Numerical analysis: Project 01

Students: Cedric Orton-Urbina (ceo29), Aditya Syam (as2839), Pratyush Sudhakar (ps2245)
Due Date: 11th March, 2024

Question 1.

Show that if z solves

$$\min_z \|A(:, 1:k)z - b\|_2^2$$

then

$$\|A \begin{bmatrix} z \\ 0 \end{bmatrix} - b\|_2 \leq \|R_{22}\|_2 \|x_{\text{LS}}\|_2 + \|b - Ax_{\text{LS}}\|_2,$$

where x_{LS} solves

$$\min_x \|Ax - b\|_2^2.$$

In other words, the residual from the sparse solution cannot be much larger than the optimal residual.

Answer 1.

To begin, let $A(:, 1:k) = A_k$ and $A(:, k+1:n) = A_m$ for simplicity. We begin identifying the meaning behind each term in the above question.

1. z is a vector containing the coefficients that best approximate b using the columns of A_k.
2. $A \begin{bmatrix} z \\ 0 \end{bmatrix} - b$ is the spare solution z extended to the full column space of A
3. $\|b - Ax_{\text{LS}}\|_2$ is the norm of the minimum achievable residual of the least squares problem in this question.
4. R_{22} represents the portion of A not involved in the initial k column solution.

We begin by noting that the residual $\|A \begin{bmatrix} z \\ 0 \end{bmatrix} - b\|_2$ contains contributions from two distinct parts of A. This first contribution is attributed to the first k columns of A, which are scaled by z. The residual $\|b - A_k z\|_2$ is bounded below by the minimal residual $r_{min} = \|b - Ax_{\text{LS}}\|_2$. The case where these two residuals would be equal corresponds to an A where all other columns are 0 aside from the first k. In other words, these residuals are equal when $A_m = 0$, but otherwise,

$$\|b - A_k z\|_2 \geq r_{min}$$

We now rewrite the original inequality as follows:

$$\|A \begin{bmatrix} z \\ 0 \end{bmatrix} - b\|_2 - \|b - Ax_{\text{LS}}\|_2 \leq \|R_{22}\|_2 \|x_{\text{LS}}\|_2$$

Therefore, we must quantify the increased accuracy in our least squares solution when A_m is included in our minimization. q

Question 2.

Show that there exists a matrix B such that

$$\|A - A(:, 1:k)B\|_2 = \|R_{22}\|_2$$

and

$$\|A(:, 1:k) - A(:, 1:k)B(:, 1:k)\|_2 = 0.$$

The second condition ensures that k columns of A are exactly represented by the rank-k factorization $A(:, 1:k)B$.

Answer 2.

Question 3.

Show that

$$\|R_{22}\|_2 \leq \sigma_1(A)$$

and that for any m, n, and k with $m \geq n$ and $k < \min(m, n)$ there exists a matrix A such that

$$\|R_{22}\|_2 = \sigma_1(A).$$

This shows that while R_{22} is bounded in size by the scale of A, there are examples where it is that large.

Answer 3.

We begin by noting that $\sigma_1(A)$ is the largest singular value of A, and is also equal to the 2-norm of A.

Since R_{22} is a partition of the QR factorization of A, and QR factorization does not increase any spectral properties of A (magnitude of singular values), $\|R_{22}\|_2$ can at most be σ_1. This would be the case if R_{22} contained the singular value σ_1, otherwise, $\|R_{22}\|_2$ will be less than σ_1. Thus,

$$\|R_{22}\|_2 \leq \sigma_1(A)$$

To show there exists A such that $\|R_{22}\|_2 = \sigma_1(A)$, consider a matrix $A \in \mathbb{R}^{m\times n}$ with entries only along the diagonal (A_{ii} for $i \leq min(m,n)$). We then place the largest of these entries (σ_1 in the case of a diagonal matrix) in the $(k+1)^{th}$ diagonal of A. Since the QR decomposition of a diagonal matrix A with excess rows or columns is just the identity matrix times A, we know that the largest singular value σ_1 will be in the R_{22} partition of the QR factorization of A. The 2-norm of R_{22} will be equal to it's largest singular value, which in this case, is the same largest singular value of A. Thus, there exists $A \in \mathbb{R}^{m\times n}$ such that

$$\|R_{22}\|_2 \ = \sigma_1(A)$$

Question 4.

$$A \begin{bmatrix} \Pi_1 & \Pi_2 \end{bmatrix} = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} \\ & R_{22} \end{bmatrix} \tag{1}$$

Show that if we have a factorization of the form eq. (1) then there exists a matrix B such that

$$\|A - A(:,\mathcal{C})B\|_2 = \|R_{22}\|_2$$

and

$$\|A(:,\mathcal{C}) - A(:,\mathcal{C})B(:,\mathcal{C})\|_2 = 0.$$

Here we are just rewriting the earlier result using the permutations. A similar thing can be done for the least-squares problem.

Answer 4.

Question 5.

The greedy heuristic we will follow is to remove the largest column from R_{22} at each step. Specifically, lets say we have completed k steps of our process have have the partial factorization

$$Q^{(k)} \cdots Q^{(1)} A \Pi^{(1)} \cdots \Pi^{(k)} = \begin{bmatrix} R_{11} & R_{12} \\ & M \end{bmatrix}, \tag{2}$$

where $R_{11} \in \mathbb{R}^{k\times k}$ is upper triangular and $M \in \mathbb{R}^{(m-k)\times(n-k)}$ is dense (because we have not yet reduced it to upper triangular form).

Show that if we continue the factorization in eq. (2) without any additional permutations then $\|R_{22}\|_2 = \|M\|_2$.

Answer 5.

Question 6.

Show that Π_1 in eq. (1) only depends on $\Pi^{(1)}$ through $\Pi^{(k)}$. In other words, if we just want/need $\mathcal{C}$ we can stop the factorization after k steps. We can also recover the matrix B in our low-rank factorization without proceeding further.

Answer 6.