshev polynomials is a parallel theory. For $z = e^{\mathrm{i}\theta}$, define $g(z) = f(\frac{1}{2}(z + z^{-1}))$. Then $g(z)$ is a function on the unit circle with a Laurent series.

$$g(z) = \sum_k a_k T_k(\frac{1}{2}(z + z^{-1})) = \sum_k \frac{a_k}{2}(z^k + z^{-k}).$$

For $\theta \in [-\pi, \pi]$, define $h(\theta) = g(e^{\mathrm{i}\theta}) = f(\frac{1}{2}(e^{\mathrm{i}\theta} + e^{-\mathrm{i}\theta}))$. Then $h(\theta)$ is a 2π-periodic function with a Fourier series.

$$h(\theta) = \sum_k a_k T_k(\frac{1}{2}(e^{\mathrm{i}\theta} + e^{-\mathrm{i}\theta})) = \sum_k \frac{a_k}{2}(e^{\mathrm{i}k\theta} + e^{-\mathrm{i}k\theta}).$$

Lemma 3.3 (Orthogonality property). *$\{T_k\}_k$ are orthogonal under a particular choice of inner product.*

$$\frac{2}{\pi}\int_{-1}^{1}\frac{T_k(x)T_\ell(x)}{\sqrt{1-x^2}}\mathrm{d}x = \begin{cases} 1 & k=\ell \\ 0 & \textit{otherwise} \end{cases}$$

Proof. Substituting $x=\cos(\theta)$ and $\mathrm{d}x = -\sin(\theta)\mathrm{d}\theta = -\sqrt{1-x^2}\mathrm{d}\theta$,

$$\begin{aligned}\int_{-1}^{1}\frac{T_k(x)T_\ell(x)}{\sqrt{1-x^2}}\mathrm{d}x &= \int_{\pi}^{0} -T_k(\cos(\theta))T_\ell(\cos(\theta))\mathrm{d}z \\ &= \int_{0}^{\pi} -\cos(k\theta)\cos(\ell\theta)\mathrm{d}z = \begin{cases} \frac{\pi}{2} & k=\ell \\ 0 & \text{otherwise} \end{cases}\end{aligned}$$

(When $k=\ell=0$, we have π instead.) □

Lemma 3.4 (Chebyshev coefficients, [Tre19, Theorem 3.1]). *For $f(x)$ with a Chebyshev series $f(x)=\sum_{k\geq 0}a_kT_k(x)$, the Chebyshev coefficients can be computed with the integral*

$$a_k = \frac{2}{\pi}\int_{-1}^{1}\frac{f(x)T_k(x)}{\sqrt{1-x^2}}\mathrm{d}x. \tag{2}$$

For $k=0$ the same formula holds with the factor $2/\pi$ changed to $1/\pi$.

Lemma 3.5 (Chebyshev coefficients are bounded). *If $\|f(x)\|_{[-1,1]}\leq 1$, then*

$$\begin{aligned}|a_k| &= \frac{2}{\pi}\Big|\int_{-1}^{1}\frac{f(x)T_k(x)}{\sqrt{1-x^2}}\mathrm{d}x\Big| \\ &\leq \frac{2}{\pi}\int_{-1}^{1}\frac{\|f\|_{[-1,1]}\|T_k\|_{[-1,1]}}{\sqrt{1-x^2}}\mathrm{d}x \\ &\leq \frac{2}{\pi}\int_{-1}^{1}\frac{1}{\sqrt{1-x^2}}\mathrm{d}x \\ &\leq 2.\end{aligned}$$

This does not hold for polynomials in the monomial basis: $\|T_n\|_{[-1,1]}=1$, but the monomial of the leading coefficient is 2^{n-1}.

3.2 Approximating functions from Chebyshev series

Definition 3.6 (Chebyshev truncation). For a function $f:[-1,1]\to\mathbb{C}$ written as a Chebyshev series $f(x)=\sum_{k=0}^{\infty}a_kT_k(x)$, we denote the degree-$n$ *Chebyshev truncation* of f as

$$f_n(x) = \sum_{k=0}^{n}a_kT_k(x).$$

If the function f one wishes to approximate is standard, closed forms of the Chebyshev coefficients may be known, so one can take a Chebyshev truncation and explicitly bound the error:

$$\|f - f_n\|_{[-1,1]} = \Big\| \sum_{k=n+1}^{\infty} a_k T_k(x) \Big\|_{[-1,1]} \leq \sum_{k=n+1}^{\infty} |a_k| \|T_k(x)\|_{[-1,1]} = \sum_{k=n+1}^{\infty} |a_k|.$$

In other words, by choosing n such that the coefficient tail sum is bounded by ε, we obtain an ε-uniform approximation on $[-1, 1]$. If we had these Chebyshev coefficients explicitly, we could bound them directly to get a polynomial approximation. However, we may not know explicitly what the Chebyshev coefficients of our desired function is, so we can't easily bound them. The following shows that the Chebyshev coefficient tail is exponential, provided that the function is analytic around $[-1, 1]$.

Theorem 3.7 ([Tre19, Theorems 8.1 and 8.2])**.** *Let f be an analytic function in $[-1, 1]$ and analytically continuable to the interior of the Bernstein ellipse $E_\rho = \{\frac{1}{2}(z + z^{-1}) : |z| = \rho\}$, where it satisfies $|f(x)| \leq M$. Then its Chebyshev coefficients satisfy $|a_0| \leq M$ and $|a_k| \leq 2M\rho^{-k}$ for $k \geq 1$.*

Corollary 3.8. *Consequently, for each $n \geq 0$, its Chebyshev projections satisfy*

$$\|f - f_n\|_{[-1,1]} \leq \sum_{k \geq n+1} |a_k| \leq 2M \sum_{k \geq n+1} \rho^{-k} = \frac{2M\rho^{-n}}{\rho - 1},$$

and choosing $n = \lceil \frac{1}{\log(\rho)} \log \frac{2M}{(\rho-1)\varepsilon} \rceil$, we have $\|f - f_n\|_{[-1,1]} \leq \varepsilon$.

Proof. Recall from (2) (and since inverting z does not change the contour integral) that for $k \geq 1$,

$$a_k = \frac{1}{\pi \mathrm{i}} \int_{|z|=1} z^{-(k+1)} f(\tfrac{1}{2}(z + z^{-1})) \mathrm{d}z.$$

The boundary of E_ρ is given by $\frac{1}{2}(z + z^{-1})$ for $|z| = \rho$, and f is analytic in E_ρ, so we may choose a different contour without affecting the value of the integral:

$$a_k = \frac{1}{\pi \mathrm{i}} \int_{|z|=\rho} z^{-(k+1)} f(\tfrac{1}{2}(z + z^{-1})) \mathrm{d}z.$$

The conclusion follows from the facts that the circumference of $|z| = \rho$ is $2\pi\rho$ and the function is bounded by M. A similar argument gives the case $k = 0$, where (2) has $2\pi\mathrm{i}$ in the denominator. $\square$

Fact 3.9. *The Bernstein ellipse E_ρ for $\rho = 1 + \delta \leq 2$ satisfies*

$$\mathrm{interior}(E_\rho) \subset \Big\{ x + \mathrm{i}y \mid x, y \in \mathbb{R},\ |x| \leq 1 + \frac{\delta^2}{2} \text{ and } |y| \leq \delta \Big\}.$$

Corollary 3.10 (Application to Hamiltonian simulation)**.** *Consider the function* $\sin(tx)$*. Then for $z = a + \mathrm{i}b$ on the interior of the Bernstein ellipse E_ρ,*

$$\begin{aligned} |\sin(tz)| &\leq \tfrac{1}{2} |e^{\mathrm{i}tz} - e^{-\mathrm{i}tz}| \\ &\leq \tfrac{1}{2}(|e^{-bt}| + |e^{bt}|) \leq e^{|bt|}. \end{aligned}$$

So, choosing $\rho = 1 + 1/t$, we can apply the theorem with $M = \mathcal{O}(1)$. When $t \geq 1$, this gives an ε-good approximation for the Chebyshev truncation of degree

$$n = \mathcal{O}(t \log \frac{t}{\varepsilon}).$$

This isn't tight; the Chebyshev coefficients of $\sin(tx)$ actually decay super-exponentially [TT23].

Remark 3.11 (Chebyshev approximation vs Taylor series approximation). You might be wondering what the difference is between truncating a Chebyshev series and truncating a Taylor series, perhaps a more commonly known tool for polynomial approximation. In fact, one can make a similar statement to the above (see [GSLW19, Corollary 66]), but it becomes difficult to apply in settings where one does not know the Taylor series. Note that this does not always give the right polynomial approximation; for example, truncating the Taylor series of e^x does not give the optimal degree.

Theorem 3.7 shows that if one can analytically continue f to a Bernstein ellipse with $\rho = 1 + \alpha$ for small α, then a degree $\approx \frac{1}{\alpha}$ polynomial obtains good approximation error on $[-1, 1]$. Unfortunately, since the approximation in Theorem 3.7 is based on Chebyshev truncation, the approximation rapidly blows up outside the range $[-1, 1]$ (i.e. growing as $O(|x|^n)$ for x sufficiently outside $[-1, 1]$). In interesting applications of the QSVT framework, this is an obstacle. For example, to use QSVT for solving a system of linear equations, we need a polynomial approximation to x^{-1} on $[\delta, 1]$ that is bounded on $[-1, 1]$. Upon linearly remapping $[\delta, 1]$ to $[-1, 1]$, this corresponds to a bounded approximation on $[-b, 1]$ for some $b > 1$, so Chebyshev truncations give us a very poor degree of control.

Chebyshev truncation is not enough for our purposes, since our criteria is *different* from uniform approximation on $[-1, 1]$. For quantum linear systems, we require a polynomial approximation close to $1/x$ on $[-1, -1/\kappa] \cup [1/\kappa, 1]$, but it merely needs to be bounded on $[-1/\kappa, 1/\kappa]$. This bounded requirement is necessary for our block-encoding machineary.

As [GSLW19] points out, there are generic ways to find approximations to piecewise smooth functions which satisfy this sort of "ε-close on smooth pieces, but bounded near points of discontinuity" requirement, with $\log \frac{1}{\varepsilon}$ scaling in the degree.

Theorem 3.12 ([TT23, Theorem 19])**.** *Let f be an analytic function in $[-1, 1]$ and analytically continuable to the interior of E_ρ where $\rho = 1 + \alpha$, where it is bounded by M. For $\delta \in (0, \frac{1}{C}\min(1, \alpha^2))$ where C is a sufficiently large constant, $\varepsilon \in (0, 1)$, and $b > 1$, there is a polynomial q of degree $O(\frac{b}{\delta} \log \frac{b}{\delta\varepsilon})$ such that*

$$\begin{aligned} \|f - q\|_{[-1,1]} &\leq M\varepsilon, \\ \|q\|_{[-(1+\delta),1+\delta]} &\leq M, \\ \|q\|_{[-b,-(1+\delta)]\cup[1+\delta,b]} &\leq M\varepsilon. \end{aligned}$$

Proof sketch.

1. Applying Theorem 3.7 gives f_n of degree $n \approx \frac{1}{\alpha}$ approximating f in the interval $[-1, 1]$, but f_n does not satisfy the other required conclusions due to its growth outside $[-1, 1]$.

2. We multiply f_n by a "threshold" r based on the Gaussian error function erf, whose tails decay much faster than the Chebyshev polynomials grow outside $[-1, 1]$. Our function r has the property that inside $[-1, 1]$, it is close to 1, and outside $[-(1 + \delta), 1 + \delta]$, it is close to 0.

3. Using bounds on the growth of erf, we show $r \cdot f_n$ is bounded on a Bernstein ellipse of radius $1 + \frac{\delta}{b}$ appropriately rescaled, and applying Theorem 3.7 once more gives the conclusion.

The final proof requires some care to obtain the claimed scalings on the windows of approximation, but we include this tedium to make the theorem statement as simple to use as possible. $\square$

Corollary 3.13. *Let $\delta, \varepsilon \in (0,1)$, and let $f(x) = |\frac{\delta}{x}|$. There exist both even and odd polynomials $p(x)$ of degree $O(\frac{1}{\delta}\log\frac{1}{\delta\varepsilon})$ such that $\|p\|_{[-1,1]} \le 3$ and $\|p - f\|_{[\delta,1]} \le \varepsilon$.*

Proof. Assume δ is sufficiently small, else taking a smaller δ only affects the bound by a constant. We rescale the region of interest: $x = \frac{1-\delta}{2}y + \frac{1+\delta}{2}$ is in $[\delta, 1]$ for $y \in [-1,1]$, so let

$$g(y) := \delta\Big(\frac{1-\delta}{2}y + \frac{1+\delta}{2}\Big)^{-c}.$$

We require a bound of g on E_ρ for $\rho = 1 + \sqrt{\delta/4}$. Since f is largest closest to the origin, g is largest at the point closest to $-\frac{1+\delta}{1-\delta}$, i.e. $-\frac{1}{2}(\rho+\rho^{-1}) > -(1+\frac{\delta}{8})$ by Fact 3.9. Further,

$$\begin{aligned}
g(-\frac{1}{2}(\rho+\rho^{-1})) &\le g(-(1+\frac{\delta}{8})) \\
&\le \delta\Big(-\frac{1-\delta}{2}(1+\frac{\delta}{8}) + \frac{1+\delta}{2}\Big)^{-1} \\
&= (1 - \frac{1-\delta}{16})^{-1} \le \frac{3}{2}.
\end{aligned}$$

Let $\widetilde{\delta} = \frac{\delta}{4C}$ for sufficiently large C, and $b = 4$. Theorem 3.12 yields $q(y)$ satisfying:

$$\|q(y) - g(y)\|_{[-1,1]} \le \varepsilon,\ \|q(y)\|_{[-(1+\widetilde{\delta}),1+\widetilde{\delta}]} \le 2,\ \|q(y)\|_{[-4,-(1+\widetilde{\delta})]\cup[1+\widetilde{\delta},4]} \le \varepsilon.$$

Shifting back $y = \frac{2}{1-\delta}(x - \frac{1+\delta}{2})$, it is clear for sufficiently large C that $y = -\frac{1+3\delta}{1-\delta}$ (which corresponds to $x = -\delta$) has $y < -(1+\widetilde{\delta})$, and $y = -\frac{3+\delta}{1-\delta}$ (which corresponds to $x = -1$) has $y > -4$. So,

$$\begin{aligned}
\Big\|q(\frac{2}{1-\delta}(x - \frac{1+\delta}{2})) - f(x)\Big\|_{[\delta,1]} &\le \varepsilon, \\
\Big\|q(\frac{2}{1-\delta}(x - \frac{1+\delta}{2}))\Big\|_{[-\delta,\delta]} &\le 2, \\
\Big\|q(\frac{2}{1-\delta}(x - \frac{1+\delta}{2}))\Big\|_{[-1,-\delta]} &\le \varepsilon.
\end{aligned} \tag{3}$$

Depending on whether we wish the final function to be even or odd, we take

$$p(x) = q(\frac{2}{1-\delta}(x - \frac{1+\delta}{2})) \pm q(\frac{2}{1-\delta}(-x - \frac{1+\delta}{2})).$$

Then the guarantees of (3) give $\|p(x) - f(x)\|_{[\delta,1]} \le 2\varepsilon$ and $\|p(x)\|_{[-1,1]} \le 3$, and we rescale ε to conclude. The final degree of the polynomial is the degree of $q(y)$: $O(\frac{1}{\delta}\log\frac{1}{\delta\varepsilon})$. $\square$

3.3 Lower bounds on polynomial approximation

There are limitations to what kinds of functions can be approximated by low-degree polynomials. The most common lower bounds to keep in mind are the *Markov brothers'* and *Bernstein* inequalities:

Theorem 3.14 ([Sch41, Theorem 1])**.** *Let $p(x)$ be a degree n polynomial such that $\|p(x)\|_{[-1,1]} \leq 1$. Then*

$$\|p'(x)\|_{[-1,1]} \leq n^2 \tag{4}$$

Theorem 3.15 ([Sch41, Theorem 2])**.** *Let $p(x)$ be a degree n polynomial such that $\|p(x)\|_{[-1,1]} \leq 1$. Then*

$$\|p'(x)\|_{[-1,1]} \leq \frac{n}{\sqrt{1-x^2}}. \tag{5}$$

Remark 3.16. This means that a bounded polynomial has derivative $\mathcal{O}(n)$ near the center of $[-1,1]$, but can be $\mathcal{O}(n^2)$ near the edge. This suggests that functions can be approximated better when its worst-conditioned pieces are on the edges of $[-1,1]$. This is true: note that the above argument shows that sufficiently good bounded polynomial approximations to δ/x on $[-1,-\delta] \cup [\delta,1]$ must have degree $\Omega(1/\delta)$.

However, suppose we have a block-encoding of $I-A$, where A is a Hermitian matrix. That $\|I-A\| \leq 1$ implies that A is PSD, and suppose its eigenvalues are in $[\delta,2]$. We can then get an approximate block-encoding of A^{-1} via a polynomial approximation of

$$\frac{1}{1-x} \text{ for } x \in [-1+\delta, 1].$$

There is a $\mathcal{O}(\sqrt{\delta})$-degree polynomial approximation of this, implying that we can invert a matrix quadratically faster if we get this block-encoding [OD21].

$$\frac{1}{x-(1+\delta)} = \frac{-2}{\sqrt{2\delta-\delta^2}} \sum_{k=0}^{\infty\,\prime} (1+\delta-\sqrt{2\delta+\delta^2})^k T_k(x)$$

References

[GSLW19] András Gilyén, Yuan Su, Guang Hao Low, and Nathan Wiebe. "Quantum singular value transformation and beyond: Exponential improvements for quantum matrix arithmetics". In: *Proceedings of the 51st ACM Symposium on the Theory of Computing (STOC)*. ACM, June 2019, pp. 193–204. DOI: 10.1145/3313276.3316366. arXiv: 1806.01838 (page 4).

[OD21] Davide Orsucci and Vedran Dunjko. "On solving classes of positive-definite quantum linear systems with quadratically improved runtime in the condition number". In: *Quantum* 5 (Nov. 2021), p. 573. DOI: 10.22331/q-2021-11-08-573. arXiv: 2101.11868 [quant-ph] (page 6).

[Sch41] A. C. Schaeffer. "Inequalities of A. Markoff and S. Bernstein for polynomials and related functions". In: *Bull. Amer. Math. Soc.* 47 (1941), pp. 565–579. ISSN: 0002-9904. DOI: 10.1090/S0002-9904-1941-07510-5 (page 6).

[Tre19] Lloyd N. Trefethen. *Approximation theory and approximation practice, extended edition*. Extended edition [of 3012510]. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2019, pp. xi+363. ISBN: 978-1-611975-93-2. DOI: 10.1137/1.9781611975949 (pages 1–3).

[TT23] Ewin Tang and Kevin Tian. *A CS guide to the quantum singular value transformation*. 2023. arXiv: 2302.14324 [quant-ph] (page 4).