CS 6210: Homework 2

Instructor: Anil Damle
Due: September 25, 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:

For square or rectangular A with $m \geq n$ and full column rank implement the computation of a (reduced) QR factorization via Householder triangularization. This now gives you a means for solving linear systems and least squares problems.

(a) For the case $m = n$ clearly illustrate that your implementation scales as $\mathcal{O}(n^3)$. Then, fix n and allow m to grow; show your implementation scales linearly with m in this regime.

(b) We will now build a set of square problems that are increasingly ill-conditioned for fixed $n = 100$ and explore the behavior of our QR factorization routine. Specifically, generate a random orthogonal matrix Q (e.g., to get Q apply a built in QR factorization routine to a matrix with i.i.d. $\mathcal{N}(0,1)$ entries) and let R_s be the upper-triangular matrix parametrized by s and defined as

$$R_s = \begin{bmatrix} 1 & & & & \\ & s & & & \\ & & s^2 & & \\ & & & \ddots & \\ & & & & s^{n-1} \end{bmatrix} \begin{bmatrix} 1 & -c & -c & \cdots & -c \\ & 1 & -c & \cdots & -c \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & -c \\ & & & & 1 \end{bmatrix},$$

where $s^2 + c^2 = 1$ and $0 < s \leq 1$. R_s is known as the Kahan matrix and becomes extremely ill-conditioned as $s \to 0$. let $A_s = QR_s$. As you vary s to increase the condition number of A_s use your code to compute a QR factorization of A and call the resulting matrices $\tilde{Q}$ and $\tilde{R}$. Compute the relative forward errors $\|Q - \tilde{Q}\|_2$ and $\|R_s - \tilde{R}\|_2 / \|R_s\|_2$ and plot the results (recall that a QR factorization is not unique, so be careful in how you do this). What happens as the condition number increases? Also plot the backward errors $\|A_s - \tilde{Q}\tilde{R}\|_2 / \|A_s\|_2$, what behavior do you observe? How do the two errors compare?

(c) Using the same set of problems, explore the relationship between the condition number of the problem and your ability to accurately solve $A_s x = b$ using your QR factorization routine. Please form b randomly (i.e., not as $A_s x$ for some fixed x) and report the relative residual $\|A_s\tilde{x} - b\|_2 / (\|A_s\|_2 \|\tilde{x}\|_2)$

Question 2:

Say we would like to compute a QR factorization of A^k for some real, square A and (potentially large) integer k.

(a) Even if it is feasible to simply construct A^k, is it be good idea numerically to compute its QR factorization directly? Justify your response.

(b) Develop a scheme to compute a QR factorization of A^k via a sequence of k QR factorizations of matrices whose condition number is never larger than that of A itself. As long as you adhere to this constraint, for the purposes of this problem you may assume that computing matrix matrix products with the resulting Q and/or R matrices is fine numerically.

Question 3:

(a) Given any $A \in \mathbb{R}^{m \times n}$ and $\lambda > 0$ prove that the solution to the regularized least squares problem

$$\min_x \|Ax - b\|_2^2 + \lambda \|x\|_2^2 \tag{1}$$

is unique.

(b) In practice, we often want to solve (1) for many values of λ. Say we are in a situation where $m \gg n$, A is full column rank (though potentially ill-conditioned), and we would like to avoid the use of the normal equations. Assume you are given the reduced QR factorization $A = QR$. Devise a scheme to solve (1) for T values of λ (i.e., $\lambda_1, \ldots, \lambda_T$ with $\lambda_i > 0$) that costs at most $\mathcal{O}(mn + Tn^3)$. Critically, note that the dependence on T only scales with n and not m.