

Numerical Computing :: Project Eight

The purpose of this homework is to see the relationship between the condition number and the numerical error when solving linear least-squares problems.

First, implement the following methods for least-squares, which you'll use in the next exercise.

1. Method of Normal Equations (uses the Cholesky factorization)
2. Method based on the Thin QR factorization

Next, load the given matrix (download from Canvas) into memory. Call the matrix A ,

$$A = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}, \quad (1)$$

where $a_i \in \mathbb{R}^m$ is the i th column of A . Define the matrices A_1, \dots, A_n as:

$$A_k = \begin{bmatrix} a_1 & \cdots & a_k \end{bmatrix}, \quad k = 1, \dots, n. \quad (2)$$

That is, A_k contains the first k columns of the matrix A (that you loaded into memory).

Now, generate the error data that you'll analyze. For k from $k_{\min} = 40$ to $k_{\max} = 65$:

1. Report the size, rank, and condition number of A_k .
2. Generate 100 random vectors $b_i \in \mathbb{R}^m$. For each b_i ,
 - (a) Use the built-in equation solver (`numpy.linalg.lstsq`) to compute the least-squares minimizers given A_k and b_i . Call this vector x_{true} , because we're just gonna trust the software on this one.
 - (b) Use your Normal Equation solver to compute the least-squares minimizer, x_{NE} . Compute the relative error with x_{true} .
 - (c) Use your QR solver to compute the least-squares minimizer, x_{QR} . Compute the relative error with x_{true} .
3. For each of QR and Normal Equations, compute the average error over all the b_i .

Make two plots on a `semilogy` scale:

- the average error versus k (how many columns in the matrix) for both QR and the Normal Equations,
- the condition number of A_k versus k .

Now tell me what's going on. More specifically:

1. What is the relationship between the error using QR versus the Normal Equations?
2. What is the relationship between the errors and the condition number of A_k ?
3. Suppose your matrix A is ill-conditioned. Which method is more favorable?