CS 6210: Homework 6

Instructor: Anil Damle
Due: December 9, 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:

Assume that we have run the Golub-Kahan bidiagonalization algorithm[1] on $A \in \mathbb{R}^{m\times n}$ with $m \geq n$ for k steps starting from and $v_0 \in \mathbb{R}^n$. This yields the matrices $U_k \in \mathbb{R}^{m\times k}$ and $V_k \in \mathbb{R}^{n\times k}$ with orthonormal columns, and the upper bidiagonal matrix $B_k \in \mathbb{R}^{k\times k}$. These matrices satisfy the relation

$$AV_k = U_kB_k \quad \text{and} \quad A^TU_k = V_kB_k^T + p_ke_k^T,$$

where $p_k = A^Tu_k - \alpha_kv_k$ and α_k is the (k,k) element of B_k. If $B_k = F_k\Gamma_kG_k^T$ then our estimates for singular values and the corresponding vectors are $\hat{\sigma}_i = \gamma_i$, $\hat{u}_i = U_kF_ke_i$, and $\hat{v}_i = V_kG_ke_i$; as usual we note that some of these may be significantly better estimates of singular values/vectors than others.

(a) Show that the right singular vector estimates $\hat{v}_i$ and associated singular value estimates $\hat{\sigma}_i$ satisfy

$$A^TA\hat{v}_i = \hat{\sigma}_i^2\hat{v}_i + \hat{\sigma}_i(e_k^TF_ke_i)p_k.$$

(b) Given what we know about perturbation theory for symmetric matrices and how the residual characterizes the quality of eigenvalue/vector estimates, how should we interpret the result from part (a) when thinking about convergence of specific singular value/vector estimates?

[1] This is written for the more general case with $m \geq n$, but it has no impact on the problem.

Question 2 (ungraded):

The following question will not be graded and you do not have to submit a solution. Nevertheless, it is a result we saw in class and it would be good to think a little bit about how you would prove a result of this type.

Assume we are using the Lanczos process to compute some eigenvalues of a real symmetric matrix A and build the Krylov space starting with vector z_0. Let A have eigenvalues $\lambda_1, \ldots, \lambda_n$ satisfying $\lambda_1 \geq \cdots \geq \lambda_n$ and associated eigenvectors $v_1, \ldots, v_n$.

(a) After k steps of the Lanczos process let T_k be the tridiagonal matrix generated by the process, prove that $\theta_1 = \lambda_1(T_k)$ (the largest magnitude eigenvalue of T_k) satisfies

$$\lambda_1 \geq \theta_1 \geq \lambda_1 - (\lambda_1 - \lambda_n) \left(\frac{\tan \phi_1}{c_{k-1}(1 + 2\rho_1)} \right)^2$$

where $\cos(\phi_1) = |z_0^T v_1|$, $\rho_1 = \frac{\lambda_1 - \lambda_2}{\lambda_2 - \lambda_n}$, and $c_{k-1}(x)$ is the Chebyshev polynomial of degree $k-1$.

(b) Compare and contrast this convergence result with that for the power method (assuming the eigenvalues are such that the power method would converge to λ_1).