Rank Revealing QR Factorization
APPM 4600 Project
University of Colorado Boulder
Department of Applied Mathematics

1 Project Summary

In class, we talked about the QR factorization. In many applications, it useful to create a low rank factorization which approximates a matrix to a user prescribed accuracy. This is useful, for example, when the low rank factorization provides a reasonable approximation to the original matrix, with a fraction of the cost when storing or computing with it. One of the most popular techniques for doing this is the rank revealing QR factorization. In fact, it is becoming a driving force in machine learning and reduced order modeling. In this project, you will learn about QR factorizations with pivoting. For the independent directions, you can consider a randomized variant of the singular value decomposition (SVD), and/or apply your low rank approximation methods (QR or SVD) to a real-world problem where the data can be compressed/approximated by such methods.

2 Project Background

There are several rank revealing factorizations that are used in practice, including the QR decomposition and the SVD. We will first define a rank revealing factorization (RRF) in a general way. The RRF of $A \in \mathbb{R}^{m\times n}$ is a factorization

$$A = XDY^T, \quad X \in \mathbb{R}^{m\times p}, \quad D \in \mathbb{R}^{p\times p}, \quad Y \in \mathbb{R}^{n\times p},$$

where $p < \min\{m, n\}$, D is diagonal and nonsingular, and X, Y are well conditioned.

The existence of an RRF is established by the existence of just one such decomposition, like the SVD ($A = U\Sigma V^T$). In that case, X, Y are U and V in the SVD, and D takes the place of Σ. Notionally, the decompositions concentrate the ill-conditioning of A into the diagonal matrix D.

If a matrix is *numerically* rank deficient, then for a given $\varepsilon > 0$ (typically small), there is a $k < n$ for which $\sigma_k > \varepsilon$. In words, the numerical rank of a matrix for a given $\varepsilon > 0$ is the number of entries on the diagonal of D which are greater than ε. As such, RRFs provide a way to compute the numerical rank of a given matrix A, and tells us which columns and rows of the decomposition matrices can be discarded to give a low rank approximation of A - one with the same numerical rank, based on the provided accuracy parameter ε.

The QR factorization is an RRF with $X = I$, $D = \text{diag}(R)$ and $Y^T = D^{-1}R$. However, it is flawed in this basic form given our definition of a RRF. To see this, consider the matrix

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

and answer the following questions.

1. What is the QR factorization of A (by inspection)?

2. What is the QR factorization of A using our definition of RRF?

You should have found that there is a discrepancy between the QR factorization of A obtained by inspection compared with our definition of RRF. The trick here is introduce pivoting into the QR factorization.

2.1 QR factorization with column pivoting

A QR factorization with column pivoting computes for a matrix $A \in \mathbb{R}^{m\times n}$ the factorization

$$AP = QR,$$

where P is a permutation matrix, $Q \in \mathbb{R}^{m\times n}$ has orthonormal columns, and $R \in \mathbb{R}^{n\times n}$ is upper triangular and satisfies

$$|r_{11}| \geq |r_{22}| \geq \cdots |r_{nn}|.$$

These inequalities are simply to say that $\text{diag}(R)$ is in decreasing order. For a given $\varepsilon > 0$, if $|r_{kk}| \geq \varepsilon \geq |r_{k+1,k+1}|$, then we can write R in block form as

$$\begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{bmatrix},$$

where $R_{11} \in \mathbb{R}^{k\times k}$, $R_{12} \in \mathbb{R}^{k\times(n-k)}$, and $R_{22} \in \mathbb{R}^{(n-k)\times(n-k)}$. If we split the Q matrix in block form to make sense with the block form of R (in terms of size), we have

$$AP = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{bmatrix} = Q_1 \begin{bmatrix} R_{11} & R_{12} \end{bmatrix} + \begin{bmatrix} 0 & Q_2 R_{22} \end{bmatrix}.$$

1. Consider the following inequality (we state without proof) between the 2-norm and Frobenius norm of the lower right block R_{22}:

   $$||R_{22}||_2 \leq ||R_{22}||_F \leq 2^{-1/2}(n-k+1)\varepsilon.$$

   Notionally, this inequality means that R_{22} is numerically negligible. Use the inequality to show that Q_1 gives a $\mathcal{O}(\varepsilon)$ approximation to the range of A.

2. What does the order r of R_{11} reveal about the numerical rank of A?

3. What are the first r columns of Q in terms of the range of A? How does this compare with when you do not use pivoting?

4. What are the first r columns of AP in terms of the columns of A?

5. Randomly generate rank deficient matrices for a range of rank deficiencies, and use an implementation of QR factorization with column pivoting to see if you can detect the numerical rank through the decomposition.

6. How does the 2-norm distance between A and the low rank approximation compare with the tolerance parameter ε (you can use a range for this parameter to generate rank deficient matrices)?

3 Software expectations

You may use software for the independent portion of this project, but must implement your own routines for the exercises. You may use built-in functions in `MATLAB` or `Python` (for computing norms, inverses, qr, etc.).

4 Independent Directions

Possible extensions to this project include, but are not limited to:

1. Come up with a real-world problem which benefits from low rank approximations to data. Compare how the approximation from QR factorization behaves between pivoting and not pivoting. Application areas of interest to some include image processing and music compression, amongst many others.

2. We know the SVD is another RRF, but is often expensive to either compute or store. A randomized version due to Martinsson, Halko and Tropp was developed to address this issue, where the RRF can be computed and stored with matrices that are small relative to A. We break the algorithm into two stages. In stage 1, we compute an approximate basis for the range of A, where $A \approx QQ^T A$, and Q has l orthonormal columns ($k \leq l \leq n$) that captures the action of A. In stage 2, we use this approximate range matrix Q to approximate the SVD of A using smaller matrices.[1] Implement this initial version of Randomized SVD (RSVD), and test its performance on either a constructed or real-world problem.

[1] A helpful tutorial can be found at `https://gregorygundersen.com/blog/2019/01/17/randomized-svd/`.

5 Helpful Sources

1. Gu, Eisenstat, Efficient algorithms for computing a strong rank-revealing QR factorization, SIAM J. Sci. Comput., 1996.

2. Chandrasekaran, Ipsen, On Rank-Revealing Factorizations, SIAM JMAA, 1994.

3. Halko, Martinsson, Tropp, Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions, SIAM Review, 2011.