CS 6210: Homework 4

Instructor: Anil Damle
Due: November 6, 2024

Policies

You may discuss the homework problems freely with other students, but please refrain from looking at their code or writeups (or sharing your own). Ultimately, you must implement your own code and write up your own solution to be turned in. Your solution, including plots and requested output from your code should be typeset and submitted via the Gradescope as a pdf file. Additionally, please submit any code written for the assignment. This can be done by either including it in your solution as an appendix, or uploading it as a zip file to the separate Gradescope assignment.

Question 1:

We are going to consider some details about how Krylov methods behave in specific situations.

(a) Given a symmetric positive definite matrix A and vector b, prove that if the Lanczos process breaks down at some point (*i.e.* $\beta_k = 0$ using the notation from class and Trefethen and Bau) then the subspace $\mathcal{K}_k(A, b)$ contains a solution to the linear system $Ax = b$. In principle we might be worried that if $\beta_k = 0$ things have gone horribly wrong since we cannot construct the next vector in our orthonormal basis. However, this result shows that in this context everything has actually gone remarkably well.

(b) Still assuming that A is symmetric positive definite, consider using CG to solve the linear system $Ax = b$ and let y satisfy $T_k y = \|b\| e_1$ (i.e., y is the solution to the CG subproblem that defines our iterate in $\mathcal{K}_k(A, b)$). Prove that the CG iterate $x^{(k)} = V_k y$ satisfies

$$\|b - Ax^{(k)}\|_2 = \beta_k (e_k^T y).$$

In addition, prove that $b - Ax^{(k)} \perp \mathcal{K}_k(A, b)$ (i.e., that the residual is orthogonal to the current Krylov subspace).

QUESTION 2:

Here, we will consider the applicability of Krylov methods to solving a set of closely related linear systems. Specifically, we are given a real symmetric $n \times n$ matrix A and a set of M real numbers $\{\sigma_i\}_{i=1}^{M}$ (you may assume none of the σ_i are eigenvalues of A), and we want to solve the set of M linear systems

$$(A - \sigma_i I)\, x_i = b$$

for $\{x_i\}_{i=1}^{M}$. You may assume we do not have any reason to use an initial guess besides $\vec{0}$ for all of the given linear systems.

Devise a Krylov subspace based iterative method to "simultaneously" solve this collection of linear systems in the sense that you construct M sequences of iterates each with the property that $x_i^{(k)} \to (A - \sigma_i)^{-1}\, b$ as $k \to \infty$. In addition, your algorithm must satisfy the following properties:

- Use no more than one matrix vector product with A at each iteration. A single iteration constitutes computing $x_i^{(k)}$ for $i = 1, \ldots, M$.
- Converge (in exact arithmetic) for any σ_i that is not an eigenvalue of A and converge in at most ℓ iterations for every i if A has ℓ distinct eigenvalues.
- Have a storage cost that is naïvely $\mathcal{O}(Mnk)$ and computational complexity per iteration that is naïvely $T_{mult}(A) + \mathcal{O}(Mnk) + \mathcal{O}(Mk^3)$, but can be improved to $\mathcal{O}(Mn) + \mathcal{O}(Mk)$ and $T_{mult}(A) + \mathcal{O}(Mn)$ respectively. You do not have to work out all the details on the improvement, but you do have to make a convincing argument that such an improvement is possible.

Given this problem and set of requirements address the following:

(a) State your algorithm for addressing the above problem and prove why it satisfies the desired criteria. Be sure to both clearly articulate your final algorithm and provide the desired proofs.

(b) Let's say we are given a set of non-singular real symmetric preconditioners $M_i^{-1} \approx (A - \sigma_i)^{-1}$ and their Cholesky factorizations $M_i^{-1} = L_i L_i^T$. Do you think that these preconditioners can be incorporated into your algorithm without adversely impacting its computational benefits (i.e., could you devise an algorithm that is faster than solving the M problems independently)? Why or why not?