

2022-11-10

1 Approximation from a subspace

Our workhorse methods for solving large-scale systems involve two key ideas: *relaxation* to produce a sequence of ever-better approximations to a problem, and *approximation from a subspace* assumed to contain a good estimate to the solution (e.g. the subspace spanned by iterates of some relaxation method). Having dealt with the former, we now deal with the latter.

Suppose we wish to estimate the solution to a linear system $Ax^{(*)} = b$ by an approximate solution $\hat{x} \in \mathcal{V}$, where \mathcal{V} is some approximation subspace. How should we choose \hat{x} ? There are three standard answers:

- *Least squares*: Minimize $\|A\hat{x} - b\|_M^2$ for some M .
- *Optimization*: If A is SPD, minimize $\phi(x) = \frac{1}{2}x^T Ax - x^T b$ over \mathcal{V} .
- *Galerkin*: Choose $A\hat{x} - b \perp \mathcal{W}$ for some test space \mathcal{W} . In *Bubnov-Galerkin*, $\mathcal{W} = \mathcal{V}$; otherwise we have a *Petrov-Galerkin* method.

These three methods are the standard approaches used in all the methods we will consider. Of course, they are not the only possibilities. For example, we might choose \hat{x} to minimize the residual in some non-Euclidean norm, or we might more generally choose \hat{x} by optimizing some non-quadratic loss function. But these approaches lead to optimization problems that cannot be immediately solved by linear algebra methods.

The three approaches are closely connected in many ways:

- Suppose \hat{x} is the least squares solution. Then the normal equations give that $A\hat{x} - b \perp MA\mathcal{V}$; this is a (Petrov-)Galerkin condition.
- Similarly, suppose \hat{x} minimizes $\phi(x)$ over the space \mathcal{V} . Then for any $\delta x \in \mathcal{V}$ we must have

$$\delta\phi = \delta x^T (Ax - b) = 0,$$

i.e. $Ax - b \perp \mathcal{V}$. This is a (Bubnov-)Galerkin condition.

- If x is the least squares solution, then by definition we minimize

$$\frac{1}{2}\|Ax - b\|_M^2 = \frac{1}{2}x^T A^T M A x - x^T A^T M b + \frac{1}{2}b^T M b,$$

i.e. we have the optimization objective for the normal equation SPD system $A^T M A x - A^T M b = 0$, plus a constant.

- Note that if A is SPD, then we can express ϕ with respect to the A^{-1} norm as

$$\phi(x) = \frac{1}{2}\|Ax - b\|_{A^{-1}}^2 - \frac{1}{2}b^T A^{-1}b,$$

so choosing \hat{x} by minimizing $\phi(x)$ is equivalent to minimizing the A^{-1} norm of the residual.

- Alternately, write $\phi(x)$ as

$$\phi(x) = \frac{1}{2}\|x - A^{-1}b\|_A^2 - \frac{1}{2}b^T A^{-1}b,$$

and so choosing \hat{x} by minimizing $\phi(x)$ is also equivalent to minimizing the A norm of the error.

When deriving methods, it is frequently convenient to turn to one or the other of these characterizations. But for computation and analysis, we will generally turn to the Galerkin formalism.

In order for any of these methods to produce accurate results, we need two properties to hold:

- *Consistency*: Does the space contain a good approximation to x ?
- *Stability*: Will our scheme find something close to the best approximation possible from the space?

We leave the consistency and the choice of subspaces to later; for now, we deal with the problem of method stability.

2 Quasi-optimality

We quantify the stability of a subspace approximation method via a *quasi-optimality bound*:

$$\|x^* - \hat{x}\| \leq C \min_{v \in \mathcal{V}} \|x^* - v\|.$$

That is, the approximation \hat{x} is quasi-optimal if it has error within some factor C of the best error possible within the space.

To derive quasi-optimality results, it is useful to think of all of our methods as defining a *solution projector* that maps x^* to the approximate solution to $A\hat{x} = Ax^* = b$. From the (Petrov-)Galerkin perspective, if $W \in \mathbb{R}^{n \times k}$ and $V \in \mathbb{R}^{n \times k}$ are bases for the trial space \mathcal{W} and \mathcal{V} , respectively, then we have

$$\begin{aligned} W^T AV \hat{y} &= W^T b, & \hat{x} &= V \hat{y} \\ \hat{x} &= V(W^T AV)^{-1} W^T b \\ &= V(W^T AV)^{-1} W^T Ax^*. \\ &= \Pi x^*. \end{aligned}$$

The *error projector* $I - \Pi$ maps x^* to the error $\hat{x} - x^*$ in approximately solving $A\hat{x} \approx Ax^* = b$. There is no error iff x^* is actually in \mathcal{V} ; that is, \mathcal{V} is the null space of $I - \Pi$. Hence, if \tilde{x} is any vector in \mathcal{V} , then

$$\hat{e} = (I - \Pi)x = (I - \Pi)(x - \tilde{x}) = (I - \Pi)\tilde{e}.$$

Therefore we have

$$\|x - \hat{x}\| \leq \|I - \Pi\| \min_{\tilde{x} \in \mathcal{V}} \|x - \tilde{x}\|,$$

and a bound on $\|I - \Pi\|$ gives a quasi-optimality result.

For any operator norm, we have

$$\|I - \Pi\| \leq 1 + \|\Pi\| \leq 1 + \|V\| \|(W^T AV)^{-1}\| \|W^T A\|;$$

and in any Euclidean norm, if V and W are chosen to have orthonormal columns, then

$$\|I - \Pi\| \leq 1 + \|(W^T AV)^{-1}\| \|A\|.$$

If A is symmetric and positive definite and $V = W$, then the interlace theorem gives $\|(V^T AV)^{-1}\| \leq \|A^{-1}\|$, and the quasi-optimality constant is bounded by $1 + \kappa(A)$. In more general settings, though, we may have no guarantee that the projected matrix $W^T AV$ is far from singular, even if A itself is nonsingular. To guarantee boundedness of $(W^T AV)^{-1}$ *a priori* requires a compatibility condition relating \mathcal{W} , \mathcal{V} , and A ; such a condition is sometimes called the *LBB* condition (for Ladyzhenskaya-Babuška-Brezzi) or the *inf-sup* condition, so named because (as we have discussed previously)

$$\sigma_{\min}(W^T AV) = \inf_{w \in \mathcal{W}} \sup_{v \in \mathcal{V}} \frac{w^T Av}{\|w\| \|v\|}.$$

The LBB condition plays an important role when Galerkin methods are used to solve large-scale PDE problems, since there it is easy to choose the spaces \mathcal{V} and \mathcal{W} in a way that leads to very bad conditioning. But for iterative solvers of the type we discuss in this course (Krylov subspace solvers), such pathologies are a more rare occurrence. In this setting, we may prefer to monitor $\|(W^T AV)^{-1}\|$ directly as we go along, and to simply increase the dimension of the space if we ever run into trouble.

3 Krylov subspaces

The *Krylov subspace* of dimension k generated by $A \in \mathbb{R}^{n \times n}$ and $b \in \mathbb{R}^n$ is

$$\mathcal{K}_k(A, b) = \text{span}\{b, Ab, \dots, A^{k-1}b\} = \{p(A)b : p \in \mathcal{P}_{k-1}\}.$$

Krylov subspaces are a natural choice for subspace-based methods for approximate linear solves, for two reasons:

- If all you are allowed to do with A is compute matrix-vector products, and the only vector at hand is b , what else would you do?
- The Krylov subspaces have excellent approximation properties.

Krylov subspaces have several properties that are worthy of comment. Because the vectors $A^j b$ are proportional to the vectors obtained in power iteration, one might reasonably (and correctly) assume that the space quickly contains good approximations to the eigenvectors associated with the largest magnitude eigenvalues. Krylov subspaces are also *shift-invariant*, i.e. for any σ

$$\mathcal{K}_k(A - \sigma I, b) = \mathcal{K}_k(A, b).$$

By choosing different shifts, we can see that the Krylov subspaces tend to quickly contain not only good approximations to the eigenvector associated with the largest magnitude eigenvalue, but to all “extremal” eigenvalues.

Most arguments about the approximation properties of Krylov subspaces derive from the characterization of the space as all vectors $p(A)b$ where $p \in \mathcal{P}_{k-1}$ and from the spectral mapping theorem, which says that if $A = V\Lambda V^{-1}$ then $p(A) = Vp(\Lambda)V^{-1}$. Hence, the distance between an arbitrary vector (say d) and the Krylov subspace is

$$\min_{p \in \mathcal{P}_{k-1}} \|V [p(\Lambda)V^{-1}b - V^{-1}d]\|.$$

As a specific example, suppose that we want to choose \hat{x} in a Krylov subspace in order to minimize the residual $A\hat{x} - b$. Writing $\hat{x} = p(A)b$, we have that we want to minimize

$$\|[Ap(A) - I]b\| = \|q(A)b\|$$

where $q(z)$ is a polynomial of degree at most k such that $q(1) = 1$. The best possible residual in this case is bounded by

$$\|q(A)b\| \leq \kappa(V)\|q(\Lambda)\| \|b\|,$$

and so the relative residual can be bounded in terms of the condition number of V and the minimum value that can bound q on the spectrum of A subject to the constraint that $q(0) = 1$.

4 Chebyshev polynomials

Suppose now that A is symmetric positive definite, and we seek to minimize $\|q(A)b\| \leq \|q(\Lambda)\| \|b\|$. Controlling $q(z)$ on all the eigenvalues is a pain, but it turns out to be simple to instead bound $q(z)$ over some interval $[\alpha_1, \alpha_n]$. The polynomial we want is the *scaled and shifted Chebyshev polynomial*

$$q_m(z) = \frac{T_m((z - \bar{\alpha})/\rho)}{T_m(-\bar{\alpha}/\rho)}$$

where $\bar{\alpha} = (\alpha_n + \alpha_1)/2$ and $\rho = (\alpha_n - \alpha_1)/2$.

The Chebyshev polynomials T_m are defined by the recurrence

$$\begin{aligned} T_0(x) &= 1 \\ T_1(x) &= x \\ T_{m+1}(x) &= 2xT_m(x) - T_{m-1}(x), \quad m \geq 1. \end{aligned}$$

The Chebyshev polynomials have a number of remarkable properties, but perhaps the most relevant in this setting is that

$$T_m(x) = \begin{cases} \cos(m \cos^{-1}(x)), & |x| \leq 1, \\ \cosh(m \cosh^{-1}(x)), & |x| \geq 1. \end{cases}$$

Thus, $T_m(x)$ oscillates between ± 1 on the interval $[-1, 1]$, and then grows very quickly outside that interval. In particular,

$$T_m(1 + \epsilon) \geq \frac{1}{2}(1 + m\sqrt{2\epsilon}).$$

Thus, we have that on $[\alpha, \alpha_n]$, $|q_m| \leq \frac{2}{1+m\sqrt{2\epsilon}}$ where

$$\epsilon = \bar{\alpha}/\rho - 1 = \frac{2\alpha_1}{\alpha_n - \alpha_1} = 2(\kappa(A) - 1)^{-1},$$

and hence

$$\begin{aligned} |q_m(z)| &\leq \frac{2}{1 + 2m/\sqrt{\kappa(A) - 1}} \\ &= 2 \left(1 - \frac{2m}{\sqrt{\kappa(A) - 1}} \right) + O\left(\frac{m^2}{\kappa(A) - 1}\right). \end{aligned}$$

Hence, we expect to reduce the optimal residual in this case by at least about $2/\sqrt{\kappa(A) - 1}$ at each step.

5 Chebyshev: Uses and Limitations

We previously sketched out an approach for analyzing the convergence of methods based on Krylov subspaces:

1. Characterize the Krylov subspace of interest in terms of polynomials, i.e. $\mathcal{K}_k(A, b) = \{p(A)b : p \in \mathcal{P}_{k-1}\}$.
2. For $\hat{x} = p(A)b$, write an associated error (or residual) in terms of a related polynomial in A .
3. Phrase the problem of minimizing the error, residual, etc. in terms of minimizing a polynomial $q(z)$ on the spectrum of A (call this $\Lambda(A)$). The polynomial q must generally satisfy some side constraints that prevent the zero polynomial from being a valid solution.
4. Let $\Lambda(A) \subset \Omega$, and write

$$\max_{\lambda \in \Lambda(A)} |q(\lambda)| \leq \max_{z \in \Omega} |q(z)|.$$

The set Ω should be simpler to work with than the set of eigenvalues. The simplest case is when A is symmetric positive definite and $\Omega = [\lambda_1, \lambda_n]$.

5. The optimization problem can usually be phrased in terms of special polynomial families. The simplest case, when Ω is just an interval, usually leads to an analysis via Chebyshev polynomials.

The analysis sketched above is the basis for the convergence analysis of the Chebyshev semi-iteration, the conjugate gradient method, and (with various twists) several other Krylov subspace methods.

The advantage of this type of analysis is that it leads to convergence bounds in terms of some relatively simple property of the matrix, such as the condition number. The disadvantage is that the approximation of the spectral set $\Lambda(A)$ by a bounding region Ω can lead to rather pessimistic bounds. In practice, the extent to which we are able to find good solutions in a Krylov subspace often depends on the “clumpiness” of the eigenvalues. Unfortunately, this “clumpiness” is rather difficult to reason about a priori! Thus, the right way to evaluate the convergence of Krylov methods in practice is usually to try them out, plot the convergence curves, and see what happens.