On approximating the inverse of a linear operator with Chebyshev polynomials

Zach Stoebner
Electrical & Computer Engineering, UT Austin
zstoebner@austin.utexas.edu

July 21, 2026

Abstract

These notes outline the steps and analysis for approximating the inverse of a linear operator with Chebyshev polynomials, with an example on Tikhonov-regularized image reconstruction. Code can be found here.

Notation Unless denoted otherwise, $||\cdot|| = ||\cdot||_2$. $\mathbb{F}$ denotes the field of numbers; the results hold for $\mathbb{R}$ and $\mathbb{C}$. Regular lowercase letters, e.g., x, denote scalars. Bolded lowercase letters, e.g., $\mathbf{x}$, denote vectors. Uppercase regular letters, e.g., X, denote matrices. Caligraphic uppercase letters, e.g., $\mathcal{X}$, denote tensors.

1 How to approximate the inverse

1.1 Normalize the operator A

Given $A \in \mathbb{S}^n_{++} \implies \exists A^{-1}$ with $\lambda_A \in [\lambda_-, \lambda_+],\ \lambda_- > 0$.
Let $\alpha = \frac{1}{2}(\lambda_+ - \lambda_-)$, $\beta = \frac{1}{2}(\lambda_+ + \lambda_-)$. Define $g : [\lambda_-, \lambda_+] \to [-1, 1]$:

$$g(\lambda) = \frac{2(\lambda - \lambda_-)}{\lambda_+ - \lambda_-} - 1 = \frac{2\lambda - (\lambda_+ + \lambda_-)}{\lambda_+ - \lambda_-} = \alpha^{-1}(\lambda - \beta)$$
$$\lambda(x) = g^{-1}(x) = \frac{1}{2}[x(\lambda_+ - \lambda_-) + (\lambda_+ + \lambda_-)] = \alpha x + \beta$$

where $\lambda(x)$ recovers the eigenvalues from the normalized operator. We map the spectrum to [-1,1] because Chebyshev polynomials are optimal in max absolute among all d-degree polynomials on this range. Since A is symmetric positive definite (SPD), we can exploit its spectral function property:

$$\begin{aligned} A \in \mathbb{S}^n_{++} &\implies A = U\Lambda U' \\ &\implies g(A) = \alpha^{-1}(A - \beta I) = U g(\Lambda) U' \end{aligned}$$

Let $\bar{A} = g(A) \implies A = \lambda(\bar{A})$. It follows, then that approximating the inverse reduces to univariate Chebyshev approximation on using $g(\bar{A})$.

1.2 Compute Chebyshev coefficients for A^{-1}

There are two objectives for approximating the inverse:

1. $$\min_{p_d} \max_{\lambda \in [\lambda_-, \lambda_+]} |\frac{1}{\lambda} - p_d(\lambda)| = |\frac{1 - \lambda p_d(\lambda)}{\lambda}| \tag{1}$$

2. $$\min_{p_d} \max_{\lambda \in [\lambda_-, \lambda_+]} |1 - \lambda p_d(\lambda)| \tag{2}$$

```
L-infinity optimized polynomial.
> Degree:   3
> Spectrum: [0.00, 93.35]
> Resulting polynomial: -1.68451184538492e-6*y**3 + 0.000314508385102814*y**2 - 0.0183504979262583*y + 0.342638203349746
```

Figure 1: Degree-3 Chebyshev polynomial that approximates the regularized inverse NUFFT. $\kappa(T(A;\mu) = A'A + \mu I) \approx 46675$, $\kappa(p_d(A;\mu)T(A;\mu)) \approx 1.9$, which implies that the approximate inverse $p_d(A;\mu)$ almost inverts the action of $T(A;\mu)$.

For inverse problems, (2) is more closely related to the propagating error, i.e., $e_k \propto (I - p_d(A)A'A)^k e_0$, for iterative solvers and is well-studied [1, 2]. The optimal polynomial is given by:

$$r_{d+1}(\lambda) = \frac{T_{d+1}(g(\lambda))}{T_{d+1}(g(0))}, \quad p_d(\lambda) = \frac{1 - r_{d+1}(\lambda)}{\lambda} \tag{3}$$

which is a direct consequence of the minimax theorem for Chebyshev polynomials [3, 4]. For a proof of the minimax theorem, check out these notes.

Remarks:

- Since $r_{d+1}(\lambda) \approx 1 - \lambda p_d(\lambda)$, solve for p_d.
- Dividing by $T_{d+1}(g(0))$ makes the $(d+1)$-degree polynomial monic, which allows invocation of the minimax theorem. Hence, $r_{d+1}(\lambda) = \frac{T_{d+1}(g(\lambda))}{T_{d+1}(g(0))}$ is the optimal polynomial.
- (1) addresses the absolute error of the approximate inverse whereas (2) addresses iterative error propagation.
- (2) has a closed-form solution w.r.t. the full Chebyshev series, whereas interpolating $\frac{1}{\lambda}$ requires truncating the Chebyshev series.
- Since Chebyshev polynomials satisfy $T_m(-x) = (-1)^m T_m(x)$ [4], $\frac{T_{d+1}(g(\lambda))}{T_{d+1}(g(0))} = \frac{T_{d+1}(-g(\lambda))}{T_{d+1}(-g(0))}$.

1.3 Evaluate the polynomial

Note that a polynomial can be refactored as follows:

$$a_0 + a_1 x + ... + a_d x^d = a_0 + x(a_1 + x(a_2 + ... + x(a_{d-1} + a_d x)...))$$

Clenshaw recurrence (Algorithm 1) for truncated Chebyshev polynomials allows evaluating a d-degree polynomial in $O(d)$ forward applications.

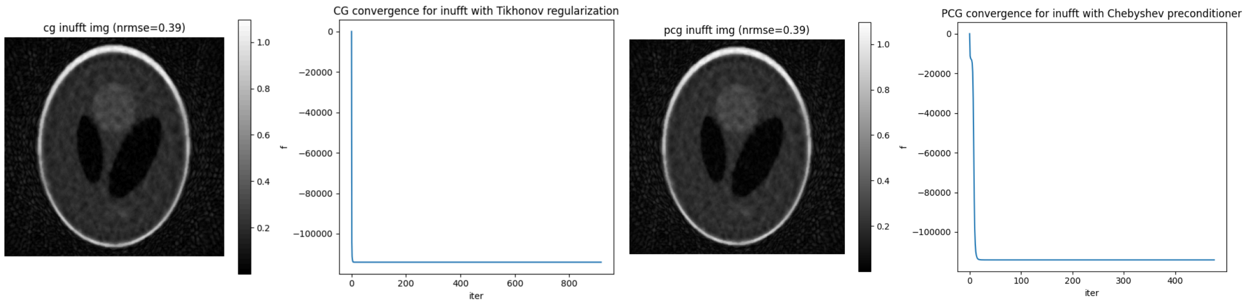
Algorithm 1 Truncated Chebyshev Clenshaw recurrence

Require: coefficients $\{c_1, ..., c_d\}$, operator $\bar{A}$
  $b_{k+1}, b_{k+2} = 0$
  for $k = d-1, ..., 0$ **do**
    $b_k = c_k I + 2\bar{A} b_{k+1} - b_{k+2}$
    $b_{k+2} = b_{k+1}$
    $b_{k+1} = b_k$
  end for
  $P = c_0 I + \bar{A} b_{k+1} - b_{k+2}$
  return P

2 Analysis

Theorem 2.1. *(convergence bound) For a linear operator $A \in \mathbb{S}^n_{++}$ and a d-degree Chebyshev polynomial given by* (3)*:*

$$||I - Ap_d(A)|| \leq 2\rho^{-(d+1)} \tag{4}$$



(a) (b)

Figure 2: Comparison of solving regularized INUFFT with CG (left) and Chebyshev PCG (right). PCG converges in fewer iterations.

where $\rho = \frac{\sqrt{\kappa}+1}{\sqrt{\kappa}-1}$.

Proof: Note that

$$
\begin{aligned}
||I - Ap_d(A)|| &= ||r_{d+1}(A)|| \\
&= \max_{\lambda_i} |r_{d+1}(\lambda_i)| \\
&\leq \frac{1}{|T_{d+1}(g(0))|} \\
&= \frac{1}{|T_{d+1}(\frac{\lambda_+ + \lambda_-}{\lambda_+ - \lambda_-})|}
\end{aligned}
$$

where the first equality comes from (3) and the second equality follows from the spectral function property and definition of the spectral norm. The inequality follows from the property $\max |T_n(x)| = 1$ [4]. The final equality follows from the property $T_m(-x) = (-1)^m T_m(x)$. Using the definition for a d-degree Chebyshev polynomial $T_d(x) = \frac{1}{2}[(x + \sqrt{x^2-1})^n + (x - \sqrt{x^2-1})^n]$, one can simplify and show that $T_{d+1}(\frac{\lambda_+ + \lambda_-}{\lambda_+ - \lambda_-}) = \frac{1}{2}(\rho^n + \rho^{-n})$. Since $\kappa \geq 1 \implies \rho^{-1} < 1$, it follows that:

$$
\frac{1}{|T_{d+1}(\frac{\lambda_+ + \lambda_-}{\lambda_+ - \lambda_-})|} = \frac{2}{\rho^{d+1} + \rho^{-(d+1)}} = \frac{2\rho^{-d+1}}{1 + \rho^{-2(d+1)}} \leq 2\rho^{-(d+1)}
$$

□

Corollary 2.1.1. *(approximate inverse error) Under the same conditions as 2.1:*

$$
||A^{-1} - p_d(A)|| \leq \frac{2}{\lambda_-}\rho^{-(d+1)} \tag{5}
$$

Proof: Follows directly from 2.1 and the sub-multiplicativity of the spectral norm:

$$
||A^{-1} - p_d(A)|| = ||A^{-1}(I - Ap_d(A))|| \leq ||A^{-1}|| \; ||I - Ap_d(A)|| \leq \frac{2}{\lambda_-}\rho^{-(d+1)}
$$

□

Theorem 2.2. *(degree complexity) With the same setting as 2.1, to achieve an ε-accurate approximate inverse requires degree d on the order of:*

$$
d = O(\sqrt{\kappa}\log(\frac{1}{\varepsilon})) \tag{6}
$$

Proof: To achieve ε accuracy requires:

$$
||I - Ap_d(A)|| \leq 2\rho^{-(d+1)} \leq \varepsilon.
$$

Solving for d on the right equality gives:

$$d \geq \frac{\log(\frac{2}{\varepsilon})}{\log \rho} - 1. \tag{7}$$

The worst-case scenario is that κ is large s.t. $\rho \approx 1 \implies \log \rho \approx 0 \implies d \to \infty$. To find the rate at which this grows, we proceed by analyzing the Taylor expansion of $\log \rho$. First, we decompose and rearrange $\log \rho$:

$$\log \rho = \log(\frac{\sqrt{\kappa}+1}{\sqrt{\kappa}-1}) = \log(\frac{1+\kappa^{-\frac{1}{2}}}{1-\kappa^{-\frac{1}{2}}}) = \log(1+\kappa^{-\frac{1}{2}}) - \log(1-\kappa^{-\frac{1}{2}})$$

Let $f(x) = \log(1+x) - \log(1-x)$. The terms have the following Maclaurin series:

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + ...$$
$$\log(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} - ...$$

Hence, $f(x) = 2(x + \frac{x^3}{3} + \frac{x^5}{5} + ...)$. Note that the Maclaurin series for $\log(x-1)$ does not exist since $\log(-1)$ so the rearrangement above is necessary. It follows then that:

$$\log \rho = \frac{2}{\kappa^{\frac{1}{2}}} + \frac{2}{3\kappa^{\frac{3}{2}}} + \frac{2}{5\kappa^{\frac{5}{2}}} + ...$$

Using the first-order Taylor approximation for large κ implies that $\log \rho \approx \frac{2}{\sqrt{\kappa}} + O(\kappa^{-\frac{3}{2}})$. Substituting into 7 yields the complexity 6. □

References

[1] S. S. Iyer, F. Ong, X. Cao, C. Liao, L. Daniel, J. I. Tamir, and K. Setsompop, "Polynomial preconditioners for regularized linear inverse problems," *SIAM Journal on Imaging Sciences*, vol. 17, no. 1, pp. 116–146, 2024.

[2] J. R. Shewchuk *et al.*, "An introduction to the conjugate gradient method without the agonizing pain," 1994.

[3] L. N. Trefethen, *Approximation theory and approximation practice, extended edition.* SIAM, 2019.

[4] J. C. Mason and D. C. Handscomb, *Chebyshev polynomials.* Chapman and Hall/CRC, 2002.